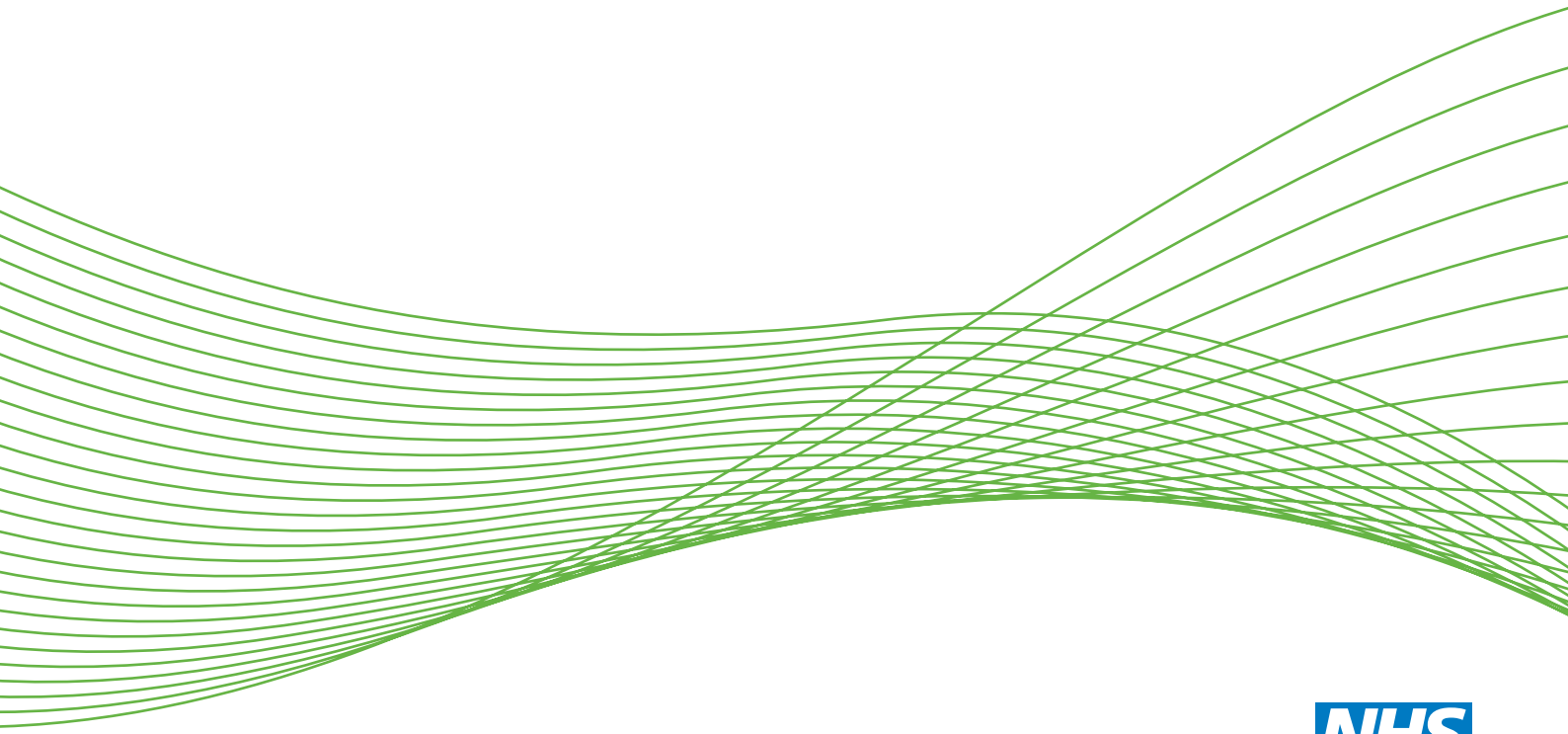
HEALTH TECHNOLOGY ASSESSMENT

VOLUME 17 ISSUE 60 DECEMBER 2013
ISSN 1366-5278

Sentinel lymph node status in vulval cancer: systematic reviews of test accuracy and decision-analytic model-based economic evaluation

C Meads, A Sutton, S Małysiak, M Kowalska, A Zapalska, E Rogozińska, P Baldwin, A Rosenthal, R Ganesan, E Borowiack, P Barton, T Roberts, S Sundar and K Khan



NHS
National Institute for Health Research

DOI 10.3310/hta17600

Sentinel lymph node status in vulval cancer: systematic reviews of test accuracy and decision-analytic model-based economic evaluation

C Meads,[1,2]* A Sutton,[3] S Małysiak,[4] M Kowalska,[4] A Zapalska,[4] E Rogozińska,[1] P Baldwin,[5] A Rosenthal,[6] R Ganesan,[7] E Borowiack,[4] P Barton,[3] T Roberts,[3] S Sundar[7] and K Khan[1]

[1]Centre for Primary Care and Public Health, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK
[2]Health Economics Research Group (HERG), Brunel University, Middlesex, UK
[3]Unit of Health Economics, University of Birmingham, Birmingham, UK
[4]Arcana Institute, Krakow, Poland
[5]Departments of Gynaecology and Gynaecological Oncology, Addenbrooke's Hospital NHS Trust, Cambridge, UK
[6]Department of Gynaecology, Barts Health NHS Trust, London, UK
[7]Pan Birmingham Gynaecological Cancer Centre, Birmingham Women's Hospital NHS Foundation Trust, Birmingham, UK

*Corresponding author

Declared competing interests of authors: none

Published December 2013
DOI: 10.3310/hta17600

This report should be referenced as follows:

Meads C, Sutton A, Małysiak S, Kowalska M, Zapalska A, Rogozińska E, *et al.* Sentinel lymph node status in vulval cancer: systematic reviews of test accuracy and decision-analytic model-based economic evaluation. *Health Technol Assess* 2013;**17**(60).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (SciSearch®) and *Current Contents®*/ Clinical Medicine.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Five-year impact factor: 5.804

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index and is assessed for inclusion in the Database of Abstracts of Reviews of Effects.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: www.hta.ac.uk/

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 09/112/03. The contractual start date was in November 2010. The draft report began editorial review in April 2012 and was accepted for publication in May 2013. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Abstract

Sentinel lymph node status in vulval cancer: systematic reviews of test accuracy and decision-analytic model-based economic evaluation

C Meads,[1,2]* A Sutton,[3] S Małysiak,[4] M Kowalska,[4] A Zapalska,[4] E Rogozińska,[1] P Baldwin,[5] A Rosenthal,[6] R Ganesan,[7] E Borowiack,[4] P Barton,[3] T Roberts,[3] S Sundar[7] and K Khan[1]

[1]Centre for Primary Care and Public Health, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK
[2]Health Economics Research Group (HERG), Brunel University, Middlesex, UK
[3]Unit of Health Economics, University of Birmingham, Birmingham, UK
[4]Arcana Institute, Krakow, Poland
[5]Departments of Gynaecology and Gynaecological Oncology, Addenbrooke's Hospital NHS Trust, Cambridge, UK
[6]Department of Gynaecology, Barts Health NHS Trust, London, UK
[7]Pan Birmingham Gynaecological Cancer Centre, Birmingham Women's Hospital NHS Foundation Trust, Birmingham, UK

*Corresponding author

Background: Vulval cancer causes 3–5% of all gynaecological malignancies and requires surgical removal and inguinofemoral lymphadenectomy (IFL). Complications affect > 50% of patients, including groin wound infection, lymphoedema and cellulitis. A sentinel lymph node (SLN) is the first groin node with the highest probability of malignancy. SLN biopsy would be useful if it could accurately identify patients in whom cancer has spread to the groin, without removing all groin nodes. SLNs can be identified by isosulfan blue dye and/or technetium-99 (^{99m}Tc) radioactive tracer during lymphoscintigraphy. The blue dye/^{99m}Tc procedure only detects SLN, not metastases – this requires histological examination, which can include ultrastaging and staining with conventional haematoxylin and eosin (H&E) or immunohistochemistry.

Objectives: To determine the test accuracy and cost-effectiveness of the SLN biopsy with ^{99m}Tc and/or blue dye compared with IFL or clinical follow-up for test negatives in vulval cancer, through systematic reviews and economic evaluation.

Data sources: Standard medical databases, including MEDLINE, EMBASE, Science Citation Index and The Cochrane Library, medical search gateways, reference lists of review articles and included studies were searched to January 2011.

Methods: For accuracy and effectiveness, standard methods were used and reported according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Searches were to January 2011, with no language restrictions. Meta-analyses were carried out with Meta-Disc version 1.4 (Javier Zamora, Madrid, Spain) for accuracy; none was appropriate for effectiveness. The economic evaluation from a NHS perspective used a decision-tree model in DATA TreeAge Pro Healthcare 2001 (TreeAge Software, Inc., Williamstown, MA, USA). Six options (blue dye with H&E, blue dye with

ultrastaging, ^{99m}Tc with H&E, ^{99m}Tc with ultrastaging, blue dye/^{99m}Tc with H&E, blue dye/^{99m}Tc with ultrastaging) were compared with IFL. Deterministic and probabilistic sensitivity analyses were conducted.

Results: For accuracy, of the 26 included studies, most evaluated ^{99m}Tc/blue dye combined. Four studies had clinical follow-up only for test negatives and five had clinical follow-up for all and IFL for test negatives. Numbers with no SLN found were difficult to distinguish from those with negative SLN biopsies. The largest group of 11 studies using ^{99m}Tc/blue dye, ultrastaging and immunohistochemistry had a pooled sensitivity of 95.6% [95% confidence interval (CI) 91.5% to 98.1%] and a specificity of 100% (95% CI 99.0% to 100%). Mean SLN detection rates were 94.6% for ^{99m}Tc, 68.7% for blue dye and 97.7% for both. One study measured global health status quality of life (QoL) and found no difference between SLN biopsy and IFL. One patient preference evaluation showed that 66% preferred IFL rather than a 5% false-negative rate from SLN biopsy. For effectiveness, of 14,038 references, one randomised controlled trial, three case–control studies and 13 case series were found. Approximately 50% died from vulval cancer and 50% from other causes during follow-ups. Recurrences were in the ratio of approximately 4 : 2 : 1 vulval, groin and distant, with more recurrences in node-positive patients. No studies reported QoL. For cost per death averted, IFL was less costly and more effective than strategies using SLN biopsy. For morbidity-free survival and long-term morbidity-free survival, ^{99m}Tc with ultrastaging was most cost-effective. Strategies with blue dye only and H&E only were never cost-effective. The incremental cost-effectiveness ratio for ^{99m}Tc with ultrastaging compared with IFL was £4300 per case of morbidity-free survival and £7100 per long-term morbidity-free survival.

Limitations: The main limitations of this study include the lack of good-quality evidence on accuracy, effectiveness and QoL. A large project such as this takes time to publish, so the most recent studies are not included.

Conclusions: A sensitive and specific combined metastatic SLN detection test and information on generic QoL in vulval cancer is urgently required.

Funding: The National Institute for Health Research Health Technology Assessment programme.

Contents

List of abbreviations

^{99m}Tc	technetium-99m
AE	adverse event
AIDS	acquired immune deficiency syndrome
CDSR	Cochrane Database of Systematic Reviews
CEAF	cost-effectiveness acceptability frontier
CENTRAL	Cochrane Central Register of Controlled Trials
CI	confidence interval
CK	cytokine
CKMNF	cytokine myocyte nuclear factor
DARE	Database of Abstracts of Reviews of Effects
DN	disease negative
DNA	deoxyribonucleic acid
DP	disease positive
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-C30
FACT-V	Functional Assessment of Cancer Therapy – Vulvar
FIGO	International Federation of Gynecology and Obstetrics
FN	false-negative
FP	false-positive
GROINSS-V	GROningen INternational Study on Sentinel nodes in Vulvar cancer
H&E	haematoxylin and eosin stain
HIV	human immunodeficiency virus
HMB	human melanoma black monoclonal antibody
HPV	human papillomavirus
HTA	health technology assessment
ICER	incremental cost-effectiveness ratio
IFL	inguinofemoral lymphadenectomy
IQR	interquartile range
LR	likelihood ratio
MEDION	medical diagnostic studies database
MeSH	medical subject heading
MM	malignant melanoma
OMNI	Organizing Medical Networked Information
PET-CT	positron emission tomography, computed tomography
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSA	probabilistic sensitivity analysis
QALY	quality-adjusted life-year
QoL	quality of life
QUADAS	quality of diagnostic accuracy studies
RCT	randomised controlled trial
ROC	receiver operating characteristic
RR	relative risk
RT	radiotherapy
RT-PCR	reverse transcription-polymerase chain reaction
SCC	squamous cell carcinoma
SD	standard deviation
SEER	US Surveillance, Epidemiology and End Results
SLN	sentinel lymph node
TN	true-negative

TNM	tumour node metastasis	VIN	vulval intraepithelial neoplasia
TP	true-positive	WHO	World Health Organization
UKGOSOC	United Kingdom Gynaecological Oncology Surgical Outcomes	WTP	willingness to pay

Scientific summary

Background

Vulval cancer is a relatively rare gynaecological malignancy most commonly seen in elderly patients. Ninety per cent of cases are squamous cell carcinomas (SCCs) and the remaining 10% are melanomas, Paget's disease, Bartholin's gland tumours, adenocarcinomas and basal cell carcinomas. Vulval cancer accounts for approximately 3–5% of all gynaecological malignancies and 1% of all cancers in women. Diagnosis is by biopsy with histological examination of the sample. This can include immunohistochemical analysis, which may enable more precise interpretation of the degree of dysplasia compared with conventional haematoxylin and eosin (H&E) staining. Vulval cancer can be locally invasive as well as spreading via the lymphatic system to the inguinal/femoral nodes. Staging is carried out using the International Federation of Gynecology and Obstetrics (FIGO) or tumour node metastasis (TNM) system. Since the late 1960s, the treatment of choice for vulval cancer has been surgical removal of the tumour and affected lymph nodes. Because of the risk of lymphatic spread to the groin nodes, lymphadenectomy of the inguinal and femoral nodes via inguinofemoral lymphadenectomy (IFL), either unilaterally or bilaterally, is undertaken depending on the stage and localisation of the cancer. Complications affect more than 50% of patients undergoing IFL and include infection of groin wounds, subsequent wound breakdown, lymphoedema and cellulitis. A sentinel lymph node (SLN) is the first lymph node that receives drainage directly from the primary tumour and, therefore, has the highest probability of containing cancer cells from the tumour in the vulva. If SLN biopsy could accurately identify those patients in whom cancer has and has not spread to the groin nodes without extensive surgical removal of all of the groin nodes, this would be of extremely high value in sparing patients from undergoing unnecessary full groin node dissection or IFL. SLNs can be identified by using a dye called isosulfan blue or a radioactive tracer called technetium-99 (^{99m}Tc) in a procedure called lymphoscintigraphy. Blue dye and ^{99m}Tc can be used alone or in combination. The blue dye/^{99m}Tc procedure only detects the SLN, but cannot determine whether or not the SLN has metastatic deposits. For this, histopathological examination is required, which can include ultrastaging (cutting thinner slices) and immunohistochemistry.

Objectives

To determine the test accuracy and cost-effectiveness of SLN biopsy with ^{99m}Tc enhanced and/or blue dye lymphoscintigraphy for diagnosis of IFL in cases of vulval cancer through systematic reviews and economic evaluation.

Methods

A protocol was developed for test accuracy and effectiveness systematic reviews and the economic evaluation. For the systematic reviews, standard methods were used and are reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. In included studies, at least 75% of women had been with diagnosed vulval cancer of FIGO stage IB or II or TNM categories T1–2, N0–2, M0. For the test accuracy reviews, any studies evaluating SLN biopsy with ^{99m}Tc or blue dye, or both, with reference standard of IFL for all, or for test positives with clinical follow-up for test negatives, were included. Quality assessment was conducted using quality of diagnostic accuracy studies (QUADAS) criteria. For the effectiveness reviews, randomised controlled trials (RCTs), cohort, case–control or case series of surgical or radiotherapy (RT) treatment with outcomes including survival, recurrence, early and late complications and quality of life (QoL) were included. Quality assessment was performed appropriate to the study designs. Inclusion decisions, quality assessment and data extraction were

performed in duplicate with disagreements resolved through discussion. Results are presented narratively and in tables. Meta-analyses were performed using Meta-Disc version 1.4 (Javier Zamora, Madrid, Spain) for test accuracy results. No meta-analysis was appropriate for effectiveness reviews.

For the economic evaluation, the model structure used was a decision tree constructed in DATA TreeAge Pro 2001 software (TreeAge Software Inc., Williamstown, MA, USA). The NHS perspective was used. Six options (blue dye with H&E, blue dye with ultrastaging, ^{99m}Tc with H&E, ^{99m}Tc with ultrastaging, blue dye and ^{99m}Tc with H&E, blue dye and ^{99m}Tc with ultrastaging) were compared with IFL for all. Inputs to the model were test accuracy and effectiveness systematic review results, test accuracy and intervention costs, costs of vulval cancer and the rate of recurrence. The primary analysis used point estimates of key parameters and extensive deterministic and probabilistic sensitivity analyses (PSA) were conducted. As no QoL information was available, the outputs were in terms of cost per death averted at 2 years, cost per patient experiencing morbidity-free survival at 2 years and cost per patient experiencing long-term morbidity-free survival at 2 years.

Data sources

Sensitive searches with both medical subject heading (MeSH) terms and text words were used in a variety of databases including MEDLINE, EMBASE, Science Citation Index, MEDION, The Cochrane Library [Cochrane Central Register of Controlled Trials (CENTRAL), health technology assessment (HTA), Database of Abstracts of Reviews of Effects (DARE), Systematic Reviews], clinical trials, medical search gateways [including Organizing Medical Networked Information (OMNI), National Cancer Institute, Google, Copernic], from inception to January 2011, with no language restrictions. Reference lists of reviews and guidelines were also searched.

Results

For the test accuracy systematic review, of 2942 references, 26 studies were included. Most studies were small, with fewer than 50 women. The largest, by Van der Zee *et al.* (Van der Zee AG, Oonk MH, de Hullu JA, Ansink AC, Vergote I, Verheijen RH, *et al.* SLN dissection is safe in the treatment of early-stage vulvar cancer. *J Clin Oncol* 2008;**26**:884–9), included 403 women, the vast majority of whom had SCC. Most studies evaluated ^{99m}Tc combined with blue dye and used a range of histopathological techniques. Four had clinical follow-up only for test negatives and five had clinical follow-up for all as well as IFL for test negatives. Reporting of results was not clear, and it was difficult at times to distinguish between the number of patients who had no SLN metastases but had metastases in other lymph nodes and patients with a negative SLN biopsy with metastases in other lymph nodes. Because of the variety of tests and immunopathological techniques used, pooling of all studies was not appropriate. In addition, test accuracy results are reported here on the basis of finding a SLN. All of the point estimates of sensitivity were above 90% for studies with IFL for all or when using groin and distant recurrences only for clinical follow-up. All of the point estimates of specificity were 100% because false-positive results were not possible. The largest group of 11 studies using ^{99m}Tc with blue dye, ultrastaging and immunohistochemistry had a pooled sensitivity of 95.6% [95% confidence interval (CI) 91.5% to 98.1%] and a specificity of 100% (95% CI 99.0% to 100%). The mean (95% CI, range) SLN detection rates were 94.6% (90.9% to 97.1%, range 76–100%) for ^{99m}Tc only, 68.7% (63.1% to 74.0%, range 53–88%) for blue dye only and 97.7% (96.6% to 98.5%, range 84–100%) for both. The results suggest that if SLN biopsy is going to be used, both tests should be performed in every patient. The only study to measure QoL found no difference between SLN biopsy and IFL groups for global health status.

For the effectiveness systematic review, of 14,038 references, one RCT, three case–control studies and 13 case series were found. The RCT compared IFL with RT to groin nodes in women undergoing surgery for SCC. Survival was better in the IFL arm. The case–control studies compared single-incision with

triple-incision IFL, RT versus no RT to the groin and hemivulvectomy with unilateral IFL to vulvectomy with bilateral IFL. The case series evaluated a variety of treatment options in vulval cancer. Most studies were small, and the largest by far (Kumar S, Shah JP, Bryant CS, Imudia AN, Morris RT, Malone JM Jr. A comparison of younger vs. older women with vulvar cancer in the United States. *Am J Obstet Gynecol* 2009;**200**:e52–5) reported results on 5620 women from the US Surveillance, Epidemiology and End Results (SEER) database. All case–control studies and case series were evaluated for survival, recurrence and adverse events (AEs). The general trends were approximately 50% of women dying from vulval cancer and 50% from other causes during the follow-ups. Recurrences were in the ratio of approximately 4 : 2 : 1 vulval, groin and distant, with more recurrences in node-positive patients. No studies reported QoL.

In the economic evaluation, the results of the base-case deterministic analyses based on the outcome of cost per death averted showed that IFL was both less costly and more effective than any of the strategies that used SLN biopsy. When considering the outcome measures of morbidity-free survival and long-term morbidity-free survival, it was found that the strategy ^{99m}Tc + ultrastaging in which ultrastaging was administered in the case of a negative H&E test was most cost-effective. Note that ultrastaging here is used as a proxy for more involved histopathological techniques such as immunohistochemistry. Moreover, it was noted that the strategies that included blue dye only as the approach to the SLN biopsy and H&E only for the histopathology were never found to be cost-effective and were always dominated by other strategies (other strategies being less costly and more effective). This finding emphasises that using blue dye and H&E for the identification of the SLN and metastasis, respectively, are not sensitive enough to be used on their own.

The incremental cost-effectiveness ratio (ICER), based on the outcome of morbidity-free survival for the strategy of blue dye + ultrastaging compared with IFL, was £2400 per case of morbidity-free survival, the ICER for ^{99m}Tc + ultrastaging compared with blue dye + ultrastaging was £4900 per case of morbidity-free survival and the ICER for ^{99m}Tc + blue dye + ultrastaging compared with ^{99m}Tc + ultrastaging was £41,000 per case of morbidity-free survival. Similarly, for the outcome measure of long-term morbidity-free survival, the strategy of blue dye + ultrastaging compared with IFL was £3700 per case of long-term morbidity-free survival, the ICER for ^{99m}Tc + ultrastaging compared with blue dye + ultrastaging was £8900 per case of long-term morbidity-free survival and the ICER for ^{99m}Tc + blue dye + ultrastaging compared with ^{99m}Tc + ultrastaging was £74,300 per case of long-term morbidity-free survival.

Limitations

Limitations of the project included lack of sufficiently accurate information on test accuracy, effectiveness of the various treatments, QoL of life and costs of SLN biopsy. A large project such as this takes time, so the search dates are relatively early and more studies may have been published since. As there were no QoL data, three outcome measures have been considered in this study (overall mortality, morbidity-free survival and long-term morbidity-free survival), so the cost-effectiveness results are not readily transferable across disciplines.

Conclusions

Compared with a strategy involving SLN biopsy in routine clinical practice, the strategy of IFL for all was found to be less costly and more effective when considering cost per death averted. Based on the findings of the current model and acknowledging the limitations that have been highlighted in terms of the inability to apply quality-adjusted life-years (QALYs) in this economic evaluation, the results of this analysis suggest that ^{99m}Tc + ultrastaging in the treatment of early-stage vulval cancer is likely to be cost-effective in terms of case of morbidity averted and long-term morbidity averted. Note that ultrastaging has been used here as a proxy for more in-depth histopathological techniques such as immunohistochemistry. There is some uncertainty regarding the acceptability of the ^{99m}Tc + blue dye + ultrastaging strategy in terms of the

outcome measures of case of morbidity and long-term morbidity averted at 2 years, as there is difficulty in attempting to apply the outcome measures used in this study to any acceptability threshold.

Implications for practice

There is insufficient evidence to suggest that SLN biopsy should be used in routine clinical practice on health economic grounds. The strategy of IFL for all was found to be less costly and more effective when considering cost per death averted.

Recommendations for further research

There needs to be further evaluation of patient preferences regarding the circumstances when patients would rather risk unremoved groin metastases by forgoing IFL should they have SLN biopsy and it is negative. This would incorporate factors including the patient age, disease stage and the aggressiveness of the malignancy in the vulval specimen.

There needs to be a robust prospective evaluation of the relative effectiveness of the different treatment strategies for vulval cancer, taking into account the uncertainty around the need for IFL in early-stage vulval cancer. As vulval cancer is uncommon, a multicentre RCT involving several countries will probably be needed to enrol sufficient patients in order to deal with the uncertainty.

There needs to be some information on the QoL in vulval cancer, using a generic QoL measure such as Euroqol EQ-5D. This analysis has highlighted the importance of obtaining overall QoL values that describe the impact of the SLN biopsy and IFL and their related complications on patients over time. A previous study has attempted to identify these values but did not find a difference in the QoL estimates between 62 patients who received either a SLN biopsy or an IFL. Intuitively there would need to be a difference in QoL between these two groups, since, if this were not the case, IFL, with its increased effectiveness at reducing the risk of a further groin recurrence and therefore patient mortality but with its much higher risk of morbidity, would always be preferred. Therefore, future in-depth work should be undertaken to examine the QoL in these treatment groups perhaps by using an alternative type of questionnaire and through a larger study that includes more patients so would have better power to determine a small difference.

Funding

The National Institute for Health Research Health Technology Assessment programme.

Chapter 1 Aim of the report

The aims of this project were as follows:

- To determine the accuracy of sentinel lymph node (SLN) biopsy with technetium-99 (^{99m}Tc) enhanced and/or blue dye lymphoscintigraphy for diagnosis of inguinofemoral lymphadenopathy (IFL) in vulval cancer through systematic review.
- To assess, through systematic review, the diagnostic and therapeutic impact of SLN biopsy with ^{99m}Tc enhanced and/or blue dye lymphoscintigraphy in:
    - changing disease staging
    - changing planned treatment
    - reducing complications associated with IFL
    - improving morbidity and disease-free survival.
- To determine the effectiveness of various interventions [e.g. surgery, radiotherapy (RT) and chemotherapy] in the management of vulval cancer.
- To evaluate the cost-effectiveness of SLN biopsy with ^{99m}Tc enhanced and/or blue dye lymphoscintigraphy versus IFL using decision-analytic modelling.

The original protocol for the project is in *Appendix 1*.

Chapter 2 Background

Description of underlying health problem

Vulval cancer is a relatively rare gynaecological malignancy diagnosed mainly among elderly patients. In 90% of cases it develops as a squamous cell carcinoma (SCC) and the remaining 10% are melanomas, Paget's disease, Bartholin's gland tumours, adenocarcinomas and basal cell carcinomas.[1] Lesions mainly occur on the inner edges of the labia majora (around 50% of the cases), less often in the labia minora and very rarely on the clitoris or in the Bartholin glands. The symptoms of vulval cancer include a lump on the vulva, vulval bleeding, itching, pain or ulceration, and approximately 90% of women present with a visible tumour.[1] Diagnosis is by biopsy with histological examination of the sample. Histological examination can include immunohistochemical analysis which may enable more a precise interpretation of the degree of dysplasia compared with conventional haematoxylin and eosin (H&E) staining.[2]

Vulval cancer can be locally invasive and can spread via the lymphatic system to the inguinal or femoral nodes. If the primary tumour is laterally located in the vulva, spread may be only to that side; if the tumour is central, spread may be to either side. Lymphatic spread is strongly related to lesion size: metastasis is present in 20–30% of tumours < 2 cm in diameter and in 44% of tumours > 2 cm.[3] Correlation of lymph node status with depth of invasion is also important. Tumours with a < 1-mm depth of invasion have a < 1% risk of nodal spread.[4,5] Lymph node status is regarded as an important predictor of survival.[6] Malignancy in the lymph nodes can result in invasion of the blood vessels of the groin, including rupture of the femoral blood vessels. Distal spread via the bloodstream is relatively rare.[2]

Once vulval cancer is diagnosed, staging is important in order to plan treatment and estimate prognosis. Two staging systems are frequently used, one developed by the International Federation of Gynecology and Obstetrics (FIGO)[7,8] and one by the tumour node metastasis (TNM) classification of malignant tumours[7] (*Tables 1* and *2*). Both staging systems have gradually evolved over time. *Table 3* gives a comparison of FIGO and TNM staging.[2]

The grade of malignancy refers to the extent of differentiation seen on microscopic examination and gives an indication of how fast the malignancy is likely to develop. In vulval cancer three grades are defined:

- Grade 1: cells are low grade or well differentiated and histologically look very much like normal vulval cells.
- Grade 2: cells are medium grade or moderately differentiated and look more abnormal than grade 1 cells, but not so much as grade 3 cells.
- Grade 3: cells are high grade or poorly differentiated or can be undifferentiated, they are very unlike normal vulval cells.

Premalignant conditions

Vulval intraepithelial neoplasia (VIN) is a precancerous condition that can occur in the vulval area and may present as pigmented lesions. There are two main types of VIN, each with their own distinctive clinical and pathological features. The classifications in use are the World Health Organization (WHO)'s classification that refers to human papillomavirus (HPV)-associated VIN as warty/basaloid VIN and grades the disease as VIN1, VIN2 and VIN3. The International Society for the Study of Vulvovaginal Disease (ISSVD) has proposed a newer classification referring to VIN usual type and VIN differentiated type. The more commonly referred to classical or usual-type VIN is a disease of relatively young women and is associated with HPV. HPV deoxyribonucleic acid (DNA), mostly type 16, is present in up to 90% of classical VIN. Typically, classical VIN involves the vulva multifocally and is associated with multicentric involvement of the vagina and cervix. Classical VIN progresses to cancer in only 3–10% of treated patients. The other type of VIN is referred to

TABLE 1 Federation of Gynaecology and Obstetrics staging of vulval cancer

Stage	FIGO staging 1969 Clinical staging	FIGO staging 1988 Surgical staging	FIGO staging 2000 Surgical staging	FIGO staging 2009 Surgical staging
0	Carcinoma in situ	Carcinoma in situ	Carcinoma in situ	Carcinoma in situ
I	Tumour confined to vulva, 2 cm or less in largest diameter and no suspicious groin nodes	Tumour confined to vulva or perineum, < 2 cm in greatest dimension and nodes are negative	Ia: tumour confined to vulva or vulva and perineum, 2 cm or less in greatest dimension and with stromal invasion no greater than 1 mm. Nodes are negative Ib: tumour confined to vulva or vulva and perineum, 2 cm or less in greatest dimension and with stromal invasion greater than 1 mm. Nodes are negative	IA: lesions ≤ 2 cm in size, with depth of invasion ≤ 1 mm IB: lesions > 2 cm in size, with depth of invasion > 1 mm
II	Tumour confined to vulva, more than 2 cm in diameter and no suspicious groin nodes	Tumour confined to vulva or perineum, > 2 cm in greatest dimension, nodes are negative	Tumour confined to vulva or vulva and perineum, more than 2 cm in greatest dimension. Nodes are negative	Tumour of any size with extension to adjacent perineal structures: one-third lower urethra, one-third lower vagina, anus. No lymph node metastases
III	Tumour of any size with (1) adjacent spread to the urethra and/or vagina, perineum, and anus, and/or (2) clinically suspicious lymph nodes in either groin	Tumour of any size with (1) adjacent spread to the lower urethra or anus, and/or (2) unilateral regional lymph nodes metastases	(1) Tumour invades any of the following: lower urethra, vagina, anus (2) Unilateral regional lymph nodes metastases	Tumour of any size, with or without extension to adjacent perineal structures, with positive lymph nodes IIIA: one lymph node metastasis (≥ 5 mm) or 1–2 lymph node metastases (< 5 mm) IIIB: at least two lymph node metastases (≥ 5 mm) or at least three lymph node metastases (< 5 mm) IIIC: lymph node metastases with extracapsular spread
IVa	IV: tumour of any size, (1) infiltrating the bladder mucosa or the rectal mucosa, or both, including the upper part of the urethral mucosa, and/or (2) fixed to the bone, and/or (3) other distant metastases	Tumour invasion of any of the following: upper urethra, bladder mucosa, rectal mucosa, pelvic bone and/or bilateral regional node metastasis	Any size tumour with bilateral regional lymph node involvement or with invasion of any of the following: upper urethra, bladder or rectal mucosa, pelvic bone	Tumour invades any of the following: upper urethral and/or vaginal mucosa, bladder or rectal mucosa, or fixed to the pelvic bone. Or fixed or ulcerated lymph nodes
IVb	–	Any distant metastasis, including pelvic lymph nodes	Any distant metastasis, including pelvic lymph nodes	Any distant metastases including pelvic nodes

TABLE 2 Tumour node metastasis classification of carcinoma of the vulva

	TNM 1969	TNM 1988	TNM 2000
Stage	Clinical staging	Surgical staging	Surgical staging
Tx	–	Primary tumour cannot be assessed	Primary tumour cannot be assessed
T0	–	No evidence of primary tumour	No evidence of primary tumour
T_{is}	–	Carcinoma in situ	Carcinoma in situ
T1	Tumour confined to the vulva, 2 cm in largest diameter	Tumour confined to the vulva and/or perineum, ≤ 2 cm in greatest dimension	T1a: tumour confined to the vulva and/or perineum, ≤ 2 cm in greatest dimension and with stromal invasion no greater than 1 mm T1b: tumour confined to the vulva and/or perineum, ≤ 2 cm in greatest dimension and with stromal invasion > 1 mm
T2	Tumour confined to the vulva, 2 cm in largest diameter	Tumour confined to the vulva and/or perineum, > 2 cm in greatest dimension	Tumour confined to the vulva and/or perineum, > 2 cm in greatest dimension
T3	Tumour of any size with adjacent spread to the urethra, and/or vagina, and/or perineum and/or anus	Tumour involves any of the following: the lower urethra, vagina, anus	Tumour invades any of the following: the lower urethra, vagina, anus
T4	Tumour of any size infiltrating the bladder mucosa, and/or the rectal mucosa, or including the upper part of the urethral mucosa and/or fixed to the bone	Tumour involves any of the following: bladder mucosa, rectal mucosa, upper urethra, pelvic bone	Tumour invades any of the following: bladder mucosa, rectal mucosa, upper urethral mucosa or is fixed to bone
Nx		Regional (i.e. femoral and inguinal) lymph nodes cannot be assessed	Regional (i.e. femoral and inguinal) lymph nodes cannot be assessed
N0	No nodes palpable	No lymph node metastases	No lymph node metastases
N1	Nodes palpable in either groin, not enlarged, mobile (not clinically suspicious for neoplasm)	Unilateral regional lymph node metastases	Unilateral regional lymph node metastases
N2	Nodes palpable in either groin, enlarged, firm and mobile (clinically suspicious for neoplasm)	Bilateral regional lymph node metastases	Bilateral regional lymph node metastases
N3	Fixed or ulcerated nodes	–	–
Mx	–	Distant metastases cannot be assessed	Distant metastases cannot be assessed
M0	No clinical metastases	No distant metastases	No distant metastases
M1	M1a: palpable deep pelvic lymph nodes M1b: other distant metastases	Distant metastases, including pelvic lymph node metastases	Distant metastases, including pelvic lymph node metastases

T_{is}, tumour in situ.

TABLE 3 Comparison of FIGO and TNM staging in vulval cancer

	TNM status		
FIGO stage	Tumour	Node	Metastases
0	T_{is}	N0	M0
IA	T1a	N0	M0
IB	T1b	N0	M0
II	T2	N0	M0
III	T3	N0	M0
	T3	N1	M0
	T1	N1	M0
	T2	N1	M0
IVA	T[a]	N2	M0
	T4	N[a]	M0
IVB	T[a]	N[a]	M1

T_{is}, tumour in situ.
a Any stage.

as simplex or differentiated VIN and was described in the 1960s, but added to ISSVD classification in 2004/5. It constitutes 2–10% of VIN diagnoses, but is seen adjacent to up to 60% of vulval cancers. It characteristically occurs in postmenopausal women and is associated with lichen sclerosis. Histologically, it is a subtle lesion and is likely to be underdiagnosed because of its high degree of differentiation.[9–14]

Aetiology

Several factors increase the risk of developing cancer of the vulva, but none confers high risk. Factors increasing risk of malignancy include HPV infection, smoking, human immunodeficiency virus (HIV) infection and other factors.

Smoking

Smoking is a risk factor for the development of many malignancies, including vulval cancer.[15,16] For women already infected with HPV, smoking further raises the risk of neoplasia. It has been reported that cigarette smoking increases the risk of vulval cancer development by 25 times for women smoking ≥ 20 cigarettes a day who also have serological evidence of HPV-16 infection.[16] Another study reported that active smoking was associated with an increased risk of developing vulval cancer [relative risk (RR) 2.03, 95% confidence interval (CI) 1.3 to 3.2].[17]

Human papillomavirus infection

Human papillomavirus infection is thought to cause up to half of vulval cancers, mostly in women under the age of 50 years. HPV is less likely to be a risk factor in older women. Vulval carcinoma development is associated with infection of high-risk HPV types. A case–control study found that over 30% of examined tumours contained HPV DNA of types 16, 18 or 33. Moreover, HPV DNA was found in 61% of tumours with adjacent intraepithelial neoplasia (VIN III) and in 13% of tumours without associated VIN III.[18] The prevailing evidence favours HPV as one causative factor in genital tract carcinoma.[19]

Human immunodeficiency virus infections

Human immunodeficiency virus causes the acquired immune deficiency syndrome (AIDS). Owing to its destructive effect on the human immune system, women burdened with it are more prone to a HPV infection, which, in turn, may be easily linked with vulval cancer development.[20]

Other factors

A higher risk of vulval cancer is associated with lower socioeconomic status and fewer years of education.[21] It has been found that a prior history of lichen sclerosus of the vulva and inflammation of the vulva or vagina are significantly associated with vulval cancer development. Some other evidence suggests that environmental exposures may also play an important role in vulval carcinogenesis.[22] In addition, some studies have shown a link between the risk of vulval cancer occurrence and psoriasis,[23] or vulval cancer occurrence and transplant immunosuppression.[24]

Epidemiology

Vulval cancer accounts for approximately 3–5% of all gynaecological malignancies and 1% of all cancers in women.[25] The worldwide prevalence of vulval cancer is around 3% and it is estimated that 27,000 women are diagnosed each year.[21] In the UK, the lifetime risk of developing vulval cancer is 1 in 316.[21] According to a recent report, the highest rate of occurrence is in North America, South America and Europe, with incidence rates of between 1 and 2 per 100,000 per year.[26]

Vulval cancer is very rare in young women aged < 25 years. Incidence rates are around 1.1 per 100,000/year, among women aged 25–39 years, rising to 3.8 per 100,000/year in those aged 60–64 years and peaking at 24.5 per 100,000/year in women aged ≥ 85 years. The proportion of women diagnosed with vulval cancer under the age of 50 years has risen from 7% in the 1970s to 14% in 2006–8.[21]

Each year in the UK, there are, on average, 373 deaths from vulval cancer. The European age-standardised death rate for vulval cancer in the UK was 0.64 per 100,000 female population over the 5-year period 2004–8. The mortality rate for vulval cancer is 0.1 per 100,000 women aged 25–44 years and 0.5 for women aged 45–64 years, rising to 6 per 100,000 women aged ≥ 65 years and 12 per 100,000 women aged ≥ 80 years. Mortality rates for vulval cancer in the UK have declined steadily since the early 1970s. The rate fell by almost half (48%) between 1971 and 2008, from 1.3 per 100,000 female population to 0.7 per 100,000.[27]

Prognosis

Vulval cancer is highly curable when diagnosed at an early stage.[28] Survival is most dependent on primary lesion diameter and the pathological status of the inguinal nodes.[1] Linking staging to survival is not precise if it is unclear whether the nodes are involved (which would be stage III) or not (which would be stage II). In a study of 588 patients treated in 1977–84,[29] survival of patients with FIGO stage I, II, III and IV disease was 98%, 87%, 75% and 29%, respectively. If nodes are involved, survival is linked to the number of nodes involved and unilaterality versus bilaterality. In patients with tumour diameter > 8 cm and three or more unilateral or two bilateral metastatic nodes, the relative survival at 5 years was 0%.[29] The following aspects have been accounted as the risk factors for node metastasis: clinical node status, age, degree of differentiation, tumour stage, tumour thickness and depth of stromal invasion.[29,30] The presence of capillary–lymphatic space invasion is associated with increased local recurrence in the vulva but not with increased risk of groin node metastases.[1]

Another early study gave 5-year survival rates for RT of only 13.0%, simple vulvectomy of 30.8%, vulvectomy and RT of 32.5%, radical vulvectomy with IFL of 55.7% and radical vulvectomy with inguinal, femoral and pelvic lymphadenectomy of 63.3%.[31]

Current treatment options

Since the late 1960s, the treatment of choice for vulval cancer has been surgical removal of the tumour and affected lymph nodes.[1] The actual treatment used for SCC depends on the stage of tumour found (*Table 4*).[3] Surgical management of non-SCC cancers vary by type of cancer. For carcinoma of the Bartholin's gland, treatment follows that for SCC tumours but the tumours are more likely to be deep and metastatic. For basal cell and verrucous carcinomas, wide local excision is usually used as metastases are rare. For malignant melanoma (MM), wide local excision is preferred, but there is a high rate of relapse.[1]

TABLE 4 Treatment options for SCC of the vulva by FIGO stage

Stage	Definition	Treatment
0	Carcinoma in situ and less advanced precancerous changes, e.g. VIN	Laser surgery, wide local excision, or a skinning vulvectomy; alternatively fluorouracil ointment may be prescribed or imiquimod
I	Treatment depend on the size and depth of the cancer and VIN occurrence	–
	IA – depth of invasion of 1 mm or less, and there are no other areas of cancer or VIN	Wide local excision. The cancer is being removed along with a 1-cm margin of the normal tissue surrounding
	IB – lesions > 2 cm in size with depth of invasion > 1 mm	Wide local excision/hemi/radical vulvectomy and inguinal lymph node dissection
		SLN biopsy may be performed instead of lymph node dissection when tumour size is < 4 cm and is unifocal
II	Cancer spread to structures near the vulva area	Partial radical vulvectomy
		Optional removal of the lymph nodes in the groin on both sides of the body
		If cancer cells are present near the margins the radiation therapy to the area of surgery is performed
		Radiation therapy with surgery in order to remove remaining cancer tissues. Chemotherapy with fluorouracil and/or cisplatin
III	Cancer spread to nearby lymph nodes	Surgical removal of cancer and lymph nodes in the groin optionally followed by radiation therapy. Radiation and chemotherapy are applied for patients not able to undergo surgery
IVA	Cancer spread to organs and tissues in the pelvis: rectum, bladder, pelvic bone, upper part of the vagina and urethra. Tumours of type T1 and T2 with less severe nearby spread but with extensive spread to nearby lymph nodes	Removal of as much as possible of the affected tissue with possible pelvic exenteration. Usually operation includes vulvectomy, removal of the pelvic lymph nodes or, alternatively, lower colon, rectum, bladder, uterus, cervix and vagina
		Gold standard: combination of surgery, radiation and chemotherapy
IVB	Cancer spread to lymph nodes in the pelvis or to organs and tissues outside the pelvis	None of the approaches is assumed to cure the malignancy though they may be helpful in relieving some symptoms emerging from the disease
Recurrent	Cancer recurrence	Treatment depends on the recurrence time and location. Local recurrence: surgery or combination of three approaches. Unresectable recurrence: chemotherapy and/or radiation therapy

Surgery to the vulva can be radical vulvectomy (complete removal of the vulva), hemivulvectomy or, more commonly, wide local excision.[1] If necessary, skin grafting using skin from the thigh can be used. The intention of surgery is complete removal of the lesion with a minimum margin of 15 mm disease-free tissue on all sides of the specimen. The side effects of vulvectomy are extensive pain, disfigurement and the risk of sphincter damage leading to urinary or faecal incontinence. There can be considerable psychological trauma and loss of psychosexual function.[32]

Because of the risk of lymphatic spread to the groin nodes, lymphadenectomy of the inguinal and femoral nodes, either unilaterally or bilaterally, is usually undertaken, depending on the stage of the disease and the laterality of the tumour. IFL can be omitted safely if the tumour depth of invasion is < 1 mm. Unilateral IFL is performed for lateral tumours that are at least at a 1-cm distance from the midline of the vulva.

Bilateral IFLs are performed for tumours encroaching within 1 cm of the midline. Complications affect over 50% of patients having IFL, including infection of groin wounds, subsequent wound breakdown, lymphoedema and cellulitis.[33–35] If the groin breaks down, patients need to stay in hospital for several days longer than otherwise and will need antibiotics and community care once discharged. As patients are often elderly, this additional morbidity can compromise overall recovery. Lymphoedema may be the most aggravating as it significantly limits overall mobility,[36] is disfiguring, causes difficulties in daily living, can lead to lifestyle becoming severely limited and may also result in psychological distress. Over the last 20 years, use of the triple-incision technique in the groin, for which three cuts are made, leaving skin bridges between so that the skin can heal more quickly than one longer single (butterfly) incision, has reduced postoperative stay and improved recovery from IFL.[37]

Patients with poor prognostic features may additionally receive RT covering the pelvis, groin and the perineum area. In cases of locally advanced disease, standard management includes surgery, chemotherapy and radiation therapy (sequentially or in combination).[1] If patients are unable to withstand surgery, RT alone or, occasionally, chemotherapy alone can be used.

Sentinel lymph node biopsy

Currently, clinical examination for the determination of metastatic involvement of groin lymph nodes is insufficiently accurate, particularly when the nodes may contain micrometastases but appear clinically normal. Out of all patients with clinically normal lymph nodes, between 16% and 24% will go on to develop metastases;[30,38] therefore, interest has been shown in the use of imaging modalities and SLN biopsy in order to determine lymphatic spread and thereby more accurately stage vulval cancer and reduce the need for unnecessary surgery. If IFL is undertaken and nodes are negative for metastases, there is considerable morbidity from the IFL, which is associated with longer hospital stays. If IFL is not undertaken and there were micrometastases that were missed, survival rates are reduced. So IFL is a surgical procedure that serves to obtain lymph nodes for histopathology as a diagnostic test for metastases and IFL can also remove clinically suspicious enlarged nodes to improve treatment success rates.

A SLN refers to any lymph node that receives drainage directly from the primary tumour and is the first in the chain of lymph nodes in the groin and, therefore, has the highest probability of containing cancer cells from the tumour in the vulva.[39] SLNs can be identified by using a dye called isosulfan blue or a radioactive tracer called ^{99m}Tc in a procedure called lymphoscintigraphy. Blue dye and ^{99m}Tc can be used alone or in combination.[40] The blue dye/^{99m}Tc procedure only detects the SLN, but cannot determine whether or not the SLN has metastatic deposits. For this, histopathological examination is required. This is best done by routine histopathology using H&E staining, although, in some centres, frozen sections may also be used. Lymph nodes can be cut in a variety of slices, with thinner slices known as ultrastaging. Immunohistochemical techniques that will enhance the ability to detect metastatic deposits can also be used.

There is a risk with SLN biopsy that malignancy may be missed. It may be that the first draining lymph node was missed, or that the malignancy developed not in the SLN but in any of the groin nodes other than the SLNs biopsied and examined with histology. There has been one small survey of vulval cancer patients evaluating the acceptability of SLN biopsy compared with IFL at different levels of risk.[41] This 106 patients who were surveyed had fully recovered from vulval cancer (99 questionnaires could be evaluated) and had received IFL as part of their treatment.[41] It was found that 66% would recommend IFL if the risk of missing metastasis from SLN biopsy was 1 in 80, and 84% would recommend IFL if the risk of missing metastasis from SLN biopsy was one in eight. Age and the presence or degree of side effects experienced by the patients surveyed, which included severe lymphoedema in 39% and with severe pain in 28%, did not affect preferences for each procedure.

The extent of SLN biopsy being undertaken in the NHS for vulval cancer is currently unclear. The most recent guideline from the UK Royal College of Obstetricians and Gynaecologists states that ‘Dye studies and lymphoscintigraphy may be of value in the detection of SLNs although the outcome of this type of intervention is awaiting the outcome of controlled clinical evaluation’.[1]

Chapter 3 Definition of the decision problem

If SLN biopsy could accurately identify those patients in whom cancer has spread to the groin nodes without extensive surgical removal of all of the groin nodes, this would be extremely valuable. If this technique was very accurate in detecting no metastases, no radical treatment would be necessary. In order to test the accuracy of SLN biopsy, the reference standard can be IFL for all node-positive and -negative patients. Alternatively, clinical follow-up could be used for node-negative patients if SLN was considered to be sufficiently accurate not to miss patients with metastases in the lymph nodes. The different possibilities are illustrated in *Figure 1* for the situations in which a SLN was found at biopsy and *Figure 2* in the case that a SLN was not found at biopsy. The histopathology of the SLN biopsy should ideally be compared with the same type of histopathology used for all of the lymph nodes examined after IFL because the histopathology is part of the SLN biopsy test as well as the reference standard. If frozen sections or immunohistochemistry were used for the SLN histopathology and routine histological techniques such as H&E staining were used only for the IFL nodes, then that would, in effect, mean a different reference standard was being used.

In the decision trees, recurrence refers to groin recurrence; however, any recurrence is important, and both groin and distant recurrence may be reduced following IFL, but not local recurrence in the vulva.

The aim of this health technology assessment (HTA) was to determine the test accuracy and cost-effectiveness of SLN biopsy in vulval cancer by systematic reviews and decision-analytic modelling.

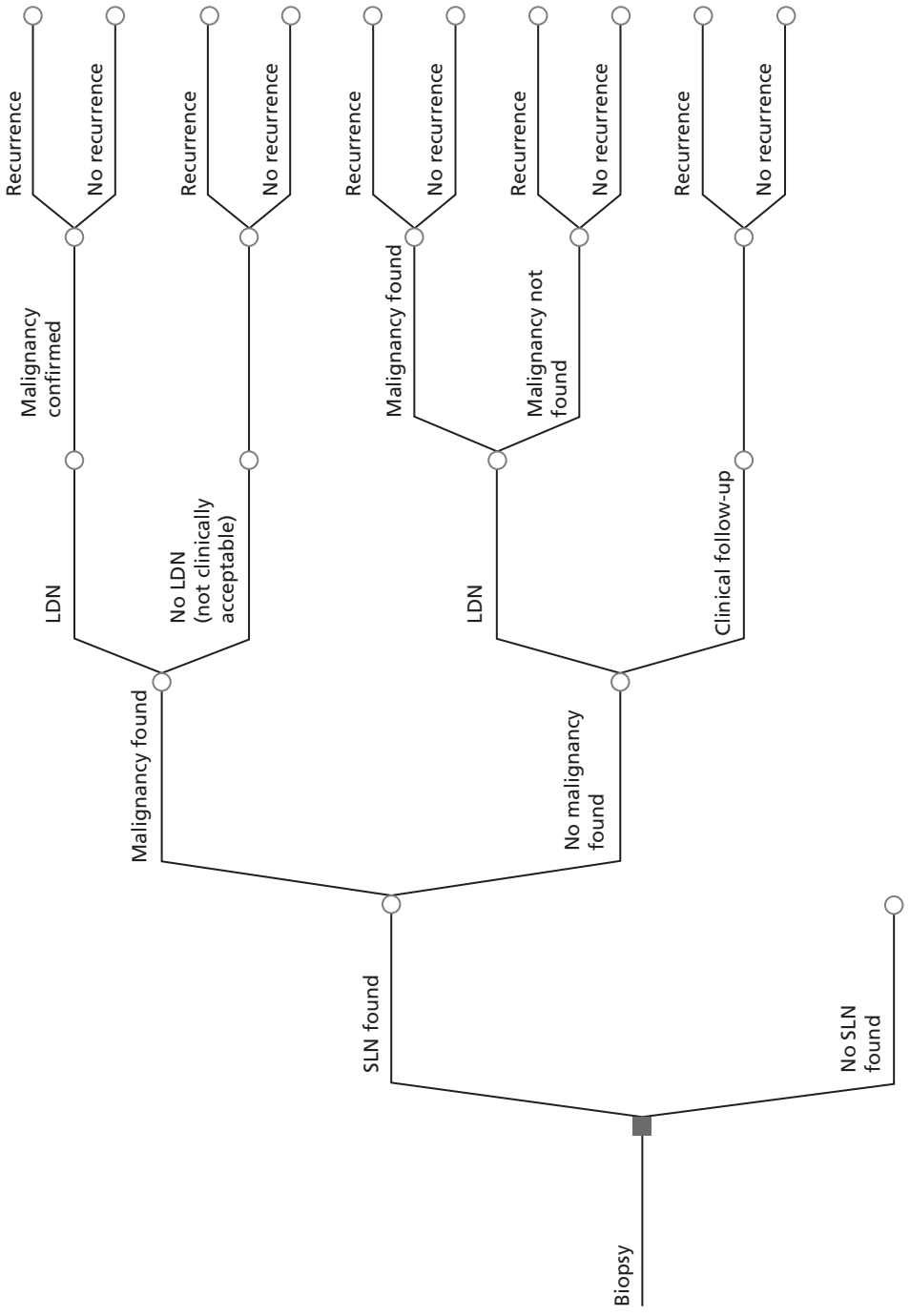


FIGURE 1 Decision tree for SLN biopsy (part 1).

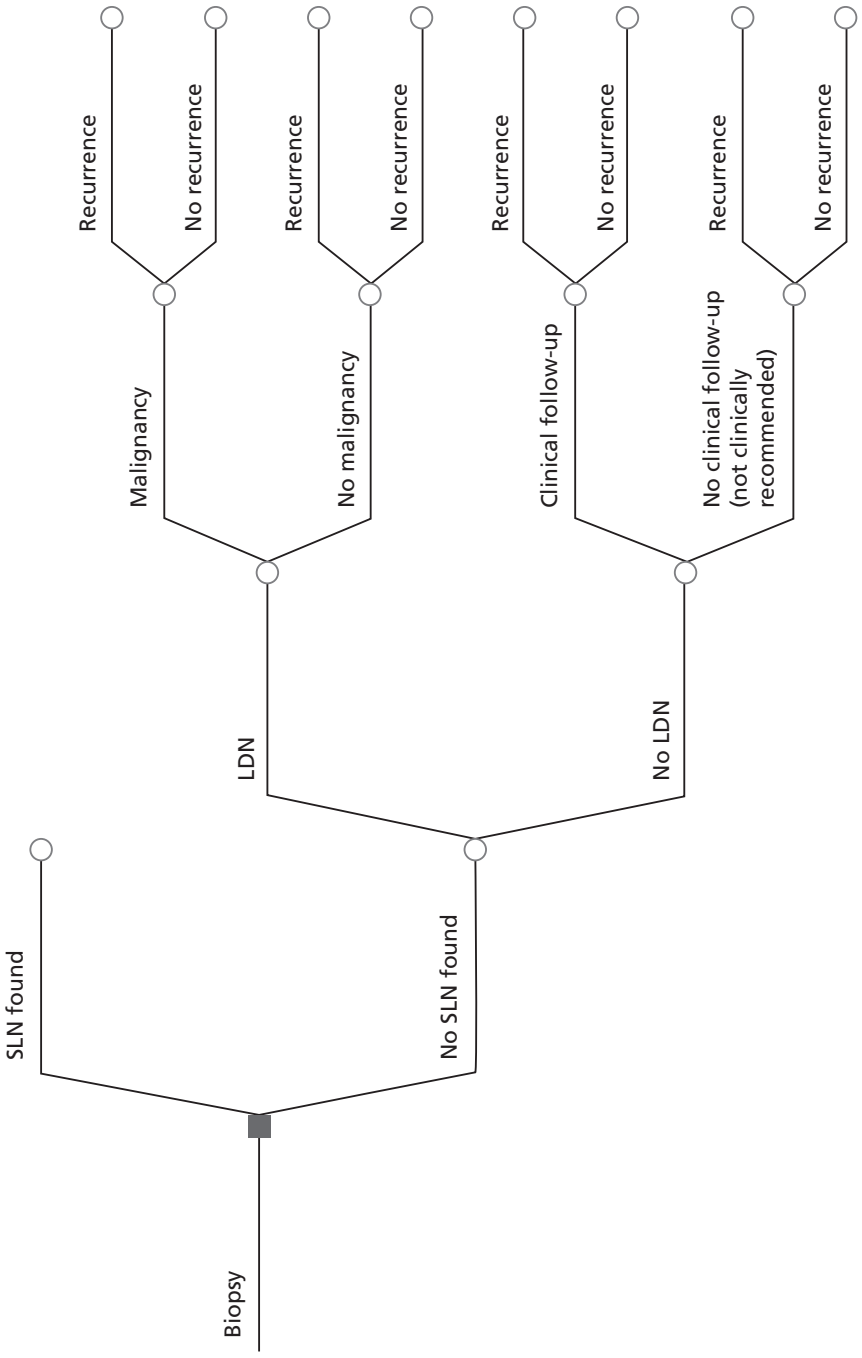


FIGURE 2 Decision tree for SLN biopsy (part 2).

Chapter 4 Systematic review methods

Protocol development and overview of review methods

A protocol was developed for undertaking systematic reviews of test accuracy, diagnostic and therapeutic impact and effectiveness of treatment for vulval cancer (see *Appendix 1*). Scoping searches for relevant systematic reviews were conducted in MEDLINE, EMBASE and The Cochrane Library [systematic reviews, HTA, Database of Abstracts of Reviews of Effects (DARE)] (see *Appendix 2*).

Systematic reviews were carried out using established methods in line with the recommendations of the NHS Centre for Reviews and Dissemination and the Cochrane Collaboration,[42,43] including those of the Cochrane Methods Working Group on Screening and Diagnostic Tests.[44] Presentation of results is according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines.[45]

Inclusion of studies, data extraction and quality assessment were carried out in duplicate, with differences resolved by consensus and/or arbitration involving a third reviewer. The selection process was piloted by applying the inclusion criteria to a sample of papers first. A two-stage process was then followed. First titles and abstracts were screened, then, for all references categorised as 'include' or 'uncertain' by both reviewers, the full text was retrieved wherever possible and final inclusion decisions were made based on the full paper. Reference Manager version 12.0 (Thomson Reuters, New York, NY, USA) software was used to construct a database of citations for all systematic reviews.

Clinical, methodological and statistical data extraction was conducted into data extraction sheets by at least two reviewers and discrepancies were resolved through discussion. If consensus could not be reached, disagreements were resolved by arbitration by a third reviewer. For diagnostic studies, information regarding study design and methods, characteristics of participants, SLN biopsy and comparison tests, and outcomes of interest were extracted using data extraction forms (see *Appendix 3*). For the effectiveness review, separate data extraction forms were used for different study designs: comparative experimental study (part A), comparative observational study (part B) and non-comparative study (part C). The data extraction sheets used are shown in *Appendix 4*. The quality assessment questions for randomised controlled trials (RCTs) were included in the data extraction sheet, but a separate form was used for case series (see *Appendix 4*). Data extraction was managed with Microsoft Word 2003 (Microsoft Corporation, Redmond, WA, USA) and Microsoft Excel 2003 (Microsoft Corporation, Redmond, WA, USA). Quality was also assessed independently by two reviewers. Any disagreements were resolved through discussion or by arbitration by the third reviewer. For each review a comprehensive master database of articles was constructed using Reference Manager 12.0 software.

Methods for test accuracy, diagnostic and therapeutic impact review

Search strategy

A sensitive search was conducted to identify all relevant published and unpublished studies and studies in progress. All databases were searched from inception to January 2011. Search strategies were designed from a series of test searches and discussions of the results of those searches among the review team. Both medical subject heading (MeSH) terms and text words were used and included 'vulva cancer', 'sentinel lymph node biopsy' and 'lymphoscintigraphy'. Search strategies can be found in *Appendix 5*. Literature was identified from several sources, including:

- General health and biomedical databases: MEDLINE (Ovid), EMBASE (Ovid), Science Citation Index; and medical diagnostic studies database (MEDION).
- Checking of reference lists of review articles and papers.
- Specialist search gateways [Organizing Medical Networked Information (OMNI) and The National Cancer Institute], general search engine (Google) and metasearch engine (Copernic).
- Searching a range of relevant databases including Clinical Trials.com and UK Clinical Research Network Portfolio to identify information about studies in progress, unpublished research or research reported in the grey literature.
- Hand-searching of *Gynecologic Oncology* journal (1980 to January 2011).
- Contact with authors of the included studies for information on any relevant published or unpublished studies.

Inclusion and exclusion criteria

Population

Included were:

- women with early stages of vulval cancer: at least 75% of population with FIGO stage I and II or TNM categories T1–2, N0, M0.

Excluded were:

- all patients with vulval melanomas
- advanced stage vulval cancer (FIGO stage IV), inoperable tumours, tumours unsuitable for primary surgery
- patients with clinical suspicion of metastases, i.e. with palpable inguinofemoral lymph nodes, enlarged lymph nodes (> 1.5 cm) on imaging or cytologically proven inguinofemoral lymph node metastases at the start of the study.

Index tests and comparator tests

Included were:

- SLN biopsy with ^{99m}Tc, blue dye or combined technique (^{99m}Tc with blue dye), with histopathology by frozen section or other routine histopathological techniques. Where studies reported any of ultrastaging, serial sections, multiple slices, additional sections or step sections, these were all classified as ultrastaging.

Excluded were:

- imaging modalities such as ultrasonography
- novel techniques such as reverse transcriptase-polymerase chain reaction (RT-PCR).

Reference standard

Included were:

- histopathology of inguinofemoral node dissection
- follow-up for groin recurrence.

Excluded were:

- imaging modalities such as ultrasonography.

Outcomes

Included were:

- diagnostic accuracy
- diagnostic impact: change in staging after SLN biopsy
- therapeutic impact: change in treatment plan including avoidance of full IFL after SLN biopsy
- complications
- morbidity
- mortality and disease-free survival
- quality of life (QoL)
- impact on surgeon's and team's skills and experience (learning curve).

Excluded were:

- non-clinical outcomes
- outcomes reported per groin only.

Study design

Included were:

- any prospective or retrospective test accuracy study designs
- studies investigating the diagnostic and therapeutic impact with or without concurrent assessment of test accuracy
- prospective cohort studies of outcomes of patients tested with ^{99m}Tc, blue dye or combined technique for SLN biopsy.

Excluded were:

- case studies
- studies with 10 or fewer patients.

Quality assessment

Test accuracy quality assessment followed the quality of diagnostic accuracy studies (QUADAS) guidelines[46] and diagnostic and therapeutic impact followed those suggested by Meads and Davenport.[47] The items of methodological quality listed in the QUADAS guidelines are representative spectrum, selection criteria clearly described, acceptable reference standard, acceptable delay between tests, partial verification avoided, differential verification avoided, reference standard independent of the index test, index test described in sufficient detail, reference standard described in sufficient detail, index test results blinded, reference standard results blinded, relevant clinical information available, uninterpretable results reported and withdrawals explained.[46]

These were tailored to assess the included studies because different aspects of quality are applicable to different topic areas. The actual quality items used for this report are listed below. For acceptable delay between tests, this included delay between the index test and reference standard (within 1 month). There will inevitably be a delay between index test and clinical follow-up (when available). Study quality was summarised in a table (see *Table 10*). No additional issues were thought to be useful in interpretetion of the results of these studies. The following items were included in study summaries and assessed using the three criteria listed under each item.[44]

1. Representative spectrum

Yes: If patients were women in early-stage squamous cell vulval cancer.

No: If a few patients were in a higher stage of vulval cancer (T3–T4) or some patients had vulval melanoma or another type of cancer rather than squamous cell cancer.

Unclear: If there is insufficient information available to make a judgement about the spectrum of patients.

2. Selection criteria clearly described

Yes: If the selection criteria are described.

No: If the selection criteria are not described.

Unclear: If there is insufficient information available to clearly know the selection criteria.

3. Acceptable reference standard

Yes: Whether or not the reference standard used (histopathology, clinical follow-up) was adequately described to permit sufficient replication and was appropriate according to advice from our clinical experts (e.g. ultrastaging used, immunohistochemistry used).

No: The reference standards used do not include ultrastaging or immunohistochemistry.

Unclear: It is unclear exactly what reference standard was used.

4. Acceptable delay between sentinel lymph nodes biopsy and histopathology

Yes: If the time between tests was shorter than 1 month, at least for an acceptably high proportion of patients.

No: If the time between tests was longer than 1 month for an unacceptably high proportion of patients.

Unclear: If information on timing of tests was not provided.

5. Partial verification avoided

Yes: If all patients, or a random selection of patients, who received the index test went on to receive verification of their disease status using a reference standard, even if the reference standard was not the same for all patients.

No: If some of the patients who received the index test did not receive verification of their true disease state, and the selection of patients to receive the reference standard was not random.

Unclear: If this information is not reported by the study.

6. Differential verification avoided

Yes: If the same reference standard was used in all patients.

No: If the choice of reference standard varied between individuals.

Unclear: If it is unclear whether or not different reference standards were used.

7. Incorporation avoided

Yes: If the index test did not form part of the reference standard.

No: If the reference standard formally included the result of the index test.

Unclear: If it is unclear whether or not the results of the index test were used in the final diagnosis.

8. Whether or not there was sufficient information to replicate index test and reference standard

Yes: Sufficient information available.

No: Insufficient information available.

Unclear: If it is unclear whether or not there is enough information to permit replication.

9. Reference standard/index test results blinded

Yes: If test results (index or reference standard) were interpreted blind to the results of the other test or blinding is dictated by the test order.

No: If it is clear that one set of test results was interpreted with knowledge of the other.

Unclear: If it is unclear whether or not blinding took place.

10. Relevant clinical information

Yes*:* If the same clinical data available when test results were interpreted as would be available when the test is used in practice.

No: If clinical data usually available was withheld or if more information than is usually available was provided.

Unclear: If information about the clinical data available was not stated.

11. Uninterpretable results reported

Yes: If the number of uninterpretable test results is stated or if the number of results reported agrees with the number of patients recruited (indicating no uninterpretable test results).

No: If it states that uninterpretable test results occurred or were excluded and does not report how many.

Unclear: If it is not possible to work out whether or not uninterpretable results occurred.

12. Withdrawals explained

Yes: If it is clear what happened to all patients who entered the study, for example, if a flow diagram of study participants is reported explaining any withdrawals or exclusions, or the numbers recruited match those in the analysis.

No: If it appears that some of the patients who entered the study did not complete the study, i.e. did not receive both the index test and reference standard, and these patients were not accounted for.

Unclear: If it is unclear how many patients entered and, hence, whether or not there were any withdrawals.

Methods of statistical analysis

RevMan version 5.0 (The Cochrane Collaboration, The Nordic Cochrane Centre, Copenhagen, Denmark) was used in the statistical analyses and Meta-Disc statistical package version 1.4 (Javier Zamora, Madrid, Spain) was used to conduct meta-analysis. Sensitivity, specificity, true-positives (TPs), false-positives (FPs), true-negatives (TNs) and false-negatives (FNs) were taken directly from the source papers. If that was not possible, values were calculated from data provided. Based on an investigation of heterogeneity, summary estimates of sensitivity, specificity and likelihood ratios (LRs) were derived as appropriate. Results are displayed graphically on summary receiver operator curve plots (see *Figure 4*). Summary SLN detection rates and their 95% CIs were calculated using Meta-Disc. Methods for meta-analysis used by Meta-Disc are as follows. Sensitivity and specificity are pooled by the formulae:

$$\mathrm{Sen}_T = \frac{\sum_i a_i}{\sum_i D_i} \tag{1}$$

$$\mathrm{Spe}_T = \frac{\sum_i d_i}{\sum_i ND_i} \tag{2}$$

where *a* are TPs and *d* are TNs, *D* is total number with disease and *ND* is the total number without disease. These formulae correspond to weighted averages in which the weight of each study is its sample size. The CIs of sensitivity and specificity are calculated using the *F* distribution method to compute the exact confidence limits for the binomial proportion (*x*/*n*) and are given by the formulae below where LL is the lower limit and UL is the upper limit:

$$\mathrm{LL} = \left(1 + \frac{n - x + 1}{xF_{2x,\,2(n-x+1),\,1-\alpha/2}}\right)^{-1} \tag{3}$$

$$\mathrm{UL} = \left(1 + \frac{n - x}{(x+1)F_{2(x+1),\,2(n-x),\,\alpha/2}}\right)^{-1} \tag{4}$$

Bivariate meta-analysis can only be conducted when there are more than four studies. Only one group of studies were eligible [IFL for all, ^{99m}Tc with blue dye – ultrastaging with immunohistochemistry (see *Table 16*)]. However, the diagnostic test results for all of the studies have no FPs, so STATA (version 12.1; StataCorp, College Station, TX, USA) will only run a bivariate meta-analysis if a continuity correction is added (changing 0 to 1 in some of the studies). This was done for the last five studies (see *Table 16*).

Methods for effectiveness reviews

Search strategy

A sensitive search was conducted to identify all relevant published and unpublished trials and trials in progress. All databases were searched from inception to January 2011. Search strategies were designed from a series of test searches in a multistep process. Both MESH terms and text words were used and included a variety of synonyms for vulval cancer and the interventions (surgery, RT, chemotherapy). Search strategies can be found in *Appendix 6*. Studies were identified from several sources, including:

- General health and biomedical databases: MEDLINE (Ovid), EMBASE (Ovid).
- Specialist electronic databases: The Cochrane Library, Cochrane Database of Systematic Reviews (CDSR), Cochrane Central Register of Controlled Trials (CENTRAL), DARE and HTA Database.
- Checking of reference lists of review articles and papers.

- Searching a range of relevant databases including ClinicalTrials.com and UK Clinical Research Network Portfolio to identify information about studies in progress, unpublished research or research reported in the grey literature.
- Hand-searching (*Gynecologic Oncology*) from 1980 to January 2011.
- Specialist search gateways (OMNI and the National Cancer Institute), general search engine (Google) and metasearch engine (Copernic) in January 2011.

Inclusion and exclusion criteria

Population

Included were:

- women with early stages of vulval cancer (including squamous cell carcinoma, basal cell carcinoma, adenocarcinoma, vulval Paget's disease, Bartholin's gland cancer): at least 75% of population with FIGO stage I and II or TNM categories T1–2, N0–1, M0.

Excluded were:

- all patients with vulval melanomas or VIN only
- patients with advanced vulval cancer, inoperable tumours and those unsuitable for primary surgery
- patients with clinical suspicion of metastases, i.e. with palpable inguinofemoral lymph nodes, enlarged lymph nodes (> 1.5 cm) on imaging or cytologically proven inguinofemoral lymph node metastases at the start of the study only
- patients with multifocal tumours only
- studies with 25% or more patients with clinical stages more advanced than FIGO stages I and II or TNM T1–2, N0–1, M0, unless the subgroup with these characteristics were clearly indicated and results given separately
- studies with all patients treated before 1980.

Intervention

Included were:

- surgery: vulvectomy (any form, with or without IFL)
- RT (any type, to vulval area or groin).

Excluded were:

- diagnostic treatment studies.

Comparator (when available)

Included were:

- surgery (any form) with RT (adjuvant or neoadjuvant) or chemotherapy.

Excluded were:

- same surgery as intervention. We did not include studies comparing different types of vulval excision for vulval cancer, as this was not relevant to the primary question to be addressed.

Outcomes

Included were:

- deaths, overall survival, disease-free survival (presented as raw numbers, survival curves, etc.)
- morbidity
- recurrence
- QoL
- early and late complications.

Excluded were:

- psychosexual outcomes.

Study design

Included were:

- RCTs
- non-RCTs
- observational studies (cohort, case–control or case series).

Excluded were:

- studies with five or fewer patients in the therapeutic group
- studies in which the majority of patients were enrolled in 1970s or earlier.

Quality assessment

Quality assessment was performed appropriate to study designs. For RCTs, quality assessment was according to the *Cochrane Handbook for Systematic Reviews of Interventions*[43] (*Table 5*). In all cases, a

TABLE 5 The Cochrane Collaboration's tool for assessing risk of bias

Section	Description	Question
Sequence generation	Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether or not it should produce comparable group	Adequate sequence generation?
Allocation concealment	Describe the method used to generate the allocation sequence in sufficient detail to determine whether or not intervention allocations could have been foreseen in advance of, or during, enrolment	Allocation concealment?
Blinding of participants, personnel and outcome assessors	Describe all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provide any information relating to whether or not the intended blinding was effective	Blinding? (Self-reported outcomes) Blinding? (Objective outcomes)
Incomplete outcome data	Describe the completeness of outcome data for each main outcome, including attrition and exclusions from analysis. State whether or not attrition and exclusions were reported, the numbers in each intervention group (compare with total randomised participants), reasons for attrition/exclusions when reported and any reinclusions in analyses performed by the review authors	Incomplete outcome data addressed?
Selective outcome reporting	State how the possibility of selective outcome reporting was examined by the review authors and what was found	Free of selective reporting?
Other sources of bias	State any important concerns about bias not addressed in the other domains in the tool	Free of other bias?

'yes' answer indicated a low risk of bias and a 'no' indicated a high risk of bias. 'Unclear' was used if details were insufficient. Quality of studies was summarised in tables (see *Tables 34* and *35*). Case–control studies were evaluated using the Newcastle–Ottawa Scale[48] (*Table 6*). A study was awarded with maximum one star [*] for each numbered item within the 'selection' and 'exposure' categories and a maximum of two stars [**] in the 'comparability' category. Each evaluated study could obtain a maximum of nine stars (four for the selection part, two for the comparability part and three for the exposure part). Qualitative description was also used. The detailed coding manual for this scale is in *Appendix 7*. Quality assessment of case series used criteria from a recent HTA report on methodological characteristics of case series.[49] A checklist composed of 13 items in five categories was used, and this is reproduced in *Appendix 8*.

Methods of statistical analysis

Separate analyses were performed on randomised and observational studies. RevMan version 5.0 was used in the statistical analyses. Information was analysed based on the group to which the participants were allocated, regardless of whether or not they received the allocated intervention. For dichotomous data, results are presented as summary RR with 95% CI (for comparative observational studies odds ratios were calculated when appropriate). For case–control studies and case series, a narrative summary of the findings is presented along with the numerical results.

TABLE 6 The Newcastle–Ottawa Scale for the quality assessment of case–control studies

Section	Number	Question
Selection	1.	Is the case definition adequate?
	2.	Representativeness of the cases
	3.	Selection of controls
	4.	Definition of controls
Comparability	1.	Comparability of cases and controls on the basis of the design or analysis
Exposure	1.	Ascertainment of exposure
	2.	Same method of ascertainment for cases and controls
	3.	Non-response rate

Chapter 5 Diagnostic review

Study selection

From the searches, 2942 citations were identified, of which 82 full papers were obtained. Included were 26 relevant studies (38 publications) (*Figure 3*). Excluded full-text articles are listed in *Appendix 9* with reasons for exclusion, which were mostly because of small sample size or type of publication (reviews, abstracts). Some studies were excluded because they gave results only per groin rather than per patient.

Characteristics of included studies

Index tests and histopathological techniques used for the index tests and reference standards used in each of the studies are given in *Table 7*. Although both ^{99m}Tc and blue dye were used in a number of studies, how the results were presented varied considerably. In some studies, only one was used. For example, De Cicco *et al.*,[55] Merisio *et al.*[64] and Vidal-Sicart *et al.*[74] used only ^{99m}Tc and Levenback *et al.*[61] used only blue dye. In six studies,[39,52,62,66,67,69] a proportion of SLNs were diagnosed with blue dye or ^{99m}Tc separately and a proportion with both (see *Table 7*); the results for malignancy were given for the whole cohort irrespective of the test or tests actually used to find the malignancy. In such cases, only the sensitivity and specificity results can be given for the combination of tests used rather than only blue dye or ^{99m}Tc separately or only both used together in all patients. However, for the other 20 studies,[50,51,53–61,63–65,68–74] detection rates per groin can be given for each test separately and both tests combined (see *Table 7*). It is noticeable that the histopathological techniques used for the full IFL specimens were either not given or were less detailed than those used for the SLNs. Only De Cicco *et al.*,[55] Johann *et al.*[59] and Radziszewski *et al.*[68] appeared to use the same techniques and very little detail is given in the first two. More details of index tests and reference standards used are given in *Appendix 10*.

Details of included studies and baseline characteristics are presented in *Tables 8* and *9*. The studies were conducted in a variety of European countries and in the USA and Canada. The majority were small and from single centres. The largest was a recent multicentre study by Van der Zee *et al.* from the Netherlands [the GROningen INternational Study on Sentinel nodes in Vulvar cancer (GROINSS-V)],[73] which included 403 patients recruited between the years 2000 and 2006. Inclusion and exclusion criteria of the studies are given in *Appendix 10*.

Data collection was prospective in 19 studies and retrospective in seven. Patients were recruited consecutively in 11 studies, prospectively in nine and retrospectively in two. Achimas-Cadariu *et al.*[50] described their study as retrospective, but data were collected prospectively from an in-house tumour registry. Recruitment dates varied between 1990 and 2008 and were not given in two studies[51,67] (see *Table 8*). The percentage of patients with early-stage disease varied between 86% and 100%, being 100% in 16 studies. Median or mean ages varied between 58 and 75 years and individual ages varied between 18 and 95 years. Medians were given in most studies as vulval cancer is relatively rare in younger women. Where reported, tumour locations were relatively evenly spread between midline or lateral positions. The most commonly reported tumour type was SCC. Five studies included one or two melanomas[53,58,61,63,71] and seven included other tumour types.[39,57,61,66,71,72,74] Either TNM, FIGO or grade staging, alone or in combination, was given in all studies. Most included patients with disease of varying severity and a few only included patients with early-stage disease, such as Terada *et al.*[70]

Quality of included studies

Quality assessment is reported in *Table 10*. Of the 26 studies included, four[53,66,67,71] provided no information about the histological staining method used. Brunner *et al.*,[52] Camara *et al.*,[53]

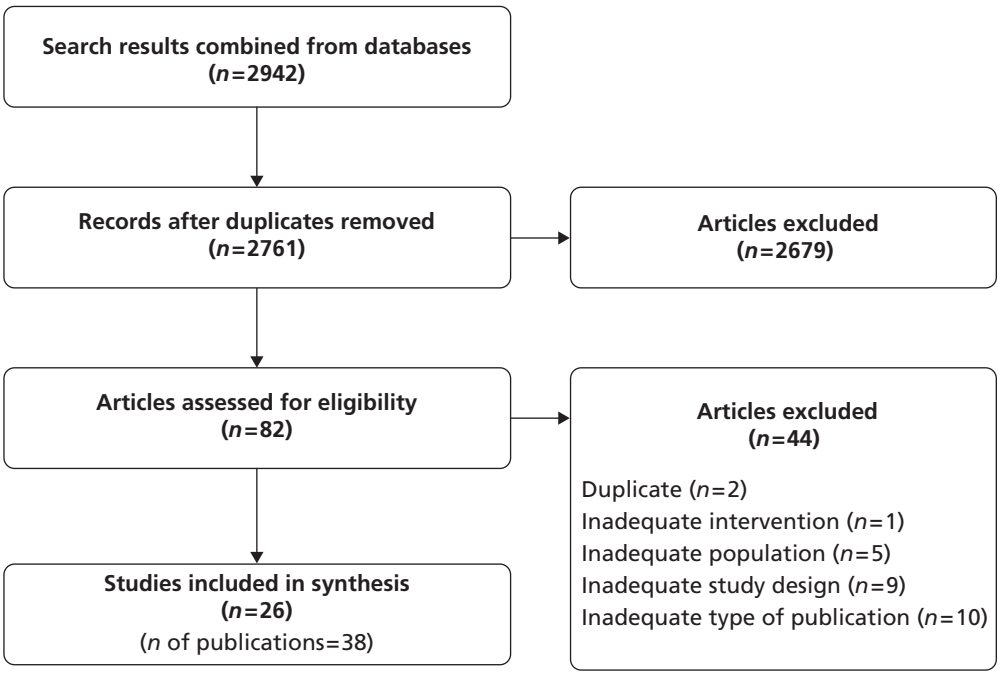


FIGURE 3 A PRISMA diagram for diagnostic review.

Hauspy *et al.*[58] and Rob *et al.*[69] used frozen section as the reference standard, rather than more routine histopathological techniques. In 16 studies, on receipt of negative results by H&E procedures, immunohistochemical tests using specific protein antibodies such as AE1, AE3, S-100, human melanoma black monoclonal antibody (HMB)-45, monoclonal antibody, cytokine myocyte nuclear factor (CKMNF), cytokine (CK)-88 and epithelial membrane antigen were conducted. In others, ultrastaging was used if samples were negative by H&E staining and standard sectioning. The thickness of slices varied from one study to another so that some studies were more likely to find small metastatic deposits than others because of the thinner sections taken.

There were four studies[50,65,70,73] in which, if the SLN was found to be negative, no IFL was performed but patients were followed up clinically instead. In a study by Van den Eynden *et al.*,[72] 10 out of 32 patients had a SLN biopsy plus full IFL. In the remaining 22 patients, an IFL was performed only if the SLN was positive or not found. In a study by Johann *et al.*[59] and another by Vidal-Sicart *et al.*,[74] some patients had SLN biopsy and full IFL regardless of node statistics and some only had IFL if the SLN was positive, but the results for the two groups were reported separately. Only the results for SLN biopsy and full IFL are reported here. In Crosbie *et al.*,[54] Klat *et al.*,[60] Martine-Palonez *et al.*,[63] Vakselj and Bebar[71] and Vidal-Sicart *et al.*,[74] clinical follow-up was reported, and, for all except Vidal-Sicart *et al.*,[74] this was reported according to whether patients had been SLN positive or negative.

Because the main aim of the included studies was the analysis of diagnostic procedures, most did not report information about the number of patients who had undergone specific types of surgery or other treatment procedures (see *Appendix 10, Table 53*, for treatment descriptions). Usually, patients underwent radical vulvectomy, wide local excision or hemivulvectomy. RT was performed in only six studies (as adjuvant therapy in Hauspy *et al.*,[58] Levenback *et al.*,[61] Moore *et al.*,[65] Vakselj and Bebar[71] and Van der Zee *et al.*[73] or as palliative treatment in Terada *et al.*[70]). Additionally, a study by Levenback *et al.*[61] mentioned that, in one patient in whom SLN was grossly positive after SLN biopsy, the surgeon aborted IFL in favour of RT. Adverse events (AEs) were reported in five studies[54,65,70,72,73] (see *Appendix 10*, *Table 55*).

With regard to blinding of index and reference test results, it would have been possible for the SLN and the full IFL nodes to be examined by different pathologists blind to each other's reports, but only

TABLE 7 Index test and reference standard details

Study	^{99m}Tc	Blue dye	Both together	Histopathological techniques: SLN	Histopathological techniques: remaining nodes	Type of surgery given (or RT)
Achimas-Cadariu *et al.*, 2009[50]	–	–	✗	H&E, ultrastaging	NR	Radical vulvectomy (58%), modified (41%)
Basta *et al.*, 2005[51]	✗[a]	✗[a]	✗	SLN immunochemical stain for micrometastases	NR	NR
Brunner *et al.*, 2008[52]	✗ (91%)	–	✗ (9%)	Frozen sections, H&E and, if negative, immunohistochemistry for cytokeratins	Routine techniques	NR
Camara *et al.*, 2009[53]	✗	✗	✗	Frozen section	NR	NR
Crosbie *et al.*, 2010[54]	–	–	✗	H&E and, if negative, with additional sections and immunohistochemistry for cytokeratins AE1–3	NR	Radical excision (47%), unclear (53%)
De Cicco *et al.*, 2000[55]	✗	–	–	H&E	H&E	Wide radical excision, hemivulvectomy or radical vulvectomy (percentages not given)
de Hullu *et al.*, 2000[56]	–	✗	✗	H&E and, if negative, with additional sections and immunohistochemistry for cytokeratins AE1–3	H&E	Radical excision (100%)
Hampl *et al.*, 2008[57]	✗	✗	✗	H&E and, if negative, with additional sections and immunohistochemistry for panycytokeratin antibody	NR	Hemivulvectomy (35%), vulvectomy (35%), local tumour resection (30%)
Hauspy *et al.*, 2007[58]	✗	–	✗	Frozen section then serial sections H&E and immunohistochemistry for cytokeratins AE1–3 for some sections	H&E	Wide local excision (76%), radical vulvectomy (20%), RT (5%)
Johann *et al.*, 2008[59]	–	–	✗	Step sectioning	Step sectioning	Radical vulvectomy (30%), hemivulvectomy (57%), wide excision (13%)

continued

TABLE 7 Index test and reference standard details (*continued*)

Study	^{99m}Tc	Blue dye	Both together	Histopathological techniques: SLN	Histopathological techniques: remaining nodes	Type of surgery given (or RT)
Klat *et al.*, 2009[60]	–	–	✗	H&E, ultrastaging and immunohistochemistry for cytokeratins AE1–3	NR	Radical surgery (100%)
Levenback *et al.*, 2001[61]	–	✗	–	Frozen section if suspicious, step sectioning and some immunohistochemistry using several protocols	NR	NR
Lindell *et al.*, 2010[39]	–	✗ (22%)	✗ (78%)	Step sections, H&E and if negative, immunohistochemistry for cytokeratin MNF116	H&E	Vulvectomy (47%), hemivulvectomy (31%), wide local excision (22%)
Louis-Sylvestre *et al.*, 2006[62]	✗ (21%)	–	✗ (79%)	Serial sections, H&E and immunohistochemistry for cytokeratins AE1 and AE3	NR	NR
Martinez-Palones *et al.*, 2006[63]	–	–	✗	0.2 mm sections, H&E and if negative, immunohistochemistry for cytokeratin and membrane epithelial antigen	NR	NR
Merisio *et al.*, 2005[64]	✗	–		H&E, ultrastaging, immunohistochemistry for cytokeratins in 50% of samples	Standard techniques	Radical vulvectomy or radical vulval excision (percentages NR)
Moore *et al.*, 2008[65]	–	–	✗	H&E and ultrastaging	NR	Radical vulvectomy or radical vulval excision (percentages NR)
Nyberg *et al.*, 2007[66]	–	✗ (20%)	✗ (80%)	Histopathology	NR	NR
Pityński *et al.*, 2003[67]	–	✗ (14%)	✗ (86%)	NR	NR	NR

Study	^{99m}Tc	Blue dye	Both together	Histopathological techniques: SLN	Histopathological techniques: remaining nodes	Type of surgery given (or RT)
Radziszewski *et al.*, 2010[68]	–	–	✗	Multiple slices, H&E in 50% slices, H&E and immunohistochemistry in the other 50% of slices	H&E in 50% slices, H&E and immunohistochemistry in other 50% slices	NR
Rob *et al.*, 2007[69]	–	✗ (27%)	✗ (73%)	Frozen section then serial sections, H&E, and immunohistochemistry on every third slide	H&E	NR
Terada *et al.*, 2006[70]	–	–	✗	Multiple slices, H&E then if negative, immunohistochemistry with cytokeratin antigen	NR	NR
Vakselj *et al.*, 2007[71]	–	–	✗	NR	NR	Tumour excised (100%)
Van den Eynden *et al.*, 2003[72]	–	–	✗	H&E then if negative, ultrastaging and immunohistochemistry for cytokeratins AE1 and AE3	NR	NR
Van der Zee *et al.*, 2008[73]	–	–	✗	Frozen section or routine histopathology, ultrastaging	H&E	Radical excision (100%)
Vidal-Sicart *et al.*, 2007[74]	–	–	✗	Multiple slices, H&E then if negative, H&E with immunohistochemistry	H&E	Radical vulvectomy or radical vulval excision (percentages NR)

NR, not reported.
a ^{99m}Tc and blue dye discrepant results in text and table.

TABLE 8 Characteristics of included studies

Study	Publications	Setting	Study design/patient selection	Recruitment dates
Achimas-Cadariu *et al.*, 2009[50]	Achimas-Cadariu *et al.*, 2009[50]	Dr Horst Schmidt Klinik, Wiesbaden, Germany	Retrospective/consecutive	June 2000 to May 2008
Basta *et al.*, 2005[51]	Basta *et al.*, 2005[51]	Department of Gynaecology, Obstetrics and Oncology, Jagiellonian University Medical College, Krakow, Poland	Prospective	NR
Brunner *et al.*, 2008[52]	Brunner *et al.*, 2008;[52] Sliutz *et al.*, 2002;[75] Hefler *et al.*, 2008[76]	Department of Obstetrics and Gynaecology at the Medical University of Vienna, Vienna, Austria	Retrospective/consecutive	January 2001 to August 2007
Camara *et al.*, 2009[53]	Camara *et al.*, 2009[53]	Friedrich-Schiller-University of Jena, Jena, Germany	Prospective	February 2003 to March 2007
Crosbie *et al.*, 2010[54]	Crosbie *et al.*, 2010[54]	Gynaecology Clinic at The Christie NHS Foundation Trust in Manchester, Manchester, UK	Prospective	2002–6
De Cicco *et al.*, 2000[55]	De Cicco *et al.*, 2000[55]	San Gerardo Hospital, Monza, Italy	Prospective/consecutive	May 1996 to September 1998
de Hullu *et al.*, 2000[56]	de Hullu *et al.*, 1998;[77] de Hullu *et al.*, 2000[56]	Groningen University Hospital, Groningen and Academic Hospital Vrije Universiteit, Amsterdam, the Netherlands	Prospective/consecutive	July 1996 to July 1999
Hampl *et al.*, 2008[57]	Hampl *et al.*, 2008[57]	Department of Gynaecology and Obstetrics of Heinrich Heine Universit (Dusseldorf), University of Jena, Medizinischen Hochschule Hannover and Women's Hospital, Regional Hospital of Altötting, Germany	Prospective/consecutive	2003–6
Hauspy *et al.*, 2007[58]	Hauspy *et al.*, 2007[58]	Sunnybrook Health Sciences Centre, University of Toronto, Toronto, ON, Canada	Prospective	April 2004 to September 2006
Johann *et al.*, 2008[59]	Johann *et al.*, 2008[59]	Department of Obstetrics and Gynaecology and of Nuclear Medicine, Inselspital, Bern University Hospital, Bern, Switzerland	Retrospective	January 1990 to March 2007

Study	Publications	Setting	Study design/patient selection	Recruitment dates
Klat *et al.*, 2009[60]	Klat *et al.*, 2009[60]	University Hospital Ostrava, Ostrava, Czech Republic	Prospective	May 2004 to November 2007
Levenback *et al.*, 2001[61]	Levenback *et al.*, 1994;[78] Levenback *et al.*, 1995;[79] Levenback *et al.*, 2001;[61] Frumovitz *et al.*, 2004[80]	Anderson Cancer Center and Southwestern Medical Centre, University of Texas, TX, USA	Prospective	1993–9
Lindell *et al.*, 2010[39]	Lindell *et al.*, 2010[39]	Department of Women's and Children's Health, of Pathology and Oncology and of Molecular Medicine and Surgery, Karolinska Institute, Stockholm, Sweden	Retrospective	2000–7
Louis-Sylvestre *et al.*, 2006[62]	Louis-Sylvestre *et al.*, 2006[62]	Centre Hospitalier Intercommunal Créteil, Créteil, France	Prospective	April 2002 to December 2005
Martinez-Palones *et al.*, 2006[63]	Martinez-Palones *et al.*, 2006[63]	Hospital Materno-infantil Vall d'Hebron, Barcelona, Spain	Prospective/consecutive	January 1995 to July 2005
Merisio *et al.*, 2005[64]	Merisio *et al.*, 2005[64]	Gynaecology Units of University of Parma and of Policlinico S Matteo of Pavia, Italy	Prospective	May 1999 to May 2003
Moore *et al.*, 2008[65]	Moore *et al.*, 2008[65]	Women and Infants' Hospital, Brown University, Providence, RI, USA	Prospective/consecutive	2002–7
Nyberg *et al.*, 2007[66]	Nyberg *et al.*, 2007[66]	Tampere University Hospital, Tampere, Finland	Retrospective	1 January 2001 to 30 June 2005
Pityński *et al.*, 2003[67]	Pityński *et al.*, 2003[67]	Department of Gynaecology and Obstetrics and of Nuclear Medicine, Jagiellonian University Medical College, Krakow, Poland	Prospective	NR
Radziszewski *et al.*, 2010[68]	Radziszewski *et al.*, 2010[68]	Maria Sklodowska-Curie Memorial Cancer Centre and Institute of Oncology, Warsaw, Poland	Prospective/consecutive	January 2002 to December 2006
Rob *et al.*, 2007[69]	Rob *et al.*, 2006;[81] Rob *et al.*, 2007[69]	Motol University Hospital, Second Medical Faculty, Charles University, Prague, Czech Republic	Prospective/consecutive	December 2001 to December 2005

continued

TABLE 8 Characteristics of included studies (*continued*)

Study	Publications	Setting	Study design/patient selection	Recruitment dates
Terada *et al.*, 2006[70]	Terada *et al.*, 2006[70]	Department of Obstetrics and Gynaecology and Cancer Research Center, University of Hawaii School of Medicine, and Department of Pathology, Queens Medical Centre, Honolulu, HI, USA,	Retrospective	1996–2003
Vakselj *et al.*, 2007[71]	Vakselj *et al.*, 2007[71]	Institute of Oncology Ljubljana, Ljubljana, Slovenia	Prospective/consecutive	March 2006 to December 2006
Van den Eynden *et al.*, 2003[72]	Van den Eynden *et al.*, 2003[72]	Department of Obstetrics-Gynaecology, University Hospital of Leuven, Leuven, Belgium	Retrospective	November 1999 to December 2002
Van der Zee *et al.*, 2008[73]	Van der Zee *et al.*, 2008;[73] Oonk *et al.*, 2009;[82] Oonk *et al.*, 2010[83]	15 centres registered at the University Medical Center Groningen, Groningen, Netherlands	Prospective/consecutive	March 2000 to May 2006
Vidal-Sicart *et al.*, 2007[74]	Vidal-Sicart *et al.*, 2002;[84] Vidal-Sicart *et al.*, 2007;[74] Vidal-Sicart *et al.*, 2009;[85] Puig-Tintore *et al.*, 2003[86]	Section of Gynecologic Oncology, Instituto Clínic de Ginecología, Obstetrícia y Neonatología, Hospital Clinic, University of Barcelona, Barcelona, Spain	Prospective	May 1998 to June 2005

NR, not reported.

TABLE 9 Baseline characteristics of patients

Study	*n*	Percentage of patients in early stage available for analysis	Age, years (median or mean, range)	Tumour location (*n*)	Histological type of tumour (*n*)	TNM (*n*)	FIGO stage (I–IV) (*n*)	Grade (1–3) (*n*)
Achimas-Cadariu *et al.*, 2009[50]	46	95%	Median 66, range 34–94	Midline: 17; lateral: 29	SCC: 46	T1–16, T2–28, T3–2	I: 16; II: 19; III: 7; IV: 4	1: 8; 2: 29; 3: 9
Basta *et al.*, 2005[51]	39	100%	NR	NR	NR	NR	All stage I and II	NR
Brunner *et al.*, 2008[52]	44	100%	Mean 70	Midline: 34; lateral: 10	SCC: 44	T1–30, T2–14/ N0–27, N1–17	NR	1: 14; 2: 27; 3: 3
Camara *et al.*, 2009[53]	17	94.1%	Median 75, range 37–83	NR	SCC: 16; melanoma: 1	T1–7, T2–9, T3–1, N1–9, N2–8	NR	NR
Crosbie *et al.*, 2010[54]	32	100%	Median 67, range 34–94	Midline: 17; lateral: 15	SCC: 32	NR	I: 7; II: 5; III: 4; IV: 3	NR
De Cicco *et al.*, 2000[55]	37	100%	NR	Midline: 19; lateral: 18	SCC: 37	T1–17, T2–20	NR	NR
de Hullu *et al.*, 2000[56]	59	100%	Median 69, range 33–92	NR	SCC: 59	T1–25, T2–34	NR	NR
Hampl *et al.*, 2008[57]	127	94.4%	Mean 61.4	Midline: 33; lateral: 49	SCC: 126; other: 1	T1–56, T2–62, T3–7, N1–88, N2 + N3–39	NR	1: 15; 2: 86; 3: 23
Hauspy *et al.*, 2007[58]	41	95%	Mean 65, range 34–92	NR	SCC: 39; melanoma: 2	T1–22, T2–19	NR	1: 18; 2: 17; 3: 4
Johann *et al.*, 2008[59]	23	86%	Median 68.4, range 34.1–86.5	NR	SCC: 23	T1–9, T2–11, T3–3, N1–11, N2–11, N3–1	NR	1: 3; 2: 14; 3: 6
Klat *et al.*, 2009[60]	23	100%	Median 67.5, range 38–92	Midline: 18; lateral: 5	SCC: 23	T1–11, T2–12	NR	NR

continued

TABLE 9 Baseline characteristics of patients (*continued*)

Study	*n*	Percentage of patients in early stage available for analysis	Age, years (median or mean, range)	Tumour location (*n*)	Histological type of tumour (*n*)	TNM (*n*)	FIGO stage (I–IV) (*n*)	Grade (1–3) (*n*)
Levenback *et al.*, 2001[61]	52	87%	Median 58, range 18–92	Midline: 25; lateral: 27	SCC: 35; melanoma: 7; other: 10	T1–22, T2–23, T3–7/N0–39, N1–9, N2–4	NR	NR
Lindell *et al.*, 2010[39]	77	98%	Mean 71.2, range 40–92	Midline: 22; lateral: 55	SCC: 77 (other: 1[a])	T1 + T2–76, T3–1	NR	1: 18; 2: 28; 3: 31
Louis-Sylvestre *et al.*, 2006[62]	38	100%	Mean 66, range 34–90	Midline: 26; lateral: 12	NR	T1–29, T2–9, N1–32, N2–6	NR	NR
Martinez-Palones *et al.*, 2006[63]	28	92.9%	Mean 71.3 ± 12, 7 (SD), range 30–84	NR	SCC: 26; melanoma: 2	T1–9, T2–19	I: 9; II: 19	1: 19; 2: 6; 3: 1
Merisio *et al.*, 2005[64]	20	100%	Mean 75, range 49–95	Midline: 11; lateral: 9	SCC: 20	T1–9, T2–11/N0–20	NR	NR
Moore *et al.*, 2008[65]	36	100%	Median 63, Mean 64, range 29–87	NR	SCC: 35	NR	I: 24; II: 8; III: 3; IV: 1	NR
Nyberg *et al.*, 2007[66]	47 (results for 25 stage I and II only)	100%	NR	NR	NR (full sample SCC: 46, other: 1)	NR	I: 11; II: 14	1: 15; 2: 8; 3: 2
Pityński *et al.*, 2003[67]	37	100%	NR	NR	NR	NR	All stage I or II	NR
Radziszewski *et al.*, 2010[68]	62	100%	Median 68, range 37–94	NR	SCC: 62	T1–20, T2–42, N1–62	NR	NR

Study	*n*	Percentage of patients in early stage available for analysis	Age, years (median or mean, range)	Tumour location (*n*)	Histological type of tumour (*n*)	TNM (*n*)	FIGO stage (I–IV) (*n*)	Grade (1–3) (*n*)
Rob *et al.*, 2007[69]	43	100%	Median 70.9, range 26–95	Midline: 21; lateral: 22	SCC: 43	T1–25, T2–18	NR	NR
Terada *et al.*, 2006[70]	21	100%	Mean 72, range 42–86	NR	SCC: 21	T1–21	NR	NR
Vakselj *et al.*, 2007[71]	35	92%	Median 65.8, range 36–88	NR	SCC: 32; melanoma: 1; other: 2	NR	I: 18; II: 6	1: 1; 2: 4; 3: 2[b]
Van den Eynden *et al.*, 2003[72]	32	100%	Mean 67, range 32–96	NR	SCC: 31; other: 1	T1 – 16, T2 – 16, N1 – 24, N2 + N3–8	NR	NR
Van der Zee *et al.*, 2008[73]	403	100%	NR	Midline: 151; lateral: 252	NR	T1 or 2–403, N0–276, N1–27	NR	NR
Vidal-Sicart *et al.*, 2007[74]	50	86%	Mean 75, range 41–95	NR	SCC: 50 (other: 8[a])	NR	Ib: 23; II: 20; III: 8	NR

NR, not reported; SD, standard deviation.
a Multifocal or advanced SCC tumour.
b Information only for positive SLNs biopsies.

TABLE 10 Quality of test accuracy studies

Study	Quality factors											
	1	2	3	4	5	6	7	8	9	10	11	12
Achimas-Cadariu *et al.*, 2009[50]	N	Y	N	Y	Y	Y	Y	N	U	Y	Y	Y
Basta *et al.*, 2005[51]	U	N	Y	Y	Y	Y	Y	N	U	N	Y	Y
Brunner *et al.*, 2008[52]	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	N
Camara *et al.*, 2009[53]	N	N	Y	Y	Y	Y	Y	N	U	Y	Y	Y
Crosbie *et al.*, 2010[54]	Y	Y	Y	Y	Y	N	Y	N	U	U	Y	Y
De Cicco *et al.*, 2000[55]	Y	Y	Y	Y	U	Y	Y	Y	N	N	Y	Y
de Hullu *et al.*, 2000[56]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Hampl *et al.*, 2008[57]	N	Y	Y	Y	U	Y	Y	N	U	Y	Y	N
Hauspy *et al.*, 2007[58]	N	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y
Johann *et al.*, 2008[59]	N	Y	U	Y	Y	Y	Y	N	U	Y	Y	N
Klat *et al.*, 2009[60]	Y	Y	Y	Y	Y	Y	Y	N	U	Y	Y	Y
Levenback *et al.*, 2001[61]	N	Y	Y	Y	Y	Y	Y	N	U	Y	Y	N
Lindell *et al.*, 2010[39]	N	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	N
Louis-Sylvestre *et al.*, 2006[62]	Y	Y	Y	U	Y	Y	Y	N	U	U	Y	N
Martinez-Palones *et al.*, 2006[63]	N	Y	Y	U	Y	Y	Y	N	U	Y	Y	Y
Merisio *et al.*, 2005[64]	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y
Moore *et al.*, 2008[65]	Y	Y	N	N	Y	N	Y	N	U	Y	Y	Y
Nyberg *et al.*, 2007[66]	Y	Y	U	Y	Y	Y	Y	N	U	U	Y	Y
Pityński *et al.*, 2003[67]	U	U	U	U	U	Y	Y	N	U	N	Y	Y
Radziszewski *et al.*, 2010[68]	Y	Y	Y	Y	Y	Y	Y	Y	U	U	Y	N
Rob *et al.*, 2007[69]	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	N
Terada *et al.*, 2006[70]	Y	Y	N	U	Y	N	Y	N	U	Y	Y	Y
Vakselj *et al.*, 2007[71]	N	N	U	Y	Y	N	Y	N	U	Y	Y	Y
Van den Eynden *et al.*, 2003[72]	Y	U	N	Y	Y	N	Y	N	U	Y	Y	N
Van der Zee *et al.*, 2008[73]	Y	Y	N	U	Y	N	Y	Y	U	Y	Y	N
Vidal-Sicart *et al.*, 2007[74]	N	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	N

N, no; U, unclear; Y, yes.
Quality factors assessed: 1, representative spectrum; 2, selection criteria clearly described; 3, acceptable reference standard; 4, acceptable delay between SLN biopsy and histopathology or clinical follow-up; 5, partial verification avoided; 6, differential verification avoided; 7, incorporation avoided; 8, sufficient information for replication; 9, reference standard/index test blinded; 10, relevant clinical information; 11, uninterpretable results reported; and 12, withdrawals explained.

de Hullu *et al.*[56] achieved this, and only De Cicco *et al.*[55] mentioned that they had not blinded pathologists. The remaining studies did not mention any blinding.

Test accuracy results

Results of test accuracy studies for which all patients had IFL as the reference standard are given in *Table 11* and for which IFL was the reference standard in test positives and clinical follow-up in test

TABLE 11 Results for test accuracy studies with IFL given to all as reference standard

Study	*n*	No. of patients with one or more SLNs found	No. of patients with SLN containing malignancies (frozen section)	No. of patients with SLN containing malignancies (histopathology)	No. with SLN found but negative and malignancy found at ILN	No. with no SLN found	No. with no SLN and malignancy from ILN (frozen section)	No. with no SLN and malignancy from ILN (histopathology)	No. with clinical follow-up	No. with malignancies at follow-up	Comments
Basta *et al.*, 2005[51]	39	38	–	12	0	1	–	0	0	–	–
Brunner *et al.*, 2008[52]	44	44	17	0	3	0	–	–	0	–	Numbers unclear
Camara *et al.*, 2009[53]	17	15	7	-	0	2	–	1	0	–	Both
	17	13	–	–	–	–	–	–	–	–	^{99m}Tc
	17	9	–	–	–	–	–	–	–	–	Blue
Crosbie *et al.*, 2010[54]	32	31	–	6	0	1	–	1	32	SLN negative: 0 SLN positive: 5	–
De Cicco *et al.*, 2000[55]	37	37	–	8	–	0	–	0	0	–	–
de Hullu *et al.*, 2000[56]	59	59	–	20	0	0	–	-	0	–	–
Hampl *et al.*, 2008[57]	127	125	0	36	3	2	–	0	0	–	Numbers unclear
Hauspy *et al.*, 2007[58]	41	39	15	0	0	2	0	1	0	–	–
Johann *et al.*, 2008[59]	23	23	?	10	1	0	–	0	0	–	Numbers unclear

continued

TABLE 11 Results for test accuracy studies with IFL given to all as reference standard (*continued*)

Study	*n*	No. of patients with one or more SLNs found	No. of patients with SLN containing malignancies (frozen section)	No. of patients with SLN containing malignancies (histopathology)	No. with SLN found but negative and malignancy found at ILN	No. with no SLN found	No. with no SLN and malignancy from ILN (frozen section)	No. with no SLN and malignancy from ILN (histopathology)	No. with clinical follow-up	No. with malignancies at follow-up	Comments
Klat *et al.*, 2009[60]	23	23	–	14	1	0	–	-	23	SLN positive: 3 SLN negative: 1	–
Levenback *et al.*, 2001[61]	52	46	–	11	0	6		2	0	–	Numbers unclear
Lindell *et al.*, 2010[39]	77	75	–	21	2	2	–	0	0	–	Numbers unclear
Louis-Sylvestre *et al.*, 2006[62]	38	36	–	11	–	2	–	?	0	–	Numbers unclear
Martinez-Palones *et al.*, 2006[63]	28	27	–	6	1	1	–	–	28	SLN positive: 4 SLN negative: 3? No SLN: 1	
Merisio *et al.*, 2005[64]	20	20	–	3	1	0	–	–	0	–	–
Nyberg *et al.*, 2007[66]	25 (47)[b]	25		4	0	0	–	0	0	–	Stage I and II only

Study	*n*	No. of patients with one or more SLNs found	No. of patients with SLN containing malignancies (frozen section)	No. of patients with SLN containing malignancies (histopathology)	No. with SLN found but negative and malignancy found at ILN	No. with no SLN found	No. with no SLN and malignancy from ILN (frozen section)	No. with no SLN and malignancy from ILN (histopathology)	No. with clinical follow-up	No. with malignancies at follow-up	Comments
Pityński *et al.*, 2003[67]	37	37	–	11	0	0	–	–	0	–	
Radziszewski *et al.*, 2010[68]	56	56	–	21	?	0	–	–	0	–	Six patients with metastases on FNA, numbers unclear
Rob *et al.*, 2007[69]	43	43	13	14	0	0	–	–	0	–	Both
	16	11	0	4	1	5	-	1	0	–	Blue only, numbers unclear
Vakselj *et al.*, 2007[71]	35	35	0	10	0	0	–	–	35	SLN positive: 6 SLN negative: 3	–
Vidal-Sicart *et al.*, 2007[74]	50	49	0	16	0	1	–	Not found	(70)[a]	Total: 5	–

?, unclear whether or not the value given is correct; FNA, fine-needle aspiration.

a This value is in brackets because it is more than the number in the study that could be used for this project. The authors included a further 20 patients and the clinical follow-up was combined for all 70.

b FIGO stages I–IV $n = 47$, results used only for stages I–II here $n = 25$.

TABLE 12 Results for test accuracy studies for which the reference standard was IFL for SLN positive and clinical follow-up for SLN negative

Study	*n*	No. of patients with 1 or more SLNs found	No. of patients with SLN with malignancies (frozen section)	No. of patients with SLN with malignancies (histopathology)	No. with SLN found but negative and malignancy found at ILN	No. with no SLN found	No. with no SLN and malignancy from ILN (frozen section)	No. with no SLN and malignancy from ILN (histopathology)	No. with clinical follow-up	No. with malignancies at follow-up	Comments
Achimas-Cadariu *et al.*, 2009[50]	46	43	–	–	–	3	–	–	46	8	Mostly per groin analysis
Moore *et al.*, 2008[65]	36	35	–	4	0	1	–	0	31 (negative SLN)	SLN negative: 2	–
Terada *et al.*, 2006[70]	21	21	0	3	0	0	–	–	21	SLN positive: 2 SLN negative: 0	–
Van den Eynden *et al.*, 2003[72]	32	27	–	10	5	5	–	?	17	SLN negative: 4	Numbers unclear
Van der Zee *et al.*, 2008[73]	403	403	–	127 (135 in Oonk *et al.* 2010[83])	36 (micromets from 33 immunohistochemistry of SLN, Oonk *et al.* 2010[83])	0	–	–	276	SLN negative: 8 groin, 34 local SLN positive: 11 groin, 34 local, 7 distant	Numbers unclear

negatives is given in *Table 12*. Most of the studies reported their results per groin rather than per patient; therefore, teasing out the results per patient was difficult in several of the papers (noted as numbers unclear in the comments column of *Tables 11* and *12*).

For calculation of sensitivity and specificity, studies have been categorised by the reference standards used, the index test used and the histopathological techniques used as follows:

- IFL for all
  - ^{99m}Tc with blue dye
    - H&E only or insufficient details to determine whether immunohistochemistry or ultrastaging were used (*Table 13*)
    - immunohistochemistry (*Table 14*)
    - frozen section only (*Table 15*)
    - immunohistochemistry with ultrastaging (*Tables 16* and *17*)
  - ^{99m}Tc only
    - H&E only or insufficient details to determine whether immunohistochemistry or ultrastaging were used (*Table 18*)
    - immunohistochemistry (*Table 19*)
  - blue dye only
    - immunohistochemistry (*Table 20*).
- IFL for SLN positive and clinical follow-up for SLN negative
  - ^{99m}Tc and blue dye
    - immunohistochemistry (*Table 21*)
    - ultrastaging (*Tables 22* and *23*).

TABLE 13 Inguinofemoral lymphadenectomy for all, ^{99m}Tc with blue dye: H&E only or insufficient details

Study	*n*	No. of patients with SLNs found	TP	FP	TN	FN	Sensitivity (95% CI)	Specificity (95% CI)
Nyberg *et al.*, 2007[66]	25	25	4	0	21	0	100% (39.8% to 100%)	100% (83.9% to 100%)
Pityński *et al.*, 2003[67]	37	37	11	0	26	0	100% (71.5% to 100%)	100% (86.8% to 100%
Vakselj *et al.*, 2007[71]	35	35	10	0	25	0	100% (69.2% to 100%)	100% (86.3% to 100%)
Pooled sensitivity = 100% (95% CI 86.3 to 100%); chi-squared test = 0.00 (degrees of freedom = 2); p = 1.000; I^2 = 0.0%.								
Pooled specificity = 100% (95% CI 95.0 to 100%); chi-squared test = 0.00 (degrees of freedom = 2); p = 1.000; I^2 = 0.0%.								

TABLE 14 Inguinofemoral lymphadenectomy for all, ^{99m}Tc with blue dye: immunohistochemistry

Study	*n*	No. of patients with SLNs found	TP	FP	TN	FN	Sensitivity (95% CI)	Specificity (95% CI)
Basta *et al.*, 2005[51]	39	38 (82%)	12	0	24	0	100% (73.5% to 100%)	100% (85.8% to 100%)

TABLE 15 Inguinofemoral lymphadenectomy for all, ^{99m}Tc with blue dye: frozen section

Study	*n*	No. of patients with SLNs found	TP	FP	TN	FN	Sensitivity (95% CI)	Specificity (95% CI)
Basta *et al.*, 2005[51]	17	15	7	0	8	0	100% (59.0% to 100%)	100% (63.1% to 100%)

TABLE 16 Inguinofemoral lymphadenectomy for all, ^{99m}Tc with blue dye: ultrastaging with immunohistochemistry

Study	*n*	No. of patients with SLNs found	TP	FP	TN	FN	Sensitivity (95% CI)	Specificity (95% CI)
Crosbie *et al.*, 2010[54]	32	31	6	0	25	0	100% (54.1% to 100%)	100% (86.3% to 100%)
de Hullu *et al.*, 2000[56]	59	59	20	0	39	0	100% (83.2% to 100%)	100% (91.0% to 100%)
Hampl *et al.*, 2008[57]	127	125	36	0	86	3	92.3% (79.1% to 98.4%)	100% (95.8% to 100%)
Hauspy *et al.*, 2007[58]	41	39	15	0	24	0	100% (78.2% to 100%)	100% (85.8% to 100%)
Johann *et al.*, 2008[59]	23	23	10	0	12	1	90.9% (58.7% to 99.8%)	100% (73.5% to 100%)
Klat *et al.*, 2009[60]	23	23	14	0	8	1	93.3% (68.1% to 99.8%)	100% (63.1% to 100%)
Lindell *et al.*, 2010[39]	77	75	21	0	52	2	91.3% (72.0% to 98.9%)	100% (93.2% to 100%)
Louis-Sylvestre *et al.*, 2006[62]	38	36	11	0	25	0	100% (71.5% to 100%)	100% (86.3% to 100%)
Martinez-Palones *et al.*, 2006[63]	28	27	6	0	20	1	85.7% (42.1% to 99.6%)	100% (83.2% to 100%)
Radziszewski *et al.*, 2010[68]	62	56	21	0	35?	0?	100% (83.9% to 100%)	100% (90.0% to 100%)
Rob *et al.*, 2007[69]	43	43	14	0	29	0	100% (76.8% to 100%)	100% (88.1% to 100%)

Pooled sensitivity = 95.6% (95% CI 91.5 to 98.1%); chi-squared test = 11.0 (degrees of freedom = 10); $p = 0.35$; $I^2 = 9.9\%$.

Pooled specificity = 100% (95% CI 99.0 to 100%); chi-squared test = 0.00 (degrees of freedom = 10); $p = 1.000$; $I^2 = 0.0\%$.

Negative predictive value = 97.8%; random-effects-positive LR = 51.368 (95% CI 22.440 to 117.586), negative LR = 0.088 (95% CI 0.053 to 0.146).

TABLE 17 Results from *Table 16* with continuity correction used for bivariate meta-analysis

Study	TP	FP	TN	FN
Crosbie *et al.*, 2010[54]	6	0	25	0
de Hullu *et al.*, 2000[56]	20	0	39	0
Hampl *et al.*, 2008[57]	36	0	86	3
Hauspy *et al.*, 2007[58]	15	0	24	0
Johann *et al.*, 2008[59]	10	0	12	1
Klat *et al.*, 2009[60]	14	0	8	1
Lindell *et al.*, 2010[39]	21	1	52	2
Louis-Sylvestre *et al.*, 2006[62]	11	1	25	0
Martinez-Palones *et al.*, 2006[63]	6	1	20	1
Radziszewski *et al.*, 2010[68]	21	1	35	0
Rob *et al.*, 2007[69]	14	1	29	0

Because bivariate meta-analysis in STATA requires that some of the FP results are greater than zero, the results were used with a continuity correction giving five of the FP results a value of 1 instead of zero. The calculated sensitivity was 0.96 (95% CI 0.91 to 0.98) and specificity 0.99 (95% CI 0.97 to 0.99). This shows very little difference from that calculated by Meta-Disc shown in *Table 16*.

TABLE 18 Inguinofemoral lymphadenectomy for all, ^{99m}Tc only: H&E only or insufficient details

Study	*n*	No. of patients with SLNs found	TP	FP	TN	FN	Sensitivity (95% CI)	Specificity (95% CI)
De Cicco *et al.*, 2000[55]	37	37	8	0	29	0	100% (63.1% to 100%)	100% (88.1% to 100%)

TABLE 19 Inguinofemoral lymphadenectomy for all, ^{99m}Tc only: immunohistochemistry with or without ultrastaging

Study	*n*	No. of patients with SLNs found	TP	FP	TN	FN	Sensitivity (95%CI)	Specificity (95%CI)
Brunner *et al.*, 2008[52]	44	44	17	0	24	3	85.0% (62.1% to 96.8%)	100% (85.8% to 100%)
Merisio *et al.*, 2005[64]	20	20	3	0	16	1	75.0% (19.4% to 99.4%)	100% (79.4% to 100%)
Vidal-Sicart *et al.*, 2007[74]	50	49	16	0	33	0	100% (79.4% to 100%)	100% (89.4% to 100%)

Pooled sensitivity = 90.0% (95% CI 76.3 to 97.2%); chi-squared = 4.60 (degrees of freedom = 2); $p = 0.10$; $I^2 = 56.5\%$.

Pooled specificity = 100% (95% CI 95.1 to 100%); chi-squared = 0.00 (degrees of freedom = 1); $p = 1.000$; $I^2 = 0.0\%$.

TABLE 20 Inguinofemoral lymphadenectomy for all, blue dye only: ultrastaging with immunohistochemistry

Study	*n*	No. of patients with SLNs found	TP	FP	TN	FN	Sensitivity (95% CI)	Specificity (95% CI)
Levenback *et al.*, 2001[61]	52	46	11	0	35	0	100% (71.5% to 100%)	100% (90.0% to 100%)
Rob *et al.*, 2007[69]	16	11	4	0	6	1	80.0% (28.4% to 99.5%)	100% (54.1% to 100%)

Pooled sensitivity = 93.8% (95% CI 69.8 to 99.8%); chi-squared test = 2.48 (degrees of freedom = 1); $p = 0.115$; $I^2 = 59.6\%$.

Pooled specificity = 100% (95% CI 91.4 to 100%); chi-squared test = 0.00 (degrees of freedom = 1); $p = 1.000$; $I^2 = 0.0\%$.

TABLE 21 Inguinofemoral lymphadenectomy for SLN positive, clinical follow-up for SLN negative, ^{99m}Tc and blue dye: ultrastaging and immunohistochemistry

Study	*n*	No. of patients with SLNs found	TP	FP	TN	FN	Sensitivity (95% CI)	Specificity (95% CI)
Moore *et al.*, 2008[65]	36	35	4	0	29	2	66.7% (22.3% to 95.7%)	100% (88.1% to 100%)
Terada *et al.*, 2006[70]	21	21	3	0	18	0	100% (29.2% to 100%)	100% (81.5% to 100%)
Van der Zee *et al.*, 2008[73]	403	403	127	0	234	42	75.1% (67.9% to 81.5%)	100% (98.4% to 100%)

Pooled sensitivity = 75.3% (95% CI 68.3 to 81.4%); chi-squared test = 1.93 (degrees of freedom = 2); $p = 0.381$; $I^2 = 0.0\%$.

Pooled specificity = 100% (95% CI 98.7 to 100%); chi-squared test = 0.00 (degrees of freedom = 2); $p = 1.000$; $I^2 = 0.0\%$.

TABLE 22 Inguinofemoral lymphadenectomy for SLN positive, clinical follow-up for SLN negative, ^{99m}Tc and blue dye: ultrastaging

Study	*n*	No. of *n* patients with SLNs found	TP	FP	TN	FN	Sensitivity (95% CI)	Specificity (95% CI)
Achimas-Cadariu *et al.*, 2009[50]	46	43 (94%)	NR	0	NR	0	Not calculable	Not calculable
Van den Eynden *et al.*, 2003[72]	32	27	10	0	12	5	66.7% (38.4% to 88.2%)	100% (73.5% to 100%)

NR, not reported.

TABLE 23 Inguinofemoral lymphadenectomy for SLN positive, clinical follow-up for SLN negative, ^{99m}Tc and blue dye: ultrastaging with or without immunohistochemistry, groin and distant recurrences only in node-negative patients

Study	*n*	No. of patients with SLNs found	TP	FP	TN	FN	Sensitivity (95% CI)	Specificity (95% CI)
Achimas-Cadariu *et al.*, 2009[50]	46	43 (94%)	NR	0	NR	0	Not calculable	Not calculable
Moore *et al.*, 2008[65]	36	35	4	0	29	2	66.7% (22.3% to 95.7%	100% (88.1% to 100%)
Terada *et al.*, 2006[70]	21	21	3	0	18	0	100% (29.%2 to 100%)	100% (81.5% to 100%)
Van den Eynden *et al.*, 2003[72]	32	27	10	0	12	5	66.7% (38.4% to 88.2%)	100% (73.5% to 100%)
Van der Zee *et al.*, 2008[73]	403	403	127	0	268	8	94.1% (88.7% to 81.5%)	100% (98.6% to 100%)

Pooled sensitivity = 90.6% (95% CI 84.9 to 94.6%); chi-squared test 11.90 (degrees of freedom = 3); $p = 0.008$; $I^2 = 74.8\%$.

Pooled specificity = 100% (95% CI 98.9 to 100%); chi-squared test 0.00 (degrees of freedom = 3); $p = 1.0000$; $I^2 = 0.0\%$.

NR, not reported.

The sensitivities and specificities in the results tables are based on the number of patients with one (or more) SLNs found, rather than the total number of patients. They could have had SLN biopsy in one or both groins; therefore, the results are giving the sensitivity and specificity of malignancy or none when a SLN (or more than one) has been found. It would not be possible for a SLN biopsy to be false-positive for malignancy, so all the point estimates of specificity are 100%. For this reason, only one receiver operating characteristic (ROC) plane is given, for the category of IFL for all, ^{99m}Tc with blue dye: ultrastaging and immunohistochemistry, in which there were 11 included studies (see *Tables 16* and *17*, *Figure 4*). All of the points are clustered along the top left-hand-side and the plot is not particularly informative. Unfortunately, there were insufficient studies of similar clinical characteristics to be able to conduct metaregression comparing ^{99m}Tc and blue dye results.

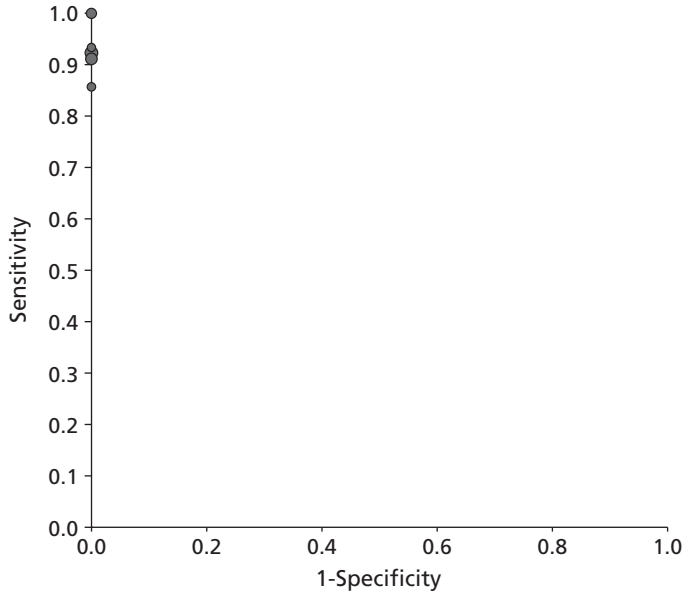


FIGURE 4 Receiver operating characteristic plane for IFL for all, ^{99m}Tc with blue dye: ultrastaging and immunohistochemistry.

The results of sensitivities and specificities show that, although the point estimates are either 100% or close to 100%, the CIs are wide, reflecting the small samples available. However, the pooled sensitivity of the studies with clinical follow-up was less than the ones with IFL for all.

In a test for cancer, it is important not to miss malignancy. If a FN diagnosis is made, the patient is at risk of developing malignancy in the groin or disseminated malignancy. If a FP diagnosis is made, the patient would undergo unnecessary IFL with resulting morbidity. So FNs may be considered to be relatively more important than FPs. In SLN biopsy, there would be no false-positive diagnoses unless an error was made in the histological examination, which would be rare, so we have to evaluate only the FN diagnoses. The results suggest that, if we evaluate SLN biopsy with clinical follow-up for node-negative patients, many more FN diagnoses will be made because of the longer time of follow-up, enabling more development of observable metastases. However, one would assume that the accuracy of a test would be a function of the test itself, rather than the method of evaluation. It is known that differential verification usually leads to increased estimates of sensitivity and specificity. The opposite is seen in this example and may be because there is more time to develop a recurrence with a long clinical follow-up. In addition, total recurrence rates were used rather than groin recurrence rates only because some studies only gave total recurrence rates. It is reasonable to assume that IFL should not influence recurrence in the vulval area. Therefore, the recurrence rates for clinical follow-up studies were recalculated, using only groin and distant recurrence, where given (only available in the study by Van der Zee *et al.*[73]). The results are shown in *Table 23* and show a higher pooled sensitivity, comparable to the estimates for which all patients received IFL as the gold standard.

The probability of curing vulval cancer is greatest when it is diagnosed at an early stage. The studies included patients with disease at a variety of FIGO stages; some included only patients with early-stage disease[51,63,67,71] and others patients with disease of all stages[50,54,65] (when reported). A study by Nyberg *et al.*[66] had a mixture of stages but reported results for each stage separately so results for stages I and II are reported here. There was a trend for the studies with early FIGO stage patients to have higher sensitivities than studies with mixed-FIGO-stage patients, but this was not consistent across all studies.

The vast majority of vulval cancer (> 90%) is SCC. Vulval melanoma is the second most common vulval malignancy but, unlike SCC, melanoma has a high risk of metastasis and the overall prognosis is poor. However, studies that included malignancies other than SCCs did not appear to have noticeably different results to those with SCCs only, but this may be a result of small sample sizes.

Sentinel lymph node detection rate

The accuracy of a diagnostic test such as SLN biopsy depends on the ability of the surgeon to identify the SLN; therefore, SLN detection rates for each of the analysed techniques (blue dye, ^{99m}Tc, blue dye/^{99m}Tc) are presented (*Table 24*). The detection rate calculated per patient was available in all included studies. Some of the studies also gave detection rates per groin for ^{99m}Tc and blue dye separately, but these are not reported here. The detection rates do not obviously vary by whether or not patient groups included non-SCCs, such as melanomas, or whether both early and late stage were included rather than early only. However, it is clear that blue dye detects fewer SLNs than ^{99m}Tc and that both used together is the most successful strategy.

Subpopulation of stages I and II

Two studies[66,74] gave detection rates for patients in the different FIGO stages separately (*Table 25*). The sample sizes are very small but suggest no obvious gradient of detection rate by FIGO stage.

Training and experience

There is a learning curve for surgeons performing SLN biopsy and IFL. Several of the included studies mention this (*Table 26*), but for most of the studies the learning curve was taken to mean that, after the

TABLE 24 Sentinel lymph node detection rate

Study	^{99m}Tc	Blue dye	Both together	Study characteristics
Achimas-Cadariu *et al.*, 2009[50]	–	–	43/46 (94%)	Only SCC, early and late stage
Basta *et al.*, 2005[51]	38/39[a] (97%)	32/39[a] (82%)	38/39 (97%)	Not reported, early stage
Camara *et al.*, 2009[53]	13/17 (76%)	9/17 (53%)	15/17 (88%)	Mostly SCC, early and late stage
Crosbie *et al.*, 2010[54]	–	–	31/32 (97%)	Only SCC, early and late stage
De Cicco *et al.*, 2000[55]	37/37 (100%)	–	–	Only SCC, early and late stage
de Hullu *et al.*, 2000[56]	–	35/59 (60%)	59/59 (100%)	Only SCC, early stage
Hampl *et al.*, 2008[57]	119/127 (94%)	80/127 (63%)	125/127 (98%)	Mostly SCC, early and late stage
Hauspy *et al.*, 2007[58]	NR	–	39/41 (95%)	Mostly SCC, early and late stage
Johann *et al.*, 2008[59]	–	–	23/23 (100%)	Only SCC, early stage
Klat *et al.*, 2009[60]	–	–	23/23 (100%)	Only SCC, early stage
Levenback *et al.*, 2001[61]	–	46/52 (88%)	–	Mostly SCC, early and late stage
Martinez-Palones *et al.*, 2006[63]	–	–	27/28 (96%)	Mostly SCC, early stage
Merisio *et al.*, 2005[64]	20/20 (100%)	–	–	Only SCC, early stage
Moore *et al.*, 2008[65]	–	–	35/36 (97%)	Only SCC, early and late stage
Radziszewski *et al.*, 2010[68]	–	–	56/62 (90%)	Only SCC, early stage
Terada *et al.*, 2006[70]	–	–	21/21 (100%)	Only SCC, early stage
Vakselj *et al.*, 2007[71]	–	–	35/35 (100%)	Mostly SCC, early stage
Van Den eynden *et al.*, 2003[72]	–	–	27/32 (84%)	Mostly SCC, early stage
Van der Zee *et al.*, 2008[73]	–	–	403/403 100%	Only SCC, early stage
Vidal-Sicart *et al.*, 2007[74]	–	–	49/50 (98%)	Mostly SCC, early and late stage
Combined rates	94.6%	68.7%	97. 7%	–
95% CI	0.909 to 0.971	0.631 to 0.740	0.966 to 0.985	–

a ^{99m}Tc and blue dye discrepant results in text and table.

Note: for Brunner *et al.*,[52] Lindell *et al.*,[39] Louis-Sylvestre *et al.*,[62] Nyberg *et al.*,[66] Pitynski *et al.*[67] and Rob *et al.*,[69] a single test was used for a proportion of patients and a combination of tests used for the remainder so the detection rate per patient is not specific to any single test or combination.

TABLE 25 Sentinel lymph node detection rates per patient according to FIGO stage

	Nyberg *et al.* 2007[66]			Vidal-Sicart *et al.* 2007[74]		
Stage	Blue dye	^{99m}Tc	Blue dye/^{99m}Tc	Blue dye	^{99m}Tc	Blue dye/^{99m}Tc
Stage I	100% (11/11)	100% (10/10)	100% (11/11)	74% (17/23)	100% (23/23)	100% (23/23)
Stage II	93% (13/14)	90% (9/10)	100% (10/10)	90% (18/20)	100% (20/20)	100% (20/20)
Stage III	100% (21/21)	90% (18/20)	100% (20/20)	71% (5/7)	86% (6/7)	NR
Stage IV	0% (0/1)	None	None	None	None	None

NR, not reported.

TABLE 26 Training and experience of physicians performing SLN biopsy

Study	Learning curve mentioned
Achimas-Cadariu *et al.*, 2009[50]	At least 20 inguinal SLN biopsies followed by full IFL regardless of node status; thereafter only SLN performed
Hampl *et al.*, 2008[57]	At least 10 SLN biopsies followed by full IFL regardless of node status (before offering SLN biopsy)
Johann *et al.*, 2008[59]	SLN biopsy followed by full IFL regardless of node status during a learning period of approximately 2 years, before SLN biopsy and full IFL if histology positive
Levenback *et al.*, 2001[61]	Success of SLN identification varied by clinical experience of the procedure such that the failure rate was 16% (4/25) in the first 2 years and 7% (2/27) in subsequent years
Van der Zee *et al.*, 2008[73]	At least 10 successful experiences of the SLN procedure with subsequent full IFL regardless of node status (this study design was SLN biopsy and full IFL if histology positive)
Vidal-Sicart *et al.*, 2007[74]	In the first 50 patients, SLN procedure followed by a full IFL was performed, as a representing learning curve, thereafter only SLN biopsy performed

first few cases, only SLN biopsy without full IFL was to be performed. Only Levenback *et al.*[61] calculated that the rate of SLN detection was worse in the first 2 years of the study but then continues with full IFL for all patients regardless of node status.

Recurrence rates

Two groups of studies gave recurrences at follow-up. The first group are those that used full IFL regardless of node status at initial operation to establish diagnostic accuracy, but also followed up patients afterwards. The second are those that used clinical follow-up for SLN-negative patients to establish diagnostic accuracy.

In group 1 are Crosbie *et al.*,[54] Klat *et al.*,[60] Martinez-Palonez *et al.*,[63] Vakselj and Bebar[71] and Vidal-Sicart *et al.*[74] Overall recurrence rates are given in *Table 11*. Vidal-Sicart *et al.*[74] did not give recurrences by SLN status. In Crosbie *et al.*,[54] 6 out of 31 patients were SLN positive at biopsy and five SLN-positive patients developed recurrences. No SLN-negative patients developed recurrences. In Klat *et al.*,[60] 14 out of 23 patients were SLN positive at biopsy (and one with a malignancy not in SLN) and three SLN-positive patients developed recurrences. One SLN-negative patient developed recurrence. In Martinez-Palonez *et al.*,[63] 6 out of 27 patients were SLN positive at biopsy (and one with malignancy not in SLN) and four SLN-positive patients developed recurrences. Recurrences also developed in possibly three SLN-negative patients and in one patient for whom the SLN was not found (the numbers are unclear from the journal article). In Vakselj and Bebar,[71] 10 out of 35 patients were SLN-positive at biopsy and six SLN-positive patients developed recurrences. Three SLN-negative patients also developed recurrences.

In group 2 are Achimas-Cadariu *et al.*,[50] Moore *et al.*,[65] Terada *et al.*,[70] Van den Eynden *et al.*[72] and Van der Zee *et al.*[73] The overall recurrence rates are given in *Tables 12* and *22*. Achimas-Cadariu *et al.*[50] did not give recurrences by SLN status, so test accuracy could not be calculated. Van der Zee *et al.*[73] gave separate results for groin and local recurrences. In addition, Van der Zee *et al.*[73] gave recurrences for node-negative and node-positive patients in separate papers (node negative, Van der Zee *et al.*;[73] node positive, Oonk *et al.*[83]). It was curious that there were 34 local recurrences in each category.

Given the recurrence results, it is reasonable to assume that the number of clinically apparent recurrences is likely to be smaller than the number of SLN-positive patients at biopsy. This may be because the subsequent IFL is removing malignancy that might otherwise develop into a recurrence. In addition, in a very small proportion of patients who do not undergo IFL, some recurrences will be distant rather than groin because the groin metastases may stay very small and not be noticeable.

In addition, it shows that some SLN-negative patients will develop recurrences so will be FNs. In Martinez-Palonez *et al.*[63] and Vakselj and Bebar,[71] there was a higher rate of recurrences in SLN-negative patients than known FN SLN biopsies. In Crosbie *et al.*[54] and Klat *et al.*,[60] the rates were the same.

Survival rates

Nine studies gave information about survival.[50,54,60,63,65,70,71,73,74]

In the study by Achimas-Cadariu *et al.*,[50] 12 out of 46 patients died during follow-up; median survival was 61.2 months for the whole cohort and 16.2 months for the eight patients with relapse.

Crosbie *et al.*[54] reported that 2 out of 32 patients died from disease during a median follow-up period of 62 months (range 33–84 months).

Klat *et al.*[60] reported that 1 out of 23 patients died from disease, with a follow-up of 8–46 months.

Martinez-Palonez *et al.*[63] did not mention any deaths in a group of 28 patients followed up for a median of 22.5 months (range 0–64 months).

Moore *et al.*[65] reported that 1 out of 35 patients died from intercurrent disease, with a median follow-up of 29 months (range 8–51 months).

Terada *et al.*[70] gave overall survival and disease-free survival curves for all patients and node-negative patients. Median follow-up of 21 patients was 4.6 years (range 2–8 years) and largely reflects losses to follow-up, as none of the node-negative patients died of cancer. Two out of three node-positive patients died of cancer and three patients died of other illnesses.

Vakselj and Bebar[71] gave the status of each of the node-positive and -negative patients and length of follow-up. Out of the 10 node-positive patients, six died of disease at follow-up times of 42, 10, unstated, 3, 18 and 10 months (total range of follow-up was 3–55 months). Of the 25 node-negative patients, one died of disease at 49 months and one died of another cause (no time given) (total range of follow-up 2–52 months).

Van der Zee *et al.*[73] gave a disease-specific survival curve for node-negative patients with a median follow-up of 35 months (range 2–87 months) for 202 of the 276 patients with at least 24 months of follow-up. Four patients were lost to follow-up, 10 patients died of vulval cancer and 16 died of intercurrent disease. Five out of 34 patients with local recurrents eventually died of distant metastases at 15, 18, 22, 41 and 41 months after primary treatment. The 3-year disease-specific survival rate for patients with unifocal vulval disease and negative SLNs was 97.0%. In the Oonk *et al.* paper,[83] in the group of patients with positive SLNs, the 5-year disease-specific survival was 77.3%. Survival was varied depending on the pathology, and was 64.9% when the SLNs were identified by routine pathology and 92.1% when identified by ultrastaging. During the follow-up of median 31 months (range 0–109 months) 15 died of other causes and 28 died of vulval cancer.

In the study by Vidal-Sicart *et al.*,[74] 1 out of 50 patients died of disease, with a mean follow-up of 20 months.

In summary, vulval cancer is largely a disease of older women, so a relatively large proportion die of other causes during follow-up. Survival rates are better in women with a node-negative status than in those with a node-positive status.

Quality of life

The Van der Zee *et al.*[73] study had a substudy investigating QoL that was published separately (Oonk *et al.*[82]). The authors used the European Organisation for Research and Treatment of Cancer QoL Questionnaire-C30 (EORTC QLQ-C30), the Functional Assessment of Cancer Therapy – Vulvar (FACT-V) and a patient's opinion questionnaire. The EORTC QLQ-C30 consists of five functional scales, global health status and nine symptom scales. The FACT-V consists of five functional scales and four symptom scales. In both of these questionnaires, a higher score indicates better functioning on the functional scales and worse functioning on the symptom scales. The patient opinion questionnaire asked what they would recommend to a friend or family member with vulval cancer: IFL or SLN biopsy (with the chance of missing metastases at a rate of 1 in 10, 1 in 100 or 1 in 1000). Patients included in the Van der Zee *et al.*[73] study between March 2000 and December 2005 were eligible. Questionnaires were sent to 37 patients with positive SLN and 37 age-matched patients with a negative SLN. Sixty-two of the 74 patients (84%) returned the completed questionnaires.

The results of the EORTC QLQ-C30 are in given *Table 27*, the results of the FACT-V in *Table 28* and the results of the patient's opinion questionnaire in *Table 29*. The most useful result was the global health status/QoL, which had a mean of 80 and a standard deviation (SD) of 18 for the SLN group and a mean of

TABLE 27 Results of the EORTC QLQ-C30 questionnaire

Parameter	SLN patients ($n = 35$), mean (SD)	IFL patients ($n = 27$), mean (SD)	p-Value
Functional scale[a]			
Physical functioning	84 (21)	80 (19)	0.43
Role functioning	87 (22)	85 (26)	0.87
Emotional functioning	90 (14)	89 (19)	0.63
Cognitive functioning	94 (11)	94 (14)	0.90
Concentration	95 (12)	96 (14)	0.44
Memory	92 (16)	91 (18)	0.83
Social functioning	96 (13)	90 (22)	0.23
Symptom scale[b]			
Fatigue	23 (22)	18 (24)	0.23
Nausea and vomiting	3 (11)	2 (10)	0.61
Nausea	6 (19)	4 (14)	0.61
Vomiting	1 (6)	1 (6)	0.85
Pain	15 (21)	14 (24)	0.63
Dyspnoea	12 (24)	12 (21)	0.79
Insomnia	14 (25)	23 (29)	0.15
Appetite loss	10 (24)	9 (22)	0.65
Constipation	7 (16)	5 (15)	0.53
Diarrhoea	9 (20)	2 (13)	0.11
Financial difficulties	2 (11)	12 (25)	0.01[c]
Global health status/QoL	**80 (18)**	**80 (23)**	**0.62**

a A high score on a functional scale represents a high/healthy level of functioning.
b A high score on a symptom scale represents a high level of symptomatology problems.
c $p < 0.05$.

TABLE 28 Results of FACT-V

Parameter	SLN, mean (SD)	IFL, mean (SD)	*p*-Value
Functional scale[a]			
Physical functioning	85 (17)	80 (18)	0.27
Discomfort groins/vulva/legs	86 (20)	69 (32)	0.03
Discomfort sitting	87 (18)	86 (28)	0.46
Discomfort bending	83 (27)	85 (20)	0.83
Sexual functioning[b]	78 (19)	81 (26)	0.67
Future perspective	70 (27)	64 (28)	0.41
Body image	43 (35)	59 (33)	0.09
Contentment	90 (27)	78 (31)	0.04[d]
Sexual activeness	6 (21)	13 (21)	0.06
Symptom scale[c]			
Vulval symptoms	14 (14)	10 (12)	0.33
Discharge/blood loss	5 (12)	1 (6)	0.17
Fetor	7 (16)	5 (12)	0.77
Itching	24 (26)	16 (19)	0.29
Pain/numbness	22 (28)	20 (27)	0.76
Oedema	12 (22)	35 (32)	0.001[d]
Complaints	12 (24)	27 (29)	0.01[d]
Stockings	12 (30)	43 (46)	0.003[d]
Urination	14 (18)	18 (18)	0.30
Incontinence	18 (30)	26 (34)	0.32
Discomfort	10 (21)	10 (18)	0.68

a A higher score indicates a higher/better level of functioning/contentment.
b Questions on sexual functioning were only answered by woman who were sexually active.
c A higher score indicates a high level of symptomatology/problems.
d $p < 0.05$.

TABLE 29 Maximum FN rate of the SLN procedure acceptable to patients

Study	Maximum acceptable FN rate (%)	Patients who accept the FN rate (%)	SLN	IFL	*p*-Value[a]
Oonk *et al.*, 2009[82]	10	69	84%	48%	0.005
	1	82	97%	62%	0.001
	0.1	87	97%	71%	0.013

a Chi-squared test.

80 and a SD of 23 for the IFL group. The authors state that 'our present study does not support our original idea that a decrease in especially long-term morbidity also translates into an improved overall quality of life'.[82] It may be that there were too few participants to detect a small difference in QoL.

The EORTC QLQ-C30 showed very few differences between the two groups; only on the financial difficulties scale was the score statistically significantly worse in the IFL group. For the FACT-V questionnaire, results were significantly worse for the contentment functional scale, oedema, complaints and stockings symptom scales.

For the patient's opinion questionnaire, the authors analysed the maximum FN rate of the SLN procedure that would be acceptable to patients. There were significant differences between the SLN biopsy patients and the IFL patients for all three FN rates, so that the SLN biopsy patients were more likely than the IFL patients to accept the FN rates.

Adverse events

Information about AEs was generally poorly reported. Eight studies provided data.[50,52,54,59,65,70,72,73] In Brunner *et al.*,[52] 8.7% of patients' groins had postoperative inguinal morbidity (inguinal seromas, abscess, wound breakdown), but this information was not given by group.

Adverse events according to surgical procedures (SLN biopsy, SLN biopsy plus IFL) and time interval (short and long term) are as follows:

Short-term AEs:

- For SLN biopsy only: transient lymph oedema (13%),[59] wound breakdown (11.7%) and wound cellulitis (4.5%).[73]
- For SLN biopsy + IFL: transient lymph oedema (39%),[59] postoperative groin lymphocele (5.5%) and cellulitis arising in the labia majora (2.8%),[65] wound cellulitis (9.5%) and seroma (4.3%),[70] wound breakdown (34%) and wound cellulitis (21.3%),[73] cellulitis (5.9%) and lymphocele (11.8%).[72]

Longer term AEs:

- For SLN biopsy only: lymphoedema (1.9%) and recurrent erysipelas (0.4%).[73]
- For SLN biopsy + IFL: wound infection (31%), wound dehiscence (5%), lymphocyst (22%) and chronic lymphoedema (16%),[54] lymphoedema (25.2%) and recurrent erysipelas (16.2%).[73]

In Achimas-Cadariu *et al.*,[50] AEs were presented according to surgical procedures and by vulva and groin locations:

- For SLN biopsy only: wound breakdown – 3% (vulva) and 3.6% (groin); haematoma – 3.6% (vulva) (no AEs from blue dye only).
- For SLN biopsy + IFL: wound breakdown – 5% (vulva) and 6% (groin); haematoma – 1% (vulva) and 3.7% (groin); chronic lymphoedema – 3.7% (vulva) and 8% (groin).
- For IFL only: wound breakdown – 2% (vulva) and 7% (groin); haematoma – 6.9% (vulva) and 6.9% (groin); chronic lymphoedema – 1.7% (vulva) and 4% (groin).

Chapter 6 Clinical effectiveness review

Study selection

The final search retrieved 14,038 potentially relevant citations, which were screened for relevance to the inclusion criteria. Relevant full-text articles for 313 citations were retrieved and 295 articles were subsequently excluded. The most common reasons for study exclusion were lack of the full-text version, wrong study design and wrong population. The list of excluded studies and reasons for their exclusion can be found in *Appendix 11*. Eighteen publications (corresponding to 17 studies) fulfilled the inclusion criteria (*Figure 5*). There was one RCT, three case–control studies and 13 case series included.

Randomised controlled trials

Characteristics of included study

Only one RCT (Stehman *et al.*[87]) met the inclusion criteria for the clinical effectiveness systematic review. It included patients with primary vulval cancer at FIGO stage I, II and III and TNM classification T1–3, N0–1, M0 (percentages in each FIGO stage not given, but likely that I and II were > 75%). Patients were enrolled between 1986 and 1990. The intervention group received radical vulvectomy or modified radical hemivulvectomy with RT to both groins. The control group received radical vulvectomy or modified radical hemivulvectomy with bilateral groin IFL. Patients were followed up for at least 3 years, or until recurrence. Inclusion criteria and baseline characteristics are in *Table 30*. Characteristics were well balanced between the groups. Three patients in the RT arm did not receive the full doses of RT; however, all patients in the control arm received IFL. All patients were analysed for recurrence, survival, deaths, AEs and number of days spent in hospital after clinical intervention.

Quality of included study

In Stehman *et al.* 1992,[87] the authors did not describe method of randomisation, apart from that it was conducted in blocks that were balanced within and between institutions (21 institutions). In addition, methods of allocation concealment or any blinding procedure that might have been used were not reported. However, it would not have been possible to blind the interventions and blinding the outcomes assessment may have made little difference. Results for all eligible patients are given.

Randomised controlled trial results

The median postoperative stay was 13 days [interquartile range (IQR) 8–20 days] and those in the RT arm had substantially shorter hospitalisations than those in the IFL arm ($p = 0.0001$). Ten patients had hospital stays of 7 days or fewer in the radiation arm compared with two in the IFL arm, whereas six patients had hospital stays of 13 or more days in the RT arm compared with 20 in the IFL arm. There were 10 episodes of grade 3 toxicity in the RT arm, including seven episodes of cutaneous toxicity. There were 22 episodes of grade 3 or 4 toxicity in the IFL arm including one death, five severe cardiovascular complications and 12 wound eruptions. There were also seven patients in this group with mild or moderate lymphoedema.

Survival curves are presented for progression-free survival, overall survival and relative survival (described as survival with intercurrent deaths censored) for up to 36 months. All three show that survival in the RT arm was worse than the IFL arm (log-rank tests: $p = 0.033$, $p = 0.035$ and $p = 0.042$, respectively). In the radiation arm there were eight deaths from vulval cancer and two from other causes. In the IFL arm there was one treatment-related death (pulmonary embolism), one death from vulval cancer and one death from other causes. The calculated relative risk for overall survival is 1.40 (95% CI 1.01 to 1.93) in favour of the IFL arm.

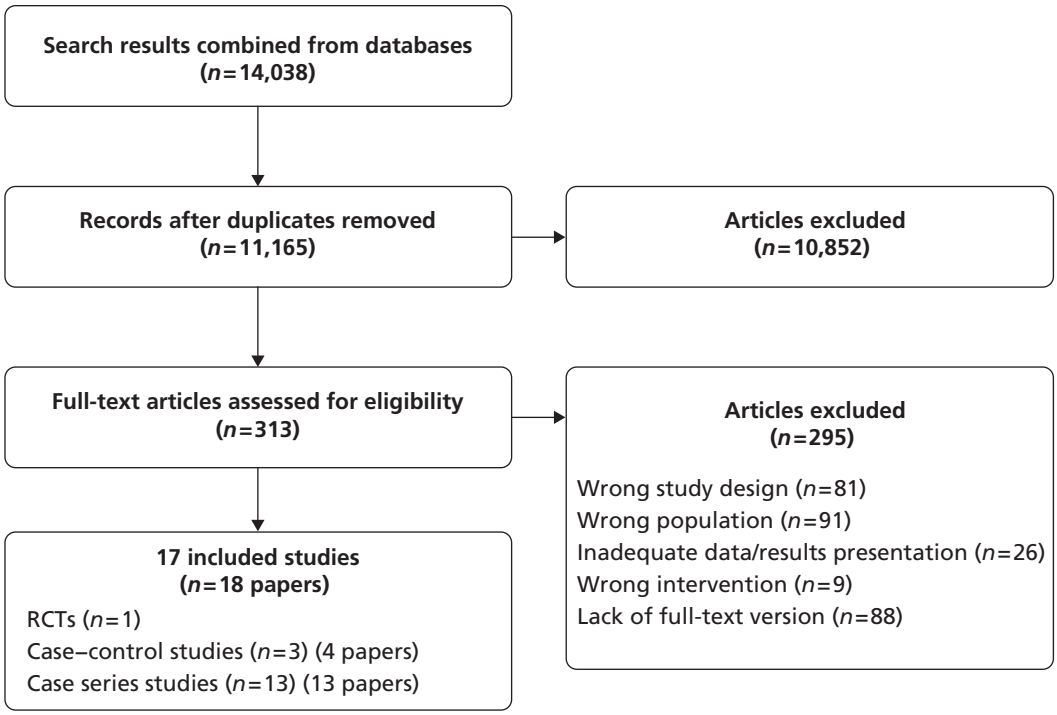


FIGURE 5 A PRISMA diagram of the selection process for the clinical effectiveness systematic review.

In the radiation arm, there was one vulval recurrence, five groin recurrences and two distant recurrences. In the IFL arm, there was one vulval recurrence and one distant recurrence. The calculated relative risk for recurrence of any sort is 0.27 (95% CI 0.07 to 0.99) in favour of the IFL arm.

Case–control studies and case series

Sixteen studies were included in this section (see *Table 31*). There was one matched case–control study (Helm *et al.*[37]), two unmatched case–control studies (Manavi *et al.*[88] and Stehman *et al.*[89]) and the remainder were retrospective case series. One study (Katz *et al.*[90]) included more than 25% FIGO stage III and IV patients but the results for recurrence for stage I and II patients only by treatment received were clear, so this subgroup was reported here. The characteristics of the included studies are shown in *Table 31*. Almost all studies were conducted in the USA or Europe and the follow-up times varied between 1 and 30 years.

The numbers of patients given different types of treatments are shown in *Table 32*. The table indicates that the different studies evaluated several interventions and many of the groups were small. Many of the case series compared radical vulvectomy with hemivulvectomy and wide excision and it was frequently difficult to determine how many patients had received unilateral or bilateral IFL. The largest study was that by Kumar *et al.*,[99] which reported results from the US Surveillance, Epidemiology and End Results (SEER) database between 1988 and 2005. With this paper, it was impossible to determine the number of women who had had IFL so the results have been placed in the surgery to primary site-only category.

The case–control study by Helm *et al.*[37] matched 32 women who had a triple incision for IFL with 32 women of similar FIGO stage, lymph node status, greatest diameter of lesion and site of lesion with a single incision for IFL. The case–control study by Manavi *et al.* 1997[88] compared women who had had inguinofemoral RT with those who had no RT. IFL was not performed on any of the patients but all had had simple vulvectomy. The case–control study by Stehman *et al.*[89] compared women who had modified radical hemivulvectomy and ipsilateral superficial inguinal lymphadenectomy with a historical comparator group of women who had had radical vulvectomy and bilateral IFL. Cases and controls had comparable

TABLE 30 Inclusion criteria and baseline characteristics of included RCT (Stehman *et al.*[87])

Parameter			Intervention group	Control group
Patients	Inclusion criteria		Primary SCC of the vulva, previously untreated, T1–3 FIGO (according to guidelines from 1971), patients with T1 were eligible only if there was capillary–lymphatic space involvement or if there was > 5 mm of invasion. Lymph nodes, if palpable, must not have been suspicious and informed consent was gained for study in accordance with local Institutional review board guidelines and with the Treaty of Helsinki	
	Exclusion criteria		Patients with distant metastasis (M1), unsuitable for operation, have received any prior radiation or chemotherapy, with any prior malignancy other than non-melanoma skin cancer of a site other than the vulva	
	Number of patients		25	27
	Age (years)	Median	64	
		Age ranges: *n* (%)	31–40: 1 (4) 41–50: 1 (4) 51–60: 5 (20) 61–70: 7 (28) 71–80: 5 (20) 81–90: 6 (24)	31–40: 3 (11.1) 41–50: 6 (22.2) 51–60: 4 (14.8) 61–70: 5 (18.5) 71–80: 6 (22.2) 81–90: 3 (11.1)
Disease	Location of primary tumour: *n* (%)		Labia: 14 (56) Clitoris: 6 (24) Perineum: 3 (12) Other: 2 (8)	Labia: 16 (59.3) Clitoris: 5 (18.5) Perineum: 6 (22.2)
	Size of tumour: *n* (%)		≤ 2 cm: 2 (8) 2.1–4 cm: 18 (72.5) ≥ 4.1 cm: 5 (20)	≤ 2 cm: 3 (11.1) 2.1–4 cm: 19 (70.4) ≥ 4.1 cm: 5 (18.5)
	Depth of invasion		Not reported	
	Tumour histology type		Squamous cell carcinoma	
	Morphology of the nodes: *n* (%)		Not palpable: 20 (80) Palpable but normal: 5 (20)	Not palpable: 20 (74.1) Palpable but normal: 7 (25.9)

stage I disease, 5 mm or less invasion, no vascular space involvement and negative inguinal and femoral nodes.

The baseline characteristics of patients in the included studies are shown in *Table 33*. Their median or mean ages were mostly of 60 years, which reflects the fact that vulval cancer is mostly a disease of older women. However, some younger women were included down to age 20 years and Kumar *et al.*[99] compared characteristics of younger women (aged < 50 years) with older women (aged ≥ 50 years). Where location of primary tumour was specified in studies, the majority were on the labium majus and were medial rather than lateral. The vast majority of malignancies were SCCs but also included were a few adenocarcinomas, MMs and others in some of the studies. Some of the studies included just FIGO stage I, whereas others had a spread of stages from I to IV. Some of the studies gave TNM classification as well as FIGO stage but others just gave one or the other. Node status was given in eight studies only.[37,89,91,92,95,96,100,101]

TABLE 31 Characteristics of included case series

Study	Study type	Population	Location	Length of follow-up [range (months)]
Anderson *et al.*, 1995[91]	Retrospective, case series	Primary vulval cancer in stages I, II and III according to TNM classification	USA	Median: 54 months (2–212)
Andrews *et al.*, 1994[92]	Retrospective, case series	Primary vulval cancer in FIGO stages I, II, III and T1, T2 according to TNM classification	USA	Mean: 5 years (1–15)
Burke *et al.*, 1995[93]	Retrospective, case series	Primary vulval cancer in stages T1, T2 according to TNM classification	USA	Median: 38 months
Busch *et al.*, 2000[94]	Retrospective, case series	Primary vulval cancer in FIGO stages I, II and T1N0–1, T2N0–1 according to TNM classification	Germany	Up to 30 years
de Hullu *et al.*, 2002[95]	Retrospective case series	Primary vulval cancer in stages T1, T2 according to TNM classification	Netherlands	Median:110 months (3–220)
DeSimone *et al.*, 2007[96]	Retrospective, case series	Primary vulval cancer in FIGO stages I, II, III and T1, T2 according to TNM classification	USA	Mean: 59 months (10–195)
Farias-Eisner *et al.*, 1994[97]	Retrospective, case series	Primary vulval cancer in FIGO stages I, II and T1N0–1M0, T2N0–1M0 according to TNM classification	USA	Median: 12 months (6–77)
Hallak *et al.*, 2007[98]	Retrospective, case series	Primary vulval cancer in FIGO stages I, II	Sweden	Mean: 101 months (19–252)
Helm *et al.*, 1992[37]	Matched case–control study	Primary vulval cancer in FIGO stages I, II, cases with triple incision IFL, controls with single incision	USA	Up to 8 years
Katz *et al.*, 2003[90]	Retrospective, case series	Primary vulval cancer in FIGO stages I, II, III and IV (stage I and II subgroup only)	USA	Up to 10 years
Kumar *et al.*, 2009[99]	Retrospective, case series	Primary vulval cancer in FIGO stages I, II	USA	Up to 220 months
Manavi *et al.*, 1997[88]	Case–control study	Primary vulval cancer in stages T1, T2 according to TNM classification, cases with inguinofemoral RT, controls without	Austria	Up to 5 years
Scheistroen *et al.*, 2002[100]	Retrospective, case series	Primary vulval cancer in FIGO stages I, II	Norway	Up to 80 months
Stehman *et al.*, 1992[89]	Case–control study	Primary vulval cancer in FIGO stage I, cases had modified radical hemivulvectomy and ipsilateral inguinal lymphadenectomy, historical controls with vulvectomy and bilateral IFL	USA	3 years or until death
Tantipalakorn *et al.*, 2009[101]	Retrospective, case series	Primary vulval cancer in FIGO stages I, II	Australia	Median: 84 months
Vavra *et al.*, 1990[102]	Retrospective, case series	Primary vulval cancer in FIGO stages I, II and T1, T2–3 according to TNM classification	Austria	Up to 5 years

TABLE 32 Case series: numbers of patients receiving each treatment

Study	Total number of patients	Surgery to primary site only	IFL – bilateral and surgery	IFL – unilateral and surgery	RT to groin and surgery to primary site	RT only	Surgery, IFL and RT	Comments
Anderson *et al.*, 1995[91]	47	13	–	23	–	11	–	–
Andrews *et al.*, 1994[92]	84	–	49	28	–	–	7	–
Burke *et al.*, 1995[93]	76	–	51	25	–	–	–	–
Busch *et al.*, 2000[94]	92	–	–	–	92	–	–	–
de Hullu *et al.*, 2002[95]	253	–	219	34	–	–	–	–
DeSimone *et al.*, 2007[96]	122	–	60	62	–	–	–	–
Farias-Eisner *et al.*, 1994[97]	74	3	58	13	–	–	–	–
Hallak *et al.*, 2007[98]	294	–	–	–	267	–	27	–
Helm *et al.*, 1992[37]	64	–	62	–	–	–	2	Matched comparison
Katz *et al.*, 2003[90]	153	–	Combined 104		153	–	29	Stage I and II subgroup
Kumar *et al.*, 2009[99]	6965	5253	–	–	1023	689	–	Number of IFL not given
Manavi *et al.*, 1997[88]	135	70	–	–	65	–	–	–
Scheistroen *et al.*, 2002[100]	216	47	161	8	–	–	–	–
Stehman *et al.*, 1992[89]	217	–	96	121	–	–	–	–
Tantipalakorn *et al.*, 2009[101]	121	19	48	54	–	–	–	–
Vavra *et al.*, 1990[102]	101	43	–	–	58	–	–	–

TABLE 33 Case series: baseline characteristic of patients

	Patients			Diseases									
								FIGO stage					
Study	*n*	Number of analysed patients	Age (years)	Location of primary tumour	Size of tumour (cm) [SD]	Depth of invasion (mm) [SD]	Histology	I	II	III	IV	TNM	Nodes status
Anderson *et al.*, 1995[91]	47	40	Median: 71 (range 20–91)	NR	NR	NR	SCC	15 (32%)	13 (28%)	12 (25%)	–	T1	Positive: 7 (20%)
Andrews *et al.*, 1994[92]	84	84	Median: 65 (range 28–89)	Labium majus: *n* = 46 (54.8%); labium minus: *n* = 26 (31%); both: *n* = 12 (14.2%)	T1: Median: 1.5 (range 0.5–2); T2: Median: 3.4 (range 2.2–9.0)	1–2 mm: 17; 2,1–3 mm: 23; > 3 mm: 32; unknown: 12	SCC	32 (38%)	31 (37%)	21 (25%)	–	T1: *n* = 37 (44%); T2: *n* = 47 (56%)	T1: 5 (13%) positive; T2: 16 (34%) positive
Burke *et al.*, 1995[93]	76	76	Mean: 62 (range 22–87)	NR	Mean: 2.6 (range 0.5–6.5)	Mean: 4.4	SCC alone	NR	NR	NR	NR	T1: *n* = 33 (43.3%); T2: *n* = 43 (56.6%)	NR
Busch *et al.*, 2000[94]	92	56	Mean: 65 (range 23–81)	Labia major: *n* = 10 (10.9%); labia minor: *n* = 38 (41.3%); clitoris: *n* = 20 (21.7%); mons pubis and perineum: *n* = 24 (26.1%)	≤ 2 cm: *n* = 42 (45.7%); > 2 cm: *n* = 50 (44.3%)	NR	NR	42 (45.7%)	50 (44.3%)	–	–	T1N0: *n* = 42 (45.7%), T2N0: *n* = 50 (44.3%)	NR

Study	Patients			Diseases				FIGO stage					
	n	Number of analysed patients	Age (years)	Location of primary tumour	Size of tumour (cm) [SD]	Depth of invasion (mm) [SD]	Histology	I	II	III	IV	TNM	Nodes status
de Hullu *et al.*, 2002[95]	253	238	Group I: Median: 71 (range 28–94) Group II: Median: 71 (range 34–94)	NR	NR	Group I: Median: 5 (range 1–4) Group II: Median: 5 (range 1–27)	SCC alone	NR	NR	NR	NR	Group I: T1 *n* = 53 (21%); T2 *n* = 115 (45.5%) Group II: T1 *n* = 25 (10%); T2 *n* = 60 (23.5%)	Invasion Group I: 29 (11.5%) Group II: 13 (5.1%)
DeSimone *et al.*, 2007[96]	122	122	Mean: 67 (range 28–91)	Labium minus: *n* = 38 (31.1%); labium majus: *n* = 68 (55.7%); both: *n* = 16 (13.2%)	T1: Median: 1.5 (range 0.5–2.0) T2: Median: 3.6 (range 2.2–9.0)	T1: Median: 4 (range 1.2–6) T2: Median: 5.4 (range 1.1–15)	SCC alone	54 (44%)	42 (35%)	26 (2%)	–	T1: *n* = 61 (50%) T2: *n* = 61 (50%)	Lymph vascular involvement –23 (19%); and perineural invasion –11 (9)
Farias-Eisner *et al.*, 1994[97]	74	74	NR	NR	NR	NR	SCC alone	39 (52.7%)	35 (47.3%)	–	–	NR	NR

continued

TABLE 33 Case series: baseline characteristic of patients (*continued*)

	Patients			Diseases									
								FIGO stage					
Study	*n*	Number of analysed patients	Age (years)	Location of primary tumour	Size of tumour (cm) [SD]	Depth of invasion (mm) [SD]	Histology	I	II	III	IV	TNM	Nodes status
Hallak *et al.*, 2007[98]	294	225	Mean: 71 (range 32–93)	NR	NR	NR	SCC: *n* = 269 (91.5%) MM: *n* = 10 (3.4%) Paget's disease: *n* = 8 (2.7%) Other: n = 7 (2.4%)	110 (37.4%)	115 (39.1%)	43 (14.6%)	26 (8.9%)	NR	NR
Helm *et al.*, 1992[37]	64	35	Mean: 65	Single incision group: vulva, *n* = 25 (39%); midline, *n* = 7 (10.9%) Triple incision group: vulva, *n* = 23 (35.9%); midline, *n* = 9 (14.2%)	Single incision group: ≤ 1: *n* = 6 (9.4%); > 1–2: *n* = 11 (17.2%); > 2–4: *n* = 10 (15.6%); > 4: *n* = 5 (7.8%) Triple incision group: ≤ 1: *n* = 5 (7.8%); > 1–2: *n* = 13 (20.3%); > 2–4: *n* = 8 (12.5%); > 4: 6 (*n* = 9.4%)	NR	SCC alone	32 (50%)	20 (31.3%)	8 (12.5%)	4 (6.2%)	NR	Negative – 56 (87.4%) Positive (bilateral) – 4 (6.3%) (unilateral) – 4 (6.3%)
Katz *et al.*, 2003[90]	153	153	NR	NR	NR	NR	NR	153 (100%)	–	–	NR	NR	

Study	Patients			Diseases				FIGO stage					
	n	Number of analysed patients	Age (years)	Location of primary tumour	Size of tumour (cm) [SD]	Depth of invasion (mm) [SD]	Histology	I	II	III	IV	TNM	Nodes status
Kumar *et al.*, 2009[99]	5620	3239	Mean: 73 (range 50–102)	NR	NR	NR	SCC alone	3239 (57.6%)	1862 (33.1%)	248 (0.4%)	–	NR	NR
Manavi *et al.*, 1997[88]	135	135	Mean: 68.5	Clitoris: group I, $n = 6$ (9.2%); group II, $n = 1$ (1.4%)	≤ 2: $n = 135$ (100%)	Group I: ≤ 2: $n = 17$ (SD 26.15); > 2: $n = 24$ (SD 36.92) Group II: ≤ 2: $n = 30$ (SD 42.8); > 2: $n = 29$ (SD 41.42)	SCC: $n = 130$ (96.3%); adenocarcinoma: $n = 1$ (1%); other: $n = 4$ (3%)	NR	NR	NR	NR	T1, N0–1: $n = 135$ (100%)	NR
Scheistroen *et al.*, 2002[100]	216	216	Mean: 70 (range 27–96)	Medial: $n = 132$ (61%); lateral: $n = 80$ (37%); multifocal: $n = 4$ (2%)	≤ 2: $n = 91$ (42%); > 2: $n = 125$ (58%)	NR	SCC	95 (44%)	121 (56%)	–	–	NR	Absent: 138 (64%); present: 31 (14%); not assessed: 47 (22%)

continued

TABLE 33 Case series: baseline characteristic of patients (*continued*)

	Patients			Diseases									
								FIGO stage					
Study	*n*	Number of analysed patients	Age (years)	Location of primary tumour	Size of tumour (cm) [SD]	Depth of invasion (mm) [SD]	Histology	I	II	III	IV	TNM	Nodes status
Stehman *et al.*, 1992[89]	217	217	Median: 61 Group 1: (range 21–40) *n* = 21 (17.3%), (range 41–60) *n* = 39 (32.3%), (range 61–80) *n* = 53 (43.8%), (range ≤ 81) n = 8 (6.6%) Group 2: (range 21–40) *n* = 4 (4.1%), (range 41–60) *n* = 29 (29.6%), (range 61–80) *n* = 57 (58.2%), (range ≤ 81) *n* = 8 (8.2%)	Group 1 labia minus: *n* = 31 (25.6%); labia majus: *n* = 58 (47.9%); labia majus and minus: *n* = 18 (14.9%); clitoris and/or labium: *n* = 2 (1.7%); perineum and/or labium: *n* = 12 (9.9%) Group 2 labia minus: *n* = 37 (37.8%); labia majus: *n* = 35 (35.7%); labia majus and minus: *n* = 12 (12.2%); clitoris and/or labium: *n* = 14 (14.3%); perineum and/or labium: *n* = 0 (0%)	Group 1: 0.1–1: *n* = 45 (37.2%); 1.1–2: *n* = 76 (62.8%) Group 2: 0.1–1: *n* = 36 (36.7%); 1.1–2: *n* = 62 (63.3%)	Group 1: 0–1 mm: *n* = 13 (10.8%); 1.1–2 mm: *n* = 43 (35.8%); 2.1–3 mm: *n* = 29 (24.2%); 3.1–4 mm: *n* = 25 (20.8%); 4.1–5 mm: *n* = 10 (8.3%); (1 missing) Group 2: 0–1 mm: *n* = 18 (18.4%); 1.1–2 mm: *n* = 22 (22.4%); 2.1–3 mm: *n* = 28 (28.6%); 3.1–4 mm: *n* = 13 (13.3%)	SCC and baso-SCC (number of patients with each not reported)	217 (100%)	–	–	–	NR	All not palpable or palpable but normal

Study	Patients			Diseases				FIGO stage					
	n	Number of analysed patients	Age (years)	Location of primary tumour	Size of tumour (cm) [SD]	Depth of invasion (mm) [SD]	Histology	I	II	III	IV	TNM	Nodes status
Tantipalakorn *et al.*, 2009[101]	121	78	Stage I: Mean: 62.9 (range 35–98) [15.4], Stage II: Mean: 67.9, (range 29–94) [16.2]	NR	Stage I: 1.6 (range 0.5–2.0) [0.4] Stage II: 3.7 (range 2.5–6.5) [1.1]	< 1: *n* = 21 (17.4%); 1.1–2.0: *n* = 26 (21.5%); 2.1–3.0: *n* = 17 (14%); 3.1–4.0: *n* = 8 (6.6%); 4.1–5.0: *n* = 1 (1%); > 5: *n* = 5 (4%)	SCC alone	78 (65%)	43 (35%)	–	–	NR	Positive: 7 (6%); negative: 114 (94.2%)
Vavra *et al.*, 1990[102]	101	81	Mean: 64.9 (range 38–75)	NR	NR	< 5: *n* = 67 (66%); > 5–10: *n* = 34 (34%)	SCC alone	54 (53%)	27 (27%)	18 (18%)	2 (2%)	T1: *n* = 58 (57%); T2 + T3: *n* = 43 (44%), N0,N1: *n* = 82 (81%); N2,N3: *n* = 19 (19%)	NR

NR, not reported.

Quality of included studies

Quality assessment of case series is shown in *Table 34* and of case–control studies in *Table 35*. Because of these study designs, it is inevitable that there will be inherent biases. In addition, they were all retrospective studies conducted by chart review and there was no reported blinding of investigators. Most of the studies were from single institutions, so were small, and, in order to obtain a reasonable sample size, they covered a number of years of recruitment. During this time, treatment methods and success rates may well have changed. Some studies attempted to reflect this by comparing different types of treatment, in particular that by Stehman *et al.*[89] However, they often compared different surgical techniques rather than whether or not patients had unilateral or bilateral IFL, or RT compared with no RT. In general, the clarity of reporting of findings in both types of study designs was poor and it was difficult to extract precise data about how the patients were selected and treated. In a number of the studies, it was also difficult to establish which patients

TABLE 34 Quality assessment of case series

	Quality factors												
Study	1	2	3	4	5	6	7	8	9	10	11	12	13
Anderson *et al.*,1995[91]	Y	Y	U	Y	Y	Y	Y	Y	N	Y	Y	Y	N
Andrews *et al.*, 1994[92]	Y	Y	Y	Y	Y	N	Y	Y	N	Y	Y	Y	N
Burke *et al.*, 1995[93]	Y	Y	Y	U	Y	U	U	Y	N	Y	Y	Y	Y
Busch *et al.*, 2000[94]	Y	Y	U	Y	Y	N	U	Y	N	Y	Y	Y	N
de Hullu *et al.*, 2002[95]	Y	Y	Y	Y	Y	Y	U	Y	Y	Y	Y	Y	N
DeSimone *et al.*, 2007[96]	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y
Farias-Eisner *et al.*, 1994[97]	Y	U	U	Y	Y	Y	Y	Y	N	Y	Y	Y	Y
Hallak *et al.*, 2007[98]	Y	Y	U	U	Y	Y	Y	Y	Y	Y	N	Y	Y
Katz *et al.*, 2003[90]	Y	Y	Y	N	Y	N	Y	Y	N	Y	Y	Y	Y
Kumar *et al.*, 2009[99]	Y	Y	U	N	N	N	Y	Y	N	Y	N	Y	N
Scheistroen *et al.*, 2002[100]	Y	Y	Y	Y	Y	N	Y	Y	N	Y	Y	Y	Y
Tantipalakorn *et al.*, 2009[101]	Y	Y	Y	Y	Y	U	Y	Y	Y	Y	Y	Y	N
Vavra *et al.*, 1990[102]	Y	Y	Y	Y	Y	N	Y	Y	N	Y	N	Y	Y

N, No; U, unclear; Y, yes.

Quality factors assessed: 1, study hypothesis; 2, population clinically described; 3, population pathologically described; 4, vulva tumour intervention described; 5, groin intervention described; 6, follow-up time reported; 7, patients at the end of follow-up; 8, clinically important outcomes reported; 9, definition of outcomes reported; 10, outcomes given; 11, data presented according to intervention; 12, adequate data presentation; 13, safety data given.

TABLE 35 Quality assessment of case–control studies

	Quality factors								
Study	1	2	3	4	5a	5b	6	7	8
Helm *et al.*, 1992[37]	Y	Y	Y	Y	Y	Y	Y	Y	Y
Manavi *et al.*, 1997[88]	N	Y	Y	Y	Y	Y	Y	N	Y
Stehman *et al.*, 1992[89]	Y	Y	N	N	N	N	Y	N	U

N, No; U, unclear; Y, yes.

Quality factors assessed: 1, case definition adequate; 2, representativeness of cases; 3, selection of controls; 4, definition of controls; 5a, comparability of cases; 5b, comparability of controls; 6, ascertainment of exposure; 7, same method of ascertainment of exposure for cases and controls; 8, non-response rate.

had which outcomes and results for the different treatment types were combined. The follow-up times were also confusingly reported in some studies so that the median follow-up was unclear and only the maximum follow-up could be obtained. Reporting of AEs was also patchy.

Effectiveness results

Deaths, survival and recurrence

All of the included studies reported either deaths, survival or recurrence. Deaths and survival are given in *Table 36* and recurrence in *Table 37*, together with the time frame that the results are reported within and any information on whether patients were node negative or positive at IFL (if conducted) and whether the deaths were following local, groin or distant recurrence (if reported). Many of the studies combined deaths or survival rates for different types of treatments. In addition, as studies were mostly small, it is difficult to establish any patterns of survival on different categories of treatment. The general trend was that there were more deaths in node-positive patients and in older patients. The two studies that compared RT with surgery on the primary site only (Manavi *et al.*[88] and Vavra *et al.*[102]) found little difference in survival between the two treatments. In Stehman *et al.*,[89] the calculated odds ratio for overall deaths was 0.66 (95% CI 0.31 to 1.40) and was not statistically significant. The paper included a survival curve for cases and controls, which showed that survival of cases was lower than of controls, but the log-rank test was not significant. There is approximately a 50 : 50 chance of dying from other causes as vulval cancer, particularly in the case series with median ages over 60 years.

Recurrences were more likely in node-positive patients. The recurrence rates show that there is an approximate ratio of 4 : 2 : 1 rate of recurrence in vulva, groin and distant. As studies were small, some had no distant recurrences. Very few gave details about node status, but results from DeSimone *et al.*[96] suggested that recurrences were more likely in node-positive patients; however, in Stehman *et al.*,[89] all recurrences in cases were in node-negative patients. In this study, survival curves for recurrence-free interval for cases and controls was given in the paper and showed significantly worse recurrence-free intervals for the cases ($p = 0.0028$).

Hospital stay

In Burke *et al.*,[93] the mean hospital stay was 10 days. In DeSimone *et al.*,[96] the mean hospital stay was 9.6 days (range 4–14 days) in patients undergoing radical vulvectomy and 5.0 days (range 2–12 days) in patients undergoing radical hemivulvectomy. In Farias-Eisner *et al.*,[97] the median hospital stay was 19 days (range 12–33 days) for patients undergoing radical vulvectomy and 11 days (range 5–17 days) for patients undergoing modified radical vulvectomy. In Stehman *et al.*,[89] the median length of hospitalisation was 7 days (IQR 5–10 days, range 3–22 days) for cases given modified hemivulvectomy and ipsilateral inguinal lymphadenectomy and 18 days for controls (ranges not given) given radical vulvectomy and bilateral IFL.

Quality of life and adverse events

No studies reported QoL and not all studies reported AEs. Of the studies that did, their reported AEs are shown in *Table 38*. In Stehman *et al.*,[89] AEs were reported only for the cases. In other studies, only some AEs were reported (see *Table 38*). The rates varied considerably between studies, for example, the rate of lymphoedema varied between 41% in Farias-Eisner *et al.*[97] and 1% in Vavra *et al.*[102] Few studies gave the total number of women who suffered an AE and, in the ones that did, the rates varied between 39%[100] and 19%.[89] Further details of AEs are reported in *Appendix 12* (*Tables 56* and *57*).

Effectiveness in younger versus older women

Kumar *et al.*[99] investigated a variety of prognostic variables in 1345 younger women (< 50 years) compared with 5620 older women (≥ 50 years). The 5-year survival rate for any treatment in younger women was 87.5% compared with 52.5% in older women. If the malignancy was localised, the survival rates were 93% versus 66%, if regional 79% versus 38% and if distant 26% versus 11%. Younger women were more likely to have been treated with surgery (92.2% vs. 84.1%) and less likely to have had RT (16.2% vs. 26.9%) than older women. In surgically treated patients, the 5-year survival rates were 90% in younger women compared with 58% in older women.

TABLE 36 Deaths from disease (or survival) and deaths from other causes (including time frame)

Study	Total number patients	Surgery on primary site only	IFL – bilateral and surgery	IFL – unilateral and surgery	RT to groin and surgery to primary site	Radiation only	Surgery, IFL and RT	Comment
Anderson *et al.*, 1995[91]	47 (40)	NR	–	NR	–	5-year overall survival 74% (range 43–100%) ($n = 11$)	–	There were 47 patients in the study and survival results given for 40 of these patients
Andrews *et al.*, 1994[92]	84	–	Combined, two died from disease within 1 year (one local, one distant) ($n = 77$)		–	–	NR	–
Burke *et al.*, 1995[93]	76	–	Combined, four died from disease, nine from other causes (median follow-up 38 months)		–	–	–	–
Busch *et al.*, 2000[94]	92	–	–	–	5-year overall survival 56%, 10-year overall survival 40%, 5-year cause-specific survival 61%, 10-year cause-specific survival 51%	–	–	
de Hullu *et al.*, 2002[95]	253	–	Combined, 37 died from disease, 34 died from other causes, (follow-up at least 4 years)		–	–	–	–
DeSimone *et al.*, 2007[96]	122	–	Combined, five died from disease of which one negative nodes (groin), four positive nodes (distant) [mean follow-up 59 months (range 10–195 months)]		–	–	–	–

Study	Total number patients	Surgery on primary site only	IFL – bilateral and surgery	IFL – unilateral and surgery	RT to groin and surgery to primary site	Radiation only	Surgery, IFL and RT	Comment
Farias-Eisner *et al.*, 1994[97]	74	Combined, 5-year overall survival 98% for stage I, 81% stage II. (Total? One death in stage I and five deaths in stage II, these five all had groin then distant recurrence)			–	–	–	Stage I node-positive 5-year overall survival 50%, node negative 100%, stage II survival by node status not given
Hallak *et al.*, 2007[98]	294	Combined, 107 died from vulval cancer, 85 from other causes (mean follow-up 101 months (range 19–252 months). 5-year overall survival 53%						
Helm *et al.*, 1992[37]	64	–	Six died from disease (all with positive nodes), one from other causes	–	–	–	One died from disease (with positive nodes) (*n*=2)	Matched comparison (follow-up median 4.7 years in one group, 2.9 years in the other group)
Katz *et al.*, 2003[90]	153	–	NR		NR	–	NR	Stage I and II subgroup
Kumar *et al.*, 2009[99]	6965	Combined, 2981 died. 5-year overall survival 87.5% in 1345 women aged under 50 years, 52.5% in 5620 older women (= 50 years)					–	Number of IFL not given
Manavi *et al.*, 1997[88]	135	5-year survival 91.4% (65 out of 70 died)	–	–	5-year survival 93.7% (61/65)	–	–	–

continued

TABLE 36 Deaths from disease (or survival) and deaths from other causes (including time frame) (*continued*)

Study	Total number patients	Surgery on primary site only	IFL – bilateral and surgery	IFL – unilateral and surgery	RT to groin and surgery to primary site	Radiation only	Surgery, IFL and RT	Comment
Scheistroen *et al.*, 2002[100]	216	Combined, 55 died from vulval cancer, 70 died from other causes. 5-year overall survival 76%, node negative 82%, node positive 57%			–	–	–	–
Stehman *et al.*, 1992[89]	217	–	–	Seven died from vulval cancer (five from groin recurrence), seven died from other causes, one unclear ($n = 121$)	–	–	–	–
Tantipalakorn *et al.*, 2009[101]	121	Combined, seven died from vulval cancer (two from vulval, three from groin and two from distant), ? two died from other causes (median follow-up 84 months)			–	–	–	–
Vavra *et al.*, 1990[102]	101	5-year survival 88.4% ($n = 43$)	–	–	5-year survival 79.3% ($n = 58$)	–	–	17 deaths in total

NR, not reported.

TABLE 37 Recurrence (local, groin, distant) and time frame

Study	Total number patients	Surgery on primary site only	IFL – bilateral and surgery	IFL – unilateral and surgery	RT to groin and surgery to primary site	Radiation only	Surgery, IFL and RT	Comment
Anderson *et al.*, 1995[91]	47 (40)	NR	–	NR	–	Three had recurrence out of the nine who had radiation therapy (33%)	–	There were 47 patients in the study and survival results given to 40 of these patients
Andrews *et al.*, 1994[92]	84	–	Combined, seven local, up to 185 months ($n = 77$)		–	–	Groin 1, distant 1 within 1 year ($n = 7$)	
Burke *et al.*, 1995[93]	76	–	Combined, nine vulval, four groin (three were negative at IFL, one was not dissected) (median follow-up 38 months)		–	–	–	
Busch *et al.*, 2000[94]	92	–	–	–	NR	–	–	
de Hullu *et al.*, 2002[95]	253	–	Combined, 38 within 2 years (of which 21 local, 11 groin, six distant) and 66 within 4 years		–	–	–	
DeSimone *et al.*, 2007[96]	122	–	Nine, of which eight local, one distant (mean follow-up 59 months (range 10–195) ($n = 60$)	Nine, of which five local, two groin, two distant (mean follow-up 59 months (range 10–195) ($n = 62$)	–	–	–	Recurrence in 12 out of 96 with negative nodes, 6 out of 26 positive nodes

continued

TABLE 37 Recurrence (local, groin, distant) and time frame (*continued*)

Study	Total number patients	Surgery on primary site only	IFL – bilateral and surgery	IFL – unilateral and surgery	RT to groin and surgery to primary site	Radiation only	Surgery, IFL and RT	Comment
Farias-Eisner *et al.*, 1994[97]	74	Recurrence stage I NR, stage II 9 recurrences: location NR [median follow-up 12 months (range 6–77 months)]			–	–	–	
Hallak *et al.*, 2007[98]	294	Combined, 127 recurrences: 80 vulval, 55 inguinal, 20 distant [mean follow-up 101 months (range 19–252 months)]						
Helm *et al.*, 1992[37]	64	–	13 recurrences (seven vulval, one distant and in five patients the site of recurrence was not reported) ($n = 62$)	–	–	–	NR ($n = 2$)	Matched comparison
Katz *et al.*, 2003[90]	153	–	17 had recurrence out of the 104 who had therapy (17%)		One had recurrence out of the 20 who had therapy (5%)	–	Five had recurrence out of the 29 who had therapy (20%)	Stage I and II subgroup
Kumar *et al.*, 2009[99]	6965	NR	–	–	NR	NR	–	Number of IFL not given

Study	Total number patients	Surgery on primary site only	IFL – bilateral and surgery	IFL – unilateral and surgery	RT to groin and surgery to primary site	Radiation only	Surgery, IFL and RT	Comment
Manavi *et al.*, 1997[88]	135	13 recurrence: five local, seven groin, one distant (up to 5 years' follow-up) ($n = 70$)	–	–	Eight recurrences: four local, three groin, one distant (up to 5 years' follow-up) ($n = 65$)	–	–	
Scheistroen *et al.*, 2002[100]	216	Combined, 66 recurrences, 41 vulva, 16 groin, nine distant (up to 80 months' follow-up)			–	–	–	
Stehman *et al.*, 1992[89]	217	–	Six vulval and one distant recurrence ($n = 96$)	10 vulval, nine groin, no distant recurrences (all node negative) ($n = 121$)	–	–	–	Follow-up 3 years or until death
Tantipalakorn *et al.*, 2009[101]	121	Combined, 28 recurrences, 26 local, two distant (median follow-up 84 months)			–	–	–	
Vavra *et al.*, 1990[102]	101	9.3% recurrence ($n = 43$)	–	–	2.6% recurrence ($n = 58$)	–	–	Five vulval, nine inguinal, two distant)

NR, not reported.

TABLE 38 Adverse events in effectiveness case series and case–control studies

Characteristic			AEs						
					Early surgical complication				
Study	Population	Intervention	Total		Wound break down/rupture (%)	Wound infection (%)	Groin breakdown (%)	Lymphocyst (%)	Cellulitis (%)
Burke *et al.*, 1995[93]	I/II	Modified vulvectomy	–		6 (8)	–	8 (11)	2	1
DeSimone *et al.*, 2007[96]	I/II	Radical vulvectomy	–		14 (23)	–	NR	5 (7)	–
		Radical hemivulvectomy	–		5 (98)	–	NR	2 (3)	–
Farias-Eisner *et al.*, 1994[97]	I/II	Radical local exision	–		2 (11)	–	5 (28)	–	–
		Radical vulvectomy	–		14 (25)	–	13 (23)	–	–
Hallak *et al.*, 2007[98]		After IFL	64 (22)		5 (19)	7 (26)	–	–	–
		No IFL			11 (6)	0 (5)	–	–	–
Helm *et al.*, 1992[37]		Single incision group	Major – 2 (6)		6 (19)	–	11 (34)	–	5(6)
		Triple incision group	Major – 1 (3)		2 (6)	–	6 (19)	–	7 (22)
Manavi *et al.*, 1997[88]	I/II	RT (for groin)	5 (8)		–	1 (2)	–	–	–
		No RT	2 (3)		–	0	–	–	–
Scheistroen *et al.*, 2002[100]	I/II	En block	85 (39)	45 (51)	21 (24)	16 (18)	–	–	–
		Triple incision group		18 (30)	8 (13)	8 (13)	–	–	–
		Individual		6 (30)	1 (5)	1 (5)	–	–	–
		Local excision		16 (34)	3 (6)	3 (6)	–	–	–
Stehman *et al.*, 1992[89]	Cases	All	23 (19)		9 (7)	14 (12)	–	–	–
Vavra *et al.*, 1990[102]		All	–	–	2 (2)	–	–	–	–

NR, not reported.

						Late surgical complication				
Bleeding (%)	None (%)	Rectovaginal fistula (%)	Major complication (%)	Inguinal pain (%)	Groin haematoma (%)	Lymphocyst (%)	Lymph oedema (%)	Lymphangitis (%)	Cellulitis (%)	Vaginal stenosis (%)
1	–	–	–	–	–	75 (80)	5 (7)	5 (4)	2 (2)	–
–	–	–	–	–	–	–	16 (26)	–	–	–
–	–	–	–	–	–	–	5 (8)	–	–	–
–	–	–	1 (1)	–	–	–	7 (38)	–	–	–
–	–	–		–	–	–	24 (41)	–	–	–
2 (7)	15 (56)	–	–	–	–	–	1 (7)	–	–	0
4 (2)	141 (77)	–	–	–	–	–	0	–	–	2 (1)
–	–	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–
–	–	1 (2)	–	1 (2)	–	–	–	–	–	2 (3)
–	–	0	–	0	–	–	–	–	–	2 (3)
–	–	–	5 (6)	–	3 (3)	–	–	–	–	–
–	–	–	1 (2)	–	1 (2)	–	–	–	–	–
–	–	–	0	–	1 (5)	–	–	–	–	–
–	–	–	3 (6)	–	2 (4)	–	–	–	–	–
–	67 (55)	–	–	–	–	–	Mild: 18 (15); moderate: 5 (4)	–	–	–
–	–	–	–	–	–	–	1 (1)	–	–	4 (4)

Chapter 7 Economic evaluation

Objective

Patients with FIGO stage I or stage II vulval cancer may have an IFL as part of the treatment in order reduce the probability of groin and distant recurrence. As has been shown previously in this report, the morbidity of this procedure is high. Moreover, it is recognised that, although there is a possibility of groin lymph node metastases before IFL is conducted, for many patients the IFL is subsequently found to have been unnecessary since no metastases were found. Therefore, an accurate method of identifying which patients need IFL would help to reduce unnecessary morbidity in many patients.

The SLN biopsy may use one or both of blue dye and ^{99m}Tc and is intended to identify one or more SLNs. The SLN is then examined using histopathological techniques to identify patients with metastases that need an IFL. The potential advantage of SLN biopsy is that there is much less morbidity because of a smaller incision if patients have no metastases in the SLN, although there is a risk that patients may test false-negative for metastasis and then go onto have a groin recurrence with an associated higher risk of mortality. The important potential advantage with SLN biopsy is that only one lymph node is removed, thereby reducing the incision, the extent of surgical dissection required and the potential for subsequent lymphocyst and lymphoedema formation as a result of fewer lymph nodes being removed.

The objective of this economic evaluation is to compare the relative cost-effectiveness of undertaking a range of SLN biopsy options that examine for the presence of metastases compared with implementing a strategy of ILN for all without first identifying whether or not patients need such radical treatment.

Developing the model structure

A decision tree was developed in TreeAge Pro 2001 software (TreeAge Software Inc., Williamstown, MA, USA). It was felt that a decision tree approach would be most appropriate for this economic evaluation for two reasons:

- The short time horizon.
- There are no examples of multiple recurrences (the same event happening to the same patient many times) seen in the model structure.

Women enter the model having been identified with a previous biopsy with presumed FIGO stage T1 or T2 unifocal vulval cancer (but not T1a). Only patients with SCC are considered because the vast majority of vulval cancer is SCC. Patients with basal cell carcinoma and MM are excluded from this analysis because of the different characteristics of these types of tumours.

The patients are assumed to follow one of seven different pathways that describe alternative approaches to the SLN biopsy and the treatment of vulval cancer. The first pathway is the comparison arm, which is the implementation of an IFL without a SLN biopsy and is used to show how this more morbid treatment compares to different SLN biopsy options. In the case of pathways 2–7, a SLN biopsy is performed using either blue dye, ^{99m}Tc, or both, in order to identify the SLN. This is followed by histopathology, which is some combination of H&E staining and ultrastaging in order to test for the presence of metastasis, in which ultrastaging can be considered to be representative of more sensitive techniques such as immunohistochemistry. For patients with metastasis, an IFL is performed, with RT also given when necessary (for example, Van der Zee *et al.*,[73] in which RT was given when more than one intranodal metastasis and/or extranodal growth was detected). Following an IFL, patients are considered to be

monitored every 3 months for 2 years for evidence of recurrence. If patients are only given a SLN biopsy and do not go on to IFL because no metastases were found, then patients are considered to be monitored every 2 months for 2 years. The seven patient treatment pathways are defined as follows:

1. *IFL*: this is performed on all patients, with no SLN biopsy being given.
2. *Blue dye test* and *H&E*: a blue dye test is administered intraoperatively to identify the SLN. This is followed by histopathology consisting of H&E staining of the SLN in order to identify the presence of metastasis.
3. *Blue dye test* and *ultrastaging*: a blue dye test administered intraoperatively to identify SLN, followed by histopathology consisting of H&E staining of the SLN. If no metastasis is detected then ultrastaging/staining is performed to confirm the absence or detect presence of metastases.
4. *^{99m}Tc* and *H&E*: a ^{99m}Tc test is administered, patient imaging preoperatively and then a radioactive probe used detect signal at surgery to identify the SLN. This is followed by histopathology consisting of H&E staining of the SLN to identify presence of metastasis.
5. *^{99m}TC* and *ultrastaging*: a ^{99m}Tc test administered, patient imaging preoperatively and then a radioactive probe used detect signal at surgery to identify the SLN. This is followed by histopathology consisting of H&E staining of the SLN. If no metastasis is seen then ultrastaging/staining is administered to confirm absence or detect presence of metastases.
6. *Blue dye* and *^{99m}Tc* and *H&E*: both blue dye and ^{99m}Tc test are administered to identify the SLN. Followed by H&E staining to identify the presence of metastasis.
7. *Blue dye* and *^{99m}Tc* and *ultrastaging*: both blue dye and ^{99m}Tc test are administered to identify the SLN, followed by H&E staining. If no metastasis is seen, then ultrastaging/staining is administered to confirm absence or detect presence of metastases.

Morbidity can occur in the short and long term as a result of complications due to a SLN biopsy or an IFL. Local recurrence can occur at any time following either of these procedures, the probability of which is informed by data (see *Table 42*), whereas groin recurrence may occur depending on the outcome of biopsy result and the treatment response. RT may be implemented alongside an IFL, or in response to a recurrence among patients who have not previously received it. Chemotherapy is assumed to be administered to all patients who have a recurrence but have already received an IFL and RT. These points are illustrated in *Figures 6–11*:

Model assumptions

A number of assumptions are required in order to develop a workable model structure and enable the analysis to be carried out. The assumptions made in this study are described below and grouped into those that refer to the general pathway, recurrence and the wider model:

General pathway: assumptions

- Patients found to be FN during the SLN biopsy (blue dye and/or ^{99m}Tc), but are then subsequently found to have metastasis receive both an IFL and RT.
- Patients are followed up every 2 months following a negative result for a SLN biopsy (and, therefore, are given no IFL) and every 3 months following an IFL.
- There are no occasions in which RT would be administered to a patient who had not previously received an IFL (apart from following a recurrence).
- Complications following a SLN biopsy (blue dye/^{99m}Tc) and then an IFL implemented during the same procedure will be the same as those experienced following just an IFL.
- Complications following all types of SLN biopsy (e.g. blue dye/^{99m}Tc) will be the same.

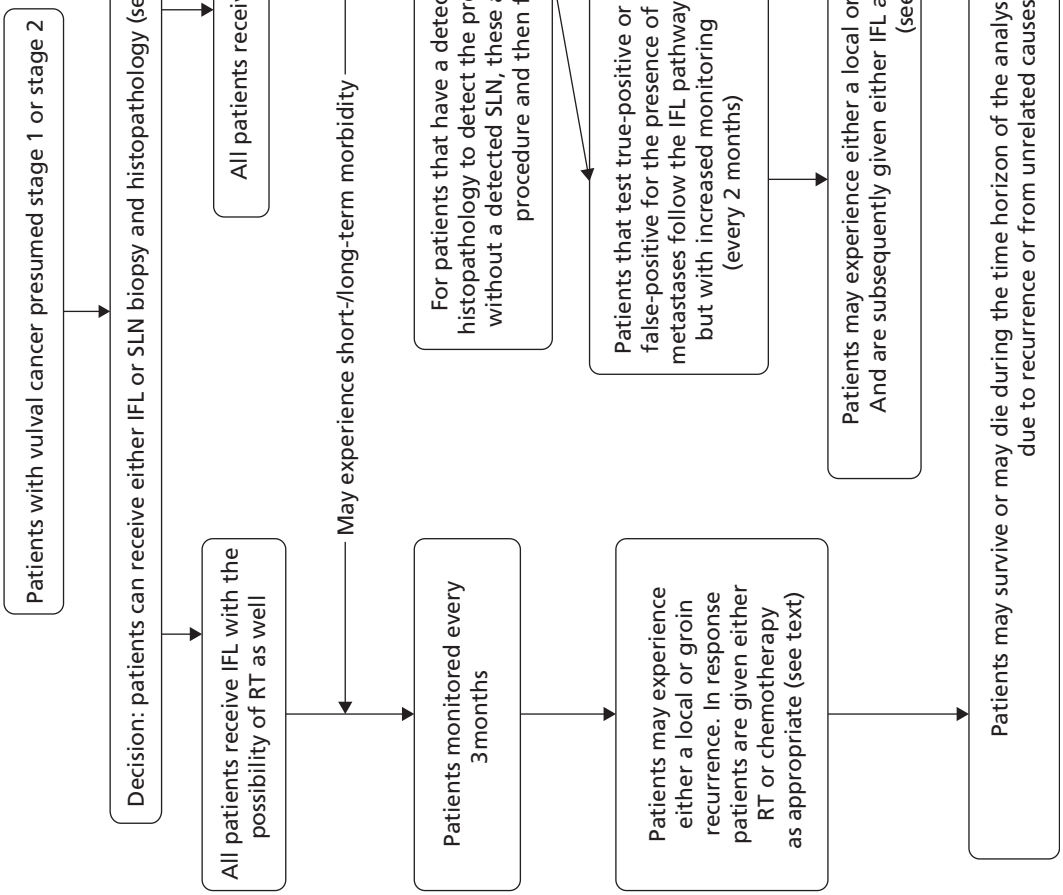


FIGURE 6 Summary of the decision pathway used in the economic model.

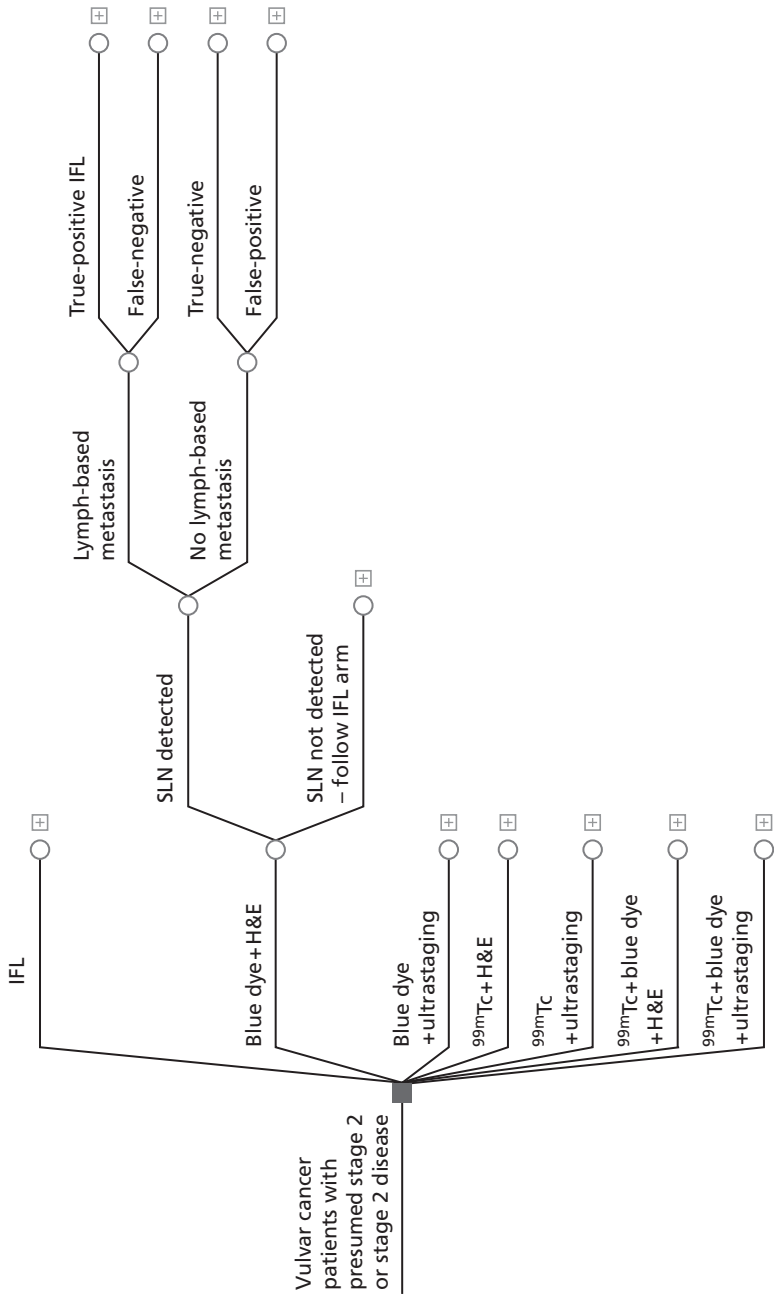


FIGURE 7 Model Structure showing each of the seven treatment pathways and the treatment pathway for blue dye + H&E. This pathway is repeated for each of the pathways that include either blue dye and/or ^{99m}Tc.

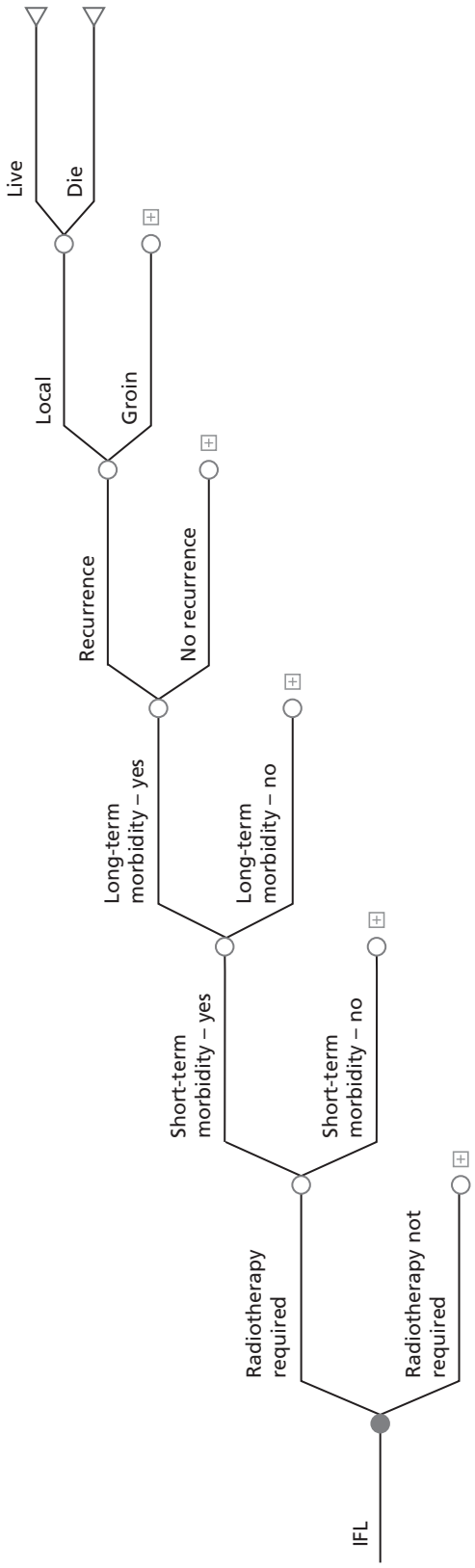


FIGURE 8 Treatment pathway for IFL. Note: patients who have a local recurrence are given a primary excision. Patients who have a groin recurrence following IFL + RT are given chemotherapy. Patients who have a groin recurrence following an IFL only are given RT.

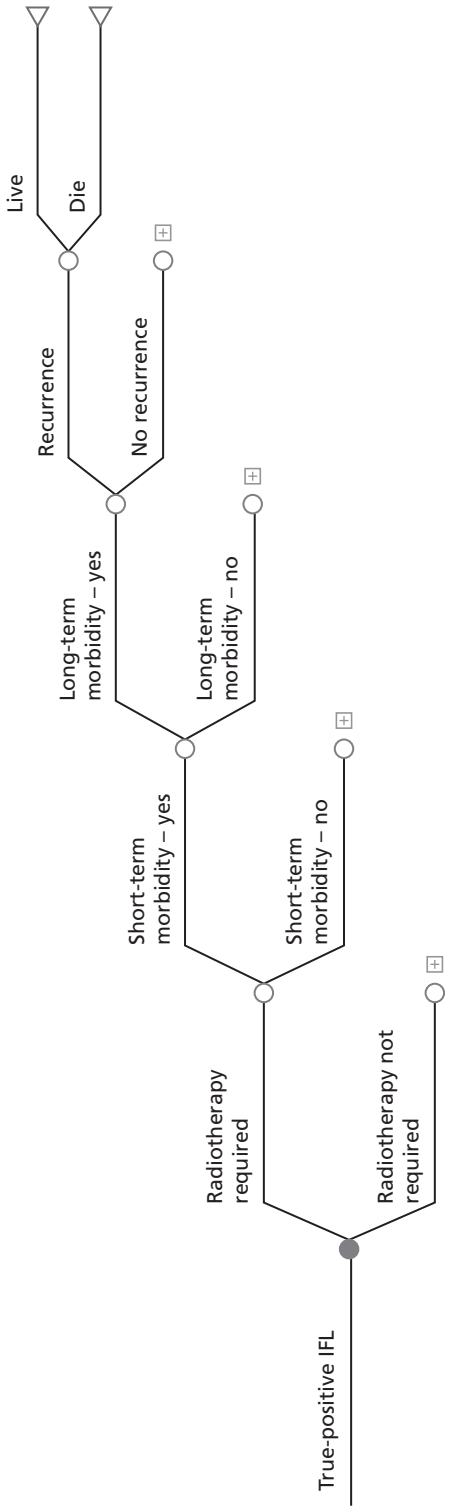


FIGURE 9 Treatment pathway for strategies 2–7 following a TP result for metastasis. Note: patients who have an IFL with or without RT may still go on to have either a local or groin recurrence. In the case of a local recurrence, all patients are given a primary excision. Patients who have a groin recurrence having previously received an IFL + RT are given chemotherapy, whereas for those patients who have only previously received a IFL RT is given.

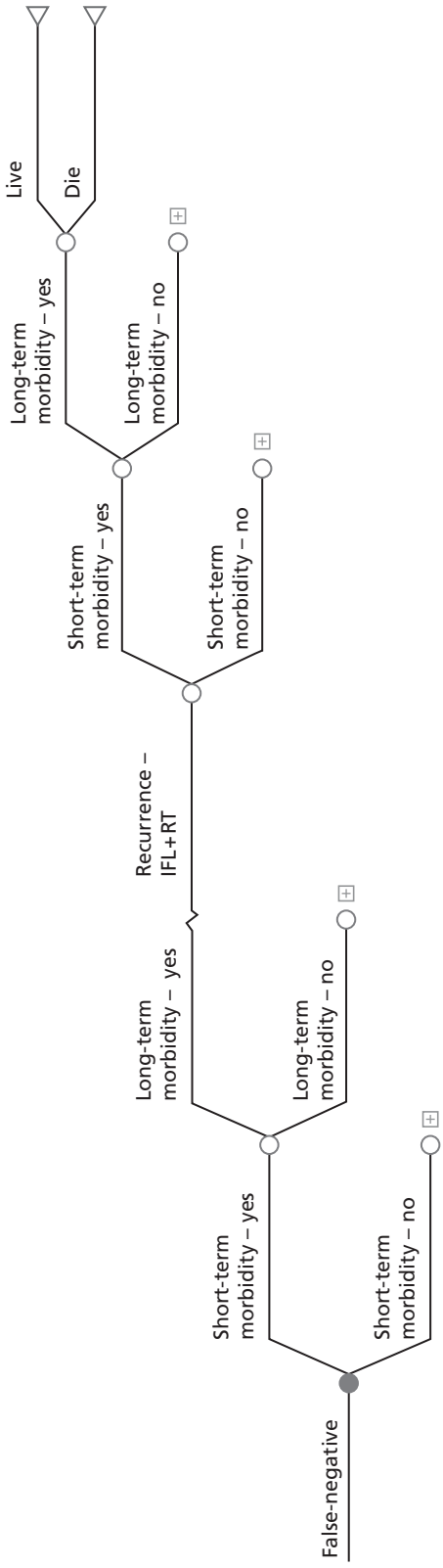


FIGURE 10 Treatment pathway for strategies 2–6 following a FN result for metastasis. Note: patients who have a FN biopsy test result, subsequently go on to have a groin recurrence. In this case, all the patients receive both an IFL + RT.

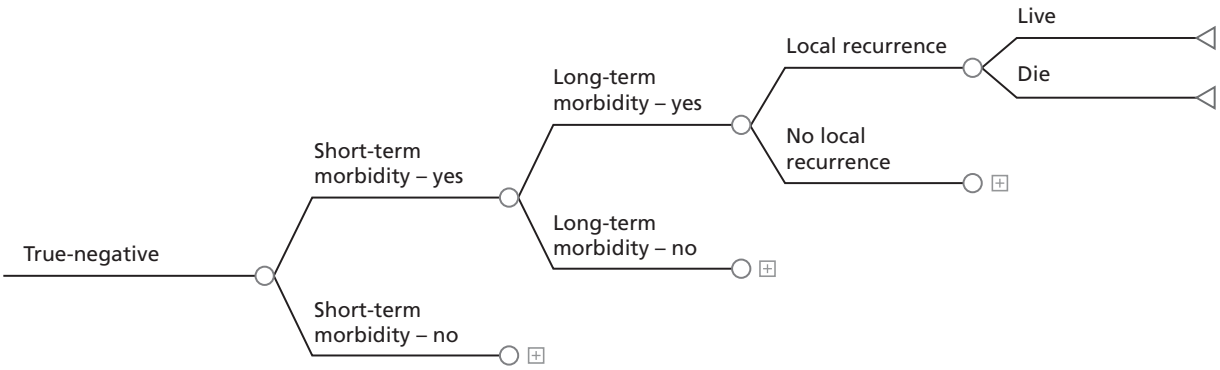


FIGURE 11 Treatment pathway for strategies 2–6 following a TN result for metastasis. Note: patients who have a TN biopsy test do not subsequently go on to have a groin recurrence. However, a local recurrence is still possible and under these circumstances the patients are given a primary excision.

Recurrence: assumptions

- Recurrence will occur only in the groin or the vulva (local); distant recurrence will not be considered. This is because any distant recurrences are likely to occur following either a local or groin recurrence and rarely occur without either.
- An additional primary excision will be required in the case of a local recurrence.
- In the case of groin recurrence, the treatment is IFL + RT if it has not been administered already and chemotherapy if it has.
- Mortality following recurrence within the 2-year time horizon is always due to vulval cancer, with these patients receiving palliative care as a result of their condition. Although it is acknowledged that the findings show that the death rate among vulval cancer patients due to vulval cancer or other causes is 50 : 50 following treatment, the risk of death following a recurrence is high enough for this assumption to be made.

Further modelling: assumptions

- For the purpose of costing follow-up, all deaths from vulval cancer and all other causes occur at 12 months following screening.
- All parameters in this model are independent of age, with the exception of the all-cause death rate. This assumption is made because of the paucity of age-specific data in this field.
- Patients experience long-term complications independently of whether or not they experience short-term complications. This assumption is made owing to the paucity of data in the literature describing what proportion of patients experience both short- and long-term complications.
- Short-term and long-term morbidity have no impact on the mortality of the patients. This assumption is made owing to the paucity of data; however, its impact is investigated through sensitivity analysis.
- All patients in the patient cohort are aged 65 years. The impact of this assumption is investigated through sensitivity analysis by examining patients aged 55 and 75 years, respectively.

Data requirements

The data requirements for the economic evaluation are fulfilled by using the findings of the systematic review. However, when the results of many studies could not be generalisable to obtain an overall parameter value owing to differences in the study protocols, etc., the findings of larger and more recent studies have been preferred.

Characteristics of the patient cohort

The age of the patients in this analysis will have an impact on their all-cause death rate over the 2-year time horizon of this study. Age 65 years was chosen as it is in broad agreement with the mean/median age of the patients in the included studies in the systematic review, who were in the range from 58–75 years (see *Table 9*). The proportion of patients with metastasis is an important parameter that is subsequently used in the model in two ways: it influences what will be seen in the results from the histopathology following the SLN biopsy and will have an effect on the probability that a patient may have a later recurrence. Patient cohort characteristics are described in *Table 39*.

Sentinel lymph node detection rate

Three approaches to the SLN biopsy are considered in this economic evaluation: blue dye and ^{99m}Tc alone, or both procedures implemented together. The aim of these procedures is to identify the SLN which can then be examined through histopathology to identify the presence/absence of metastasis. The following rates used in this study (*Table 40*) are informed by the findings of the systematic review (see *Table 24*).

Histopathology

Following the identification of the SLN(s) using a SLN biopsy, their histopathological assessment considered in this economic evaluation is the same as that described by Van der Zee *et al.*[73] H&E staining of the lymph node is used and then, if no metastasis is seen, ultrastaging with immunohistochemistry is then undertaken to confirm absence/presence of metastases. However, the impact of using routine histopathological examination alone is also considered. In this study, 80 out of 135 patients were found to have metastasis on routine examination (H&E) with the remaining positives (55 out of 135) being identified using ultrastaging (from Oonk *et al.*[83]). In Van der Zee *et al.*,[73] 6 out of 259 patients with unifocal vulval cancer and a negative SLN following H&E and ultrastaging were subsequently diagnosed with a groin recurrence, which can be used to help inform the probability of a FN test. Taking the prevalence of metastasis to be 33.5% (135 out of 403, see *Table 39*), the estimated probabilities for the different histopathology test results can be calculated as follows:

Calculation of histopathology test accuracy parameters

The following calculations describe the possible outcomes for patients who have a detected SLN that is then subject to histopathology.

As previously described, two approaches to histopathology are considered, these being H&E + ultrastaging and H&E alone. These are each considered in turn.

TABLE 39 Parameters describing the characteristics of patient cohort

Parameter	Value (range)	Reference	Notes
Patients with metastasis	33.5%	Van der Zee *et al.*, 2008[73]	–
Age of cohort (years)	65	–	Examined in sensitivity analysis

TABLE 40 Sentinel lymph node detection rates by SLN biopsy

SLN biopsy	Detection rate
Blue dye	202/294 (68.7%)
^{99m}Tc	227/240 (94.5%)
Blue dye + ^{99m}Tc	1049/1074 (97.7%)

Haematoxylin and eosin stain + ultrastaging Among patients with an identified SLN that was subject to H&E + ultrastaging, Van der Zee *et al.*[73] describe that 6 out of 259 patients with unifocal vulval cancer and a negative SLN following H&E and ultrastaging were subsequently diagnosed with a groin recurrence (see *Table 41*); therefore, by definition:

$$\text{Number of FNs} = 6 \tag{5}$$

$$\text{Number of TNs} = 253 \tag{6}$$

Taking these values for FN and TN into account, this means that the assumed prevalence of metastasis in this study cannot fall below 2.3% (6/259).

The number of disease (metastasis) negative (DN) can be calculated from the sum of the patients who test TN and FP for metastasis (TN + FP):

$$\text{DN} = \text{TN} + \text{FP} \tag{7}$$

However, given that the systematic review failed to find any evidence of patients testing FP for metastasis (see *Tables 13–23*), it is assumed here that FP = 0.

As described in *Table 38*, the proportion of patients with metastasis (p) is taken to be 0.335 (135/403). From this, the number of disease (metastasis) positive (DP) can be calculated:

$$\text{DP} = \text{DN} \times p/(1-p) \tag{8}$$

Taking the previously calculated values for FN and DP, the number of TPs can now be calculated:

$$\text{Number of TPs} = \text{DP} - \text{FN} \tag{9}$$

Now that we have values for FN, TN, FP and TP which are based values obtained from the literature that include the assumed prevalence of metastasis, it is straightforward to calculate what proportion of patients will test for each of these possibilities for H&E + ultrastaging, with the baseline parameters shown in *Table 41*.

Haematoxylin and eosin stain alone As part of the study by Oonk *et al.*,[83] 80 out of 135 patients who were found to have metastasis were found positive by H&E. This gives the sensitivity of H&E ($\text{Sens}_{\text{H\&E}}$) to be 59% (80 out of 135).

TABLE 41 Test accuracy parameters describing outcomes of H&E and ultrastaging among patients with an identified SLN

Test result	H&E	H&E if negative then ultrastaging
FN	13.6%	1.6%
TN	66.5%	66.5%
FP	0.0%	0.0%
TP	19.9%	31.9%

Taking the number of DP described above for H&E + ultrastaging, the number of TPs and FNs detected by H&E can be given as follows:

$$\text{Number of TPs } (TP_{H\&E}) = Sens_{H\&E} \times DP \quad (10)$$

$$\text{Number of FNs } (FN_{H\&E}) = DP - TP_{H\&E} \quad (11)$$

Taking the number of TNs detected by H&E ($TN_{H\&E}$) to be the same as for H&E + ultrastaging and again assuming that $FP_{H\&E} = 0$.

The values for $FN_{H\&E}$, $TN_{H\&E}$, $FP_{H\&E}$ and $TP_{H\&E}$ allow the proportion of patients who test for each of these possibilities for H&E alone to be calculated, with the values used at baseline shown in *Table 41*.

In all cases the values taken from the literature, e.g. prevalence of metastasis, are varied as part of the probabilistic sensitivity analysis (PSA) in order to show their impact on the model results.

Cancer recurrence

A recurrence of cancer may occur along any of the patient pathways with varying probabilities. These are summarised in *Table 42*.

Survival following treatment

Patient death is categorised as occurring as a result of a vulval cancer recurrence (local or groin) or from all other causes. All-cause death depends on the assumed age of the cohort with values calculated for the 2-year time horizon shown in *Table 43*.

Radiotherapy

Radiotherapy may be given to patients following an IFL or following a recurrence. However, in all situations, RT is never administered to the same patient more than once. RT is assumed to always be implemented to a patient following a recurrence that has not previously received it, with the probability of RT at other points in the decision pathway being informed by data. This is summarised in *Table 44*.

TABLE 42 Parameters describing the probabilities of recurrence used in the economic evaluation

Parameter	*n*/*N* (%)	Reference	Notes
Local recurrence	34/276 (12.3)	Van der Zee *et al.*, 2008[73]	The possibility of local recurrence is present in all arms of the model
Groin recurrence following IFL (no SLN biopsy)	1/32 (3.1)	Crosbie *et al.*, 2010[54]	Metastasis prevalence in this study found to be 6/31. Relative risk of groin recurrence given metastasis = 0.1546
Groin recurrence following negative SLN biopsy result	6/259 (2.3)	Van der Zee *et al.*, 2008[73]	Patients with unifocal disease
Groin recurrence following positive SLN biopsy and IFL	11/135 (8.1)	Oonk *et al.*, 2010[83]	–
Groin recurrence following FN test	100%	–	All patients will get a recurrence if the test is falsely negative

See *Chapter 5, Recurrences rates*, for further information.

TABLE 43 Death probabilities used in the economic evaluation

Death rate	Percentage	Reference	Note
Local recurrence	5/34 (14.75%)	Van der Zee *et al.*, 2008[73]	–
Groin recurrence	9/11 (81.8%)	Oonk *et al.*, 2010[83]	–
All cause	Age 55 years: 0.84% Age 65 years: 1.97% Age 75 years: 5.85%	Office for National Statistics (2010)[103]	Calculated from: Natural Death rates. Mid-year estimates published 30 June 2011

TABLE 44 Parameters describing the probability of patients requiring RT, depending on the clinical pathway

RT	Percentage	Reference	Notes
With an IFL Strategy 1	26/56 (46.4%)	Fonseca-Moutinho *et al.*, 2000[104]	–
After a TP SLN biopsy result and IFL	49/117 (41.9%)	Van der Zee *et al.*, 2008[73]	–
After a FP biopsy result and IFL	0%	–	See *Chapter 7, Model assumptions*
Following a recurrence if not previously administered	100%	–	See *Chapter 7, Model assumptions*

Morbidity and complications

The reported complications following a procedure are used as a proxy for the additional morbidity experienced by patients. This in turn will have an impact on the costs to the health-care provider as a result of the extra resources needed to treat the patients. The definitions of short- and long-term morbidity in this study are the same as those proposed by Van der Zee *et al.*,[73] in which short-term morbidity is defined as the occurrence of wound breakdown or wound infection (requiring antibiotics) and long-term morbidity is defined as lymphoedema present over two consecutive visits more than 1 year after primary therapy or recurrent erysipelas (more than one episode of erysipelas requiring antibiotics) (see *Chapter 5, Recurrence rates*).

Complication-related morbidity can occur following an IFL or SLN biopsy in the short or long term. The percentage of patients with short- and long-term complications is calculated assuming that the probability of experiencing one type of short-term (or long-term) complication is independent of experiencing another at the same time and that the probability of experiencing long-term morbidity is independent of whether or not short-term morbidity was previously experienced (*Table 45*). This is informed by the findings of the study by Van der Zee *et al.* 2008.[73] These assumptions were made owing to the lack of information on the proportion of patients who have more than one complication at the same time or who experience both short-term and, then, long-term complications.

TABLE 45 Parameters describing the probability of short- or long-term morbidity by treatment implemented (adapted from Van der Zee *et al.*[73])

Time frame	Procedure	Complication	% patients with complications (*n*/*N*)
Short term	IFL (with/without SLN biopsy)	Wound breakdown 34%	48.1 (22.6/47)
		Wound cellulitis 21.3%	
	SLN biopsy	Wound breakdown 11.7%	15.7 (41.4/264)
		Wound cellulitis 4.5%	
Long term	IFL (with/without SLN biopsy) and RT	Lymphoedema 25.5%	48.3 (23.7/49)
		Recurrent erysipelas 30.6%	
	IFL (with/without SLN biopsy) no RT	Lymphoedema 25.5%	29.9 (20.9/70)
		Recurrent erysipelas 5.9%	
	SLN biopsy	Lymphoedema 1.9%	2.3 (6.1/264)
		Recurrent erysipelas 0.4%	–

Cost and resource data

Three sources of data were used to parameterise the cost component of the economic evaluation: NHS Reference Costs 2009/2010,[105] information provided by the Histology Department at Birmingham City Hospital and data collected as part of United Kingdom Gynaecological Oncology Surgical Outcomes and Complications audit (UKGOSOC). The UKGOSOC is a prospective web-based audit looking at outcomes of surgery in gynaecological oncology, particularly focused on complications. All costs in this study are presented in values for the year 2010. In all cases, patients only require a maximum of one unit of each cost depending on their treatment pathway, with the itemised costs shown in *Table 46*.

All patients receive a radical excision of the primary vulval cancer which is administered along with either the SLN biopsy or IFL depending on the treatment pathway. The prices for the SLN biopsies and IFL include the cost of the radical excision; however, as a radical excision is also administered in the case of a vulval recurrence, a separate price is also given.

Costs are given for blue dye and ^{99m}Tc when administered separately; however, no cost was available for the two procedures combined and, therefore, it is assumed that the cost for both blue dye and ^{99m}Tc is 10% greater than ^{99m}Tc alone (the more expensive of the two). The impact of this assumption on model results is examined through sensitivity analysis.

Outcomes

The main focus of this economic evaluation is how the different treatment scenarios impact on the mortality and morbidity of the patients. Therefore, the following outcomes have been examined in this analysis:

- case of death avoided within 2 years
- case of morbidity-free survival within 2 years
- case of long-term morbidity-free survival within 2 years.

TABLE 46 Cost of items incorporated in the economic evaluation

Item	Code	Cost (£)	Reference	Assumption
Radical excision	MB01B	1971	UKGOSOC data	3.86 bed-days
IFL (+ radical excision)	MA06Z	4129	UKGOSOC data	5.64 bed-days
RT	SC22Z + SC56Z	1728	NHS reference costs[105]	'Day case and regular day/night' 3 weeks of treatment, 5 days each week (assumption)
Chemotherapy	SB12Z + SB15Z	1270	NHS reference costs[105]	'Inpatient'. Assume drugs from regime in band 6 procurement + delivery £779 + £207 + £284
Monitoring of patients	503	171	NHS reference costs[105]	Per consultation
SLN biopsy (+ radical excision)				
Blue dye	MA06Z	3574	UKGOSOC data	3.86 bed-days
^{99m}Tc	MA06Z + RA36Z	3836	UKGOSOC data	3.86 bed-days
Blue dye + ^{99m}Tc	MA06Z + RA36Z	4219	UKGOSOC data	3.86 bed-days 10% greater than ^{99m}Tc (assumption)
Morbidity and mortality				
Short term	MA06Z	1635	NHS reference costs[105]	5.24 bed-days
Long term	MA06Z TPCTCLFUMFF 502 gynaecology	702	NHS reference costs[105]	Three outpatient visits + 1 bed-day (assumed)
Vulval cancer related death	SD01A	436	NHS reference costs[105]	Specialist palliative care, inpatient
Histopathology				
H&E		74.50	Histology Department Birmingham City Hospital	–
Ultrastaging		86.75	Histology Department Birmingham City Hospital	–

In each case, these are compared again the comparison scenario of implementing an IFL to all patients.

As has been highlighted in the systematic reviews, there are no studies that have measured a generic QoL such as Euroqol EQ-5D in vulval cancer. Furthermore, the only study to measure global health status (QoL)[82] showed no difference for patients who received the different procedures that are considered in this analysis. Therefore, the option of using the quality-adjusted life-year (QALY) as an outcome measure was not available in this economic evaluation.

The outcome measures in this study are considered to be reasonable given that the main focus of this study is whether or not a negative result from a SLN biopsy and no further treatment, with the possibility of recurrence, is preferable to the highly morbid IFL which has a lower risk of recurrence after the procedure.

Analysis

The model used in the economic evaluation begins with a hypothetical cohort of women who are presumed to have FIGO stage I or stage II vulval cancer. The model estimates the mean costs associated with each of the treatment strategies and the base case assumes that all women entering the model are aged 65 years. The time horizon of the model is 2 years, which was chosen as it was felt that any groin recurrences that might appear as a result of a FN SLN biopsy would be detected within this time frame. Owing to this short time horizon, and with the majority of costs occurring in the first year, no discounting was applied. This economic evaluation takes the form of a cost-effectiveness analysis and is carried out from the UK NHS perspective. Therefore, only direct costs and resources associated with the intervention and outcomes are incorporated in the analysis. The results of the analysis are presented in terms of incremental cost-effective ratios (ICERs) for each of the three outcome measures considered.

Sensitivity analysis

The results described by the cost-effectiveness point estimates do not consider any uncertainty in relation to the model input parameters. PSA was therefore undertaken to assess the impact of the uncertainty in the model parameters on the results and conclusions obtained from the model. The costs in the model are all unit costs for specific procedures and are treated as fixed; however, the number of bed-days was varied. The probabilities in the tree, these being the proportions of patients who follow each branch, were also varied.

The standard distribution used in this analysis for the proportions is the beta distribution (*Table 47*). The beta distribution is described by two parameters, a and b. A beta distribution is able to precisely represent the uncertainty in a proportion when the only available information is alpha-positive cases and beta-negative cases. In all cases, in this study, exact numbers were available and so these were used to inform the parameters of each beta distribution directly. The bed-days were described by a gamma distribution. The method of moments approach was used to estimate the parameters of the gamma distribution, where:

$$\alpha = \frac{(\text{mean}^2)}{(\text{standard error}^2)} \tag{12}$$

$$\beta = \frac{(\text{mean})}{(\text{standard error}^2)} \tag{13}$$

The following one-way sensitivity analyses were also carried out.

- Age of the cohort: as a baseline. It is assumed that the age of the cohort is 65 years. The impact of the alternative ages of 55 and 75 years is examined as the all-cause death rate of these groups may impact on the conclusions drawn from the model.
- Increased mortality due to patient morbidity. It has been assumed in this study that the morbidity experienced by the patients at baseline has no impact on their overall survival. In this sensitivity analysis, this assumption is relaxed, and instead it is assumed that, for all pathways in which the patients experience morbidity, their death rate is increased by 20%. To balance this out, the death rates for pathways in which there is no morbidity are reduced by 20%. The purpose of this analysis is solely to illustrate how this assumption may impact on the final conclusions drawn from the model.
- Varying the cost of implementing ^{99m}Tc and blue dye together. It is assumed that the cost of implementing ^{99m}Tc and blue dye together is 10% more than the price of ^{99m}Tc alone and is £4219. The impact of this assumption on model results is examined by instead assuming that implementing ^{99m}Tc and blue dye together is equal to the cost of ^{99m}Tc (£3836) and by assuming that it costs 50% more than ^{99m}Tc (£5754).
- Groin recurrence rate following a negative SLN biopsy result. Van der Zee *et al.*[73] found that 6 out of 259 (2.3%) patients with unifocal disease experienced a groin recurrence following a negative SLN

TABLE 47 Distributions used in the PSA

Parameter	Distribution	Alpha	Beta
Patients with metastasis	Beta	135	268
Blue dye detection rate	Beta	202	92
^{99m}Tc detection rate	Beta	227	13
Blue dye + ^{99m}Tc detection rate	Beta	1050	25
Negative predictive value of H&E + ultrastaging	Beta	253	6
Sensitivity of H&E	Beta	80	55
Local recurrence	Beta	34	242
Groin recurrence following IFL (no SLN biopsy)	Beta	1	31
Groin recurrence following positive SLN biopsy and IFL	Beta	11	124
Death following a local recurrence	Beta	5	29
Death following a groin recurrence	Beta	9	2
RT following IFL in the comparison arm	Beta	26	30
RT with IFL following a TP histopathology result	Beta	49	68
Short-term morbidity following IFL	Beta	22.6	24.4
Short-term morbidity following SLN biopsy	Beta	41.4	222.6
Long-term morbidity following IFL + RT	Beta	23.7	25.3
Long-term morbidity following IFL without RT	Beta	20.9	49.1
Long-term morbidity following SLN Biopsy	Beta	6.1	257.9
Bed-days following a primary excision/SLN biopsy	Gamma	1.925	2.007
Bed-days following a IFL	Gamma	3.504	1.6103

biopsy result. Owing to the size of this study, this was used as the baseline figure in this economic evaluation. However, the results from the systematic review suggest that this result may be slightly low (see *Chapter 5, Recurrence rates*, in the test accuracy systematic review). A study by Moore *et al.*[65] found that 2 out of 31 (6.5%) patients experienced a groin recurrence following a negative SLN biopsy. This value was not used at baseline owing to the small sample size and the lack of clarity as to whether or not the disease was unifocal. The impact of this alternative value on the model results is examined.

Results

The base-case deterministic results for the seven different treatment strategies are calculated based on the following outcomes additional cost per case of patient survival at 2 years, cost per case of morbidity-free survival at 2 years and cost per case of survival free of long-term morbidity at 2 years. Incremental cost-effectiveness analysis then follows.

Deterministic results: base case

Table 48 shows the deterministic results obtained from the model. This shows the cost of each treatment strategy and its effectiveness in terms of each of the three outcome measures.

Outcomes

Overall survival at 2 years following IFL was found to be the most effective strategy. This result is not surprising given that this procedure seeks to reduce the potential for future recurrences at the expensive of

TABLE 48 Deterministic analysis results for the seven strategies across three different outcome measures

Strategy	Cost (£)	Overall survival at 2 years	Morbidity-free survival at 2 years	Survival free of long-term morbidity at 2 years
IFL	9367	0.9645	0.3512	0.6423
Blue dye + ultrastaging	9775	0.9427	0.5241	0.7534
Blue dye + H&E	9826	0.8782	0.5015	0.7105
^{99m}Tc + ultrastaging	10,175	0.9345	0.6054	0.7985
^{99m}Tc + H&E	10,245	0.8457	0.5744	0.7395
^{99m}Tc + blue dye + ultrastaging	10,576	0.9335	0.6151	0.8039
^{99m}Tc + blue dye + H&E	10,648	0.8418	0.5830	0.7430

increased patient morbidity. For all types of morbidity-free survival, the ^{99m}Tc + blue dye + ultrastaging strategy was found to be the most effective. Again this is not a surprising result given that this uses the most robust procedures for identifying both the SLN (highest detection rate) and metastasis (highest sensitivity).

Costs

The IFL strategy was found to be the cheapest, costing £9367 per woman treated for presumed type I or type II vulval cancer. The most expensive was found to be the ^{99m}Tc + blue dye + H&E strategy, costing £10,648 per patient. Although this is not the most expensive treatment to administer, it is likely that these costs are due to extra costs associated with undetected recurrences that will occur with this type of treatment regimen.

Incremental analysis

A strategy is dominated by an alternative strategy if it is more expensive and less effective. It is not normally necessary to consider dominated strategies since they are supplanted by other strategies that are more cost-effective. The analysis below describes the treatment pathways that are dominated and, then, an incremental analysis is undertaken for the remaining strategies.

Results: overall survival at 2 years outcome

As can be seen in *Figure 12* for the overall survival at 2 years outcome, the IFL strategy dominates all other strategies as it is both less expensive and averts the greatest mortality. Therefore, a further incremental analysis for overall survival for the baseline deterministic results is not undertaken.

Results: overall morbidity-free survival at 2 years outcome

For the outcome of morbidity-free survival at 2 years, only the strategies of blue dye + ultrastaging, ^{99m}Tc + ultrastaging and ^{99m}Tc + blue dye + ultrastaging remain undominated by any of the alternative treatment strategies. *Table 49* presents the deterministic analysis for the morbidity-free survival at 2 years outcome restricted to the non-dominated competing strategies.

In terms of morbidity-free survival at 2 years, the least expensive strategy is the base-case scenario of IFL. The most effective strategy is the ^{99m}Tc + blue dye + ultrastaging strategy, but this comes at a greater cost, generating an ICER of £41,200, i.e. the strategy requires an investment of £41,200 to generate one additional case of morbidity-free survival compared with the strategy of ^{99m}Tc + ultrastaging. The strategy of ^{99m}Tc + ultrastaging is both slightly less effective in terms of overall morbidity-free survival and slightly less costly than ^{99m}Tc + blue dye + ultrastaging. The ICER for ^{99m}Tc + ultrastaging is approximately £4900, i.e. a financial outlay of £4900 is necessary to generate one additional case of morbidity-free survival compared with the strategy of the blue dye + ultrastaging. The ICER for blue dye + ultrastaging is approximately £2400, i.e. a financial outlay of £2400 is necessary to generate one additional case of morbidity-free survival compared with the strategy of IFL.

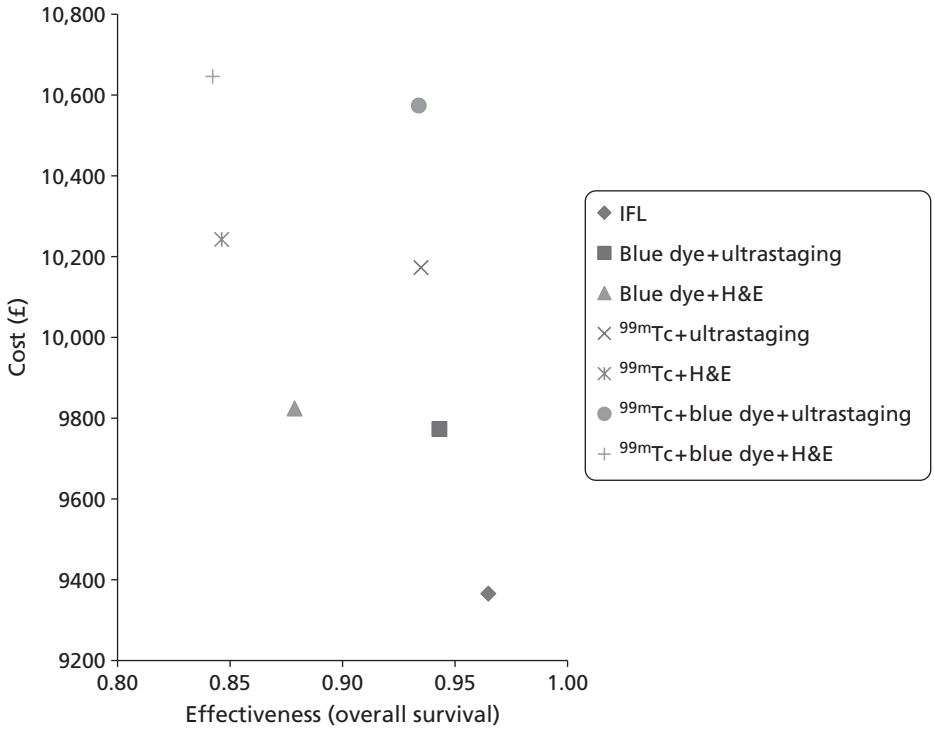


FIGURE 12 Cost-effectiveness plane showing the results of the deterministic analysis for the seven strategies with overall survival at 2 years as the outcome measure.

TABLE 49 Deterministic results for the non-dominated strategies for morbidity-free survival at 2 years

Strategy	Cost (£)	Incremental cost (£)	Effectiveness	Incremental effectiveness	ICER (£)
IFL	9367	–	0.3512	–	–
Blue dye + ultrastaging	9775	408	0.5241	0.1729	2400
^{99m}Tc + ultrastaging	10,175	400	0.6054	0.0813	4900
^{99m}Tc + blue dye + ultrastaging	10,576	400	0.6151	0.0097	41,200

Incremental cost and effectiveness calculated with respect to the strategy on the previous line.

Figure 13 shows the total costs and the effectiveness in terms of morbidity-free survival for the different treatment strategies considered in this analysis. The line on the graph joins the non-dominated strategies of blue dye + ultrastaging, ^{99m}Tc + ultrastaging and ^{99m}Tc + blue dye + ultrastaging. Any strategy that appears above this line is not considered cost-effective in relation to the non-dominated alternatives.

It is noted that the strategies blue dye + H&E, ^{99m}Tc + H&E and ^{99m}Tc + blue dye + H&E are sufficiently close to the boundary of dominance that the impact of parameter uncertainty on the model results between these dominated alternatives should be examined in addition to the non-dominated options.

Results: long-term morbidity-free survival at 2 years outcome

The dominance seen for morbidity-free survival at 2 years is repeated for long-term morbidity-free survival at 2 years in that the strategies of blue dye + ultrastaging, ^{99m}Tc + ultrastaging and ^{99m}Tc + blue dye + ultrastaging remain undominated by any of the alternative treatment strategies (*Figure 14*).

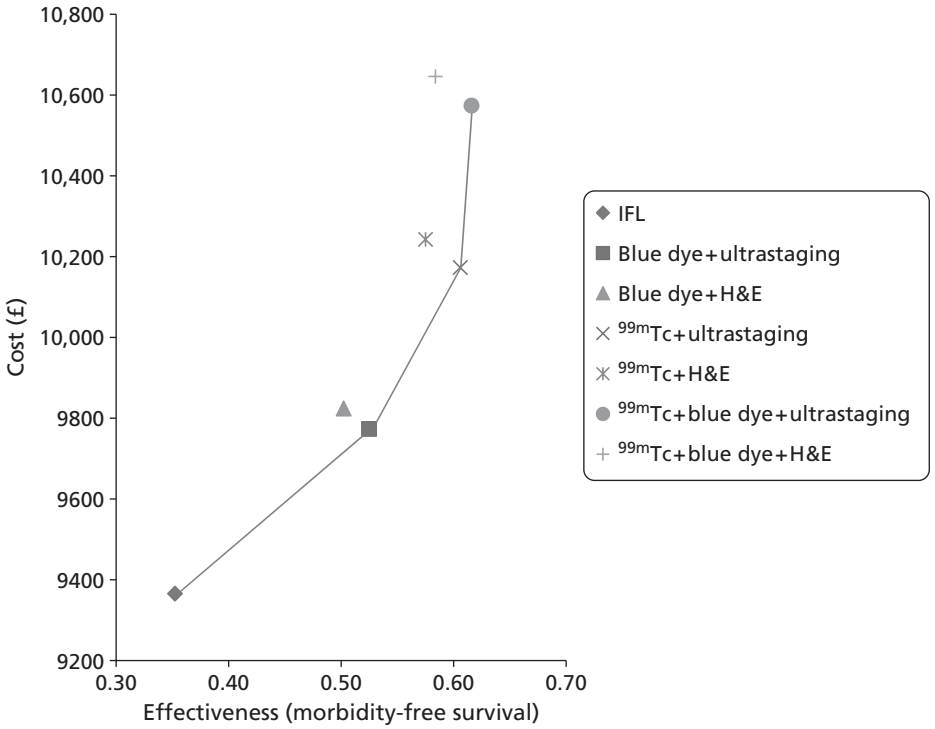


FIGURE 13 Cost-effectiveness plane showing the results of the deterministic analysis for the seven strategies with morbidity-free survival at 2 years as the outcome measure.

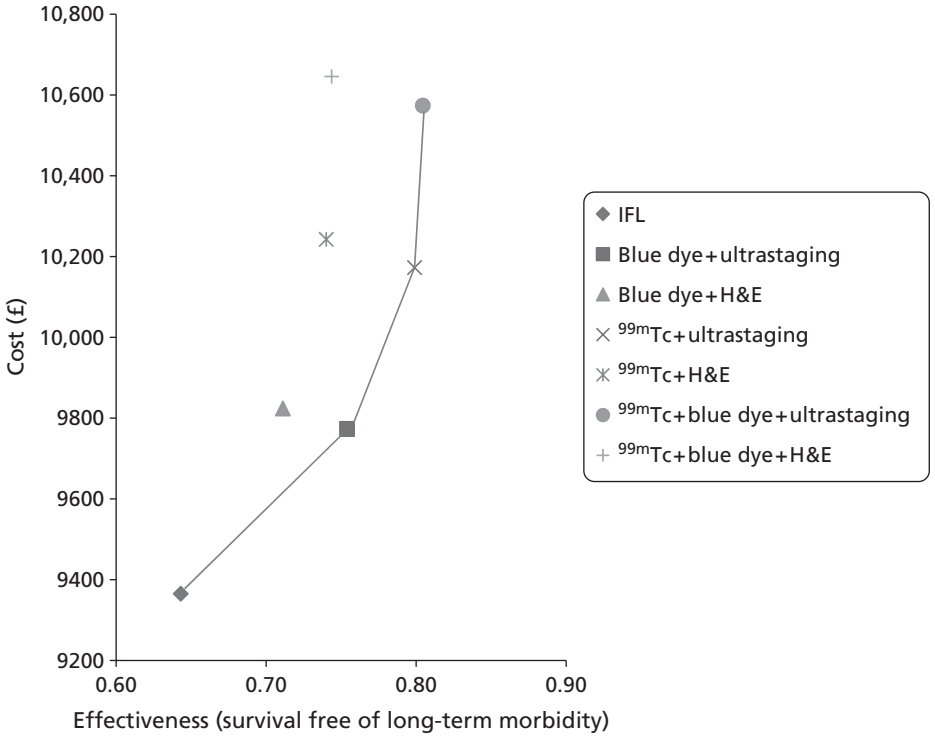


FIGURE 14 Cost-effectiveness plane showing the results of the deterministic analysis for the seven strategies with long-term morbidity-free survival at 2 years as the outcome measure.

Considering the outcome measure to be long-term morbidity-free survival at 2 years does nothing to change the costs of the different strategies and so the cheapest strategy is still the base-case scenario of IFL (*Table 50*). The most effective strategy is the ^{99m}Tc + blue dye + ultrastaging strategy, but this comes at a greater cost, generating an ICER of £74,300, i.e. the strategy requires an investment of £74,300 to generate one additional case of long-term morbidity-free survival compared with the strategy of ^{99m}Tc + ultrastaging. The strategy of ^{99m}Tc + ultrastaging is both slightly less effective in terms of overall long-term morbidity-free survival and slightly less costly than ^{99m}Tc + blue dye + ultrastaging. The ICER for ^{99m}Tc + ultrastaging is approximately £8900, i.e. an additional financial outlay of £8900 is necessary to generate one case of long-term morbidity-free survival compared with the strategy of blue dye + ultrastaging. The strategy of blue dye + ultrastaging is again both slightly less effective in terms of overall long-term morbidity-free survival and slightly less costly than ^{99m}Tc + ultrastaging. The ICER for blue dye + ultrastaging is approximately £3700, i.e. an additional financial outlay of £3700 is necessary to generate one case of long-term morbidity-free survival compared with the strategy of IFL.

Figure 14 shows the total costs and the effectiveness in terms of long-term morbidity-free survival at 2 years for the different treatment strategies considered in this analysis. The line on the graph joins the non-dominated strategies of blue dye + ultrastaging, ^{99m}Tc + ultrastaging and ^{99m}Tc + blue dye + ultrastaging. Any strategy that appears above this line is not considered cost-effective in relation to the non-dominated alternatives.

Once again, it is noted that the strategies blue dye + H&E, ^{99m}Tc + H&E and ^{99m}Tc + blue dye + H&E are sufficiently close to the boundary of dominance that the impact of parameter uncertainty on the model results between these dominated alternatives should be examined in addition to the non-dominated options.

Probabilistic sensitivity analysis

The PSA is undertaken to assess the impact of the uncertainty in the model parameters on the results and conclusions obtained from the model. As with the deterministic analysis, the outcome measures of the overall survival at 2 years, morbidity-free survival at 2 years and survival free long-term morbidity at 2 years are considered.

Probabilistic sensitivity analysis: overall survival at 2 years outcome

Figure 15 illustrates the overall uncertainty in the model results with the outcome measure being overall survival at 2 years. It is clear from the degree of overlap of the results obtained from the different treatment strategies that there is uncertainty regarding which one may be considered most cost-effective when a range of values is sampled from the distributions that describe the data values. This output is therefore used to examine the overall uncertainty related to the optimal decision across a range of plausible willingness-to-pay (WTP) values, in which, for this outcome, the WTP is measured in pounds per additional case of survival achieved.

Table 50 Deterministic results for the non-dominated strategies for overall long-term morbidity-free survival at 2 years

Strategy	Cost (£)	Incremental cost (£)	Effectiveness	Incremental effectiveness	ICER (£)
IFL	9367		0.6423		
Blue dye + ultrastaging	9775	408	0.7534	0.1111	3700
^{99m}Tc + ultrastaging	10,175	400	0.7985	0.0451	8900
^{99m}Tc + blue dye + ultrastaging	10,576	400	0.8039	0.0054	74,300

Incremental cost and effectiveness calculated with respect to the strategy on the previous line.

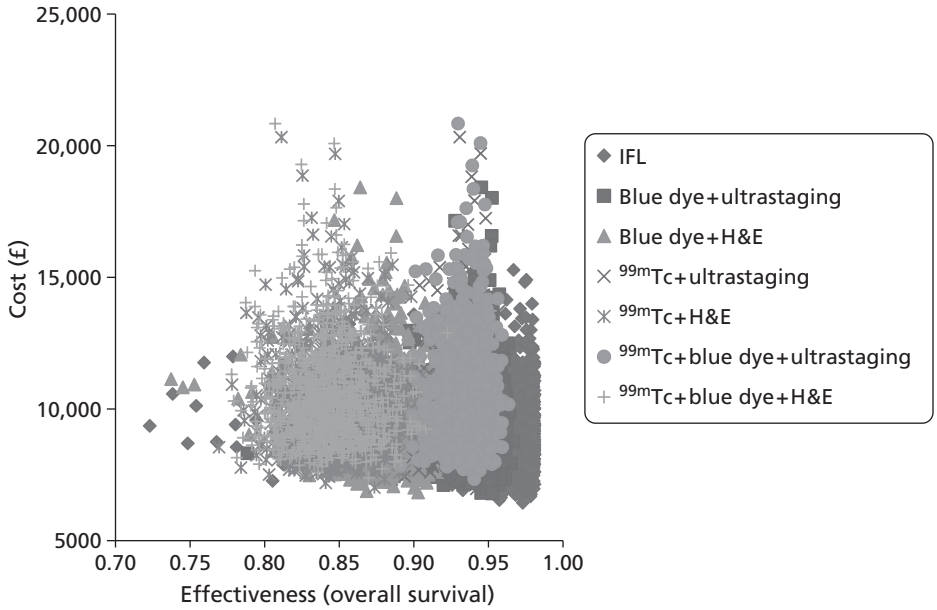


FIGURE 15 Scatterplot showing the uncertainty in costs and effectiveness within the model for each of the seven strategies for 1000 runs with overall survival at 2 years as the outcome measure.

Figure 16 shows the cost-effectiveness acceptability frontier (CEAF) for the outcome of additional case of survival achieved and is generated as follows. First, for any value of the WTP, the optimal solution is obtained based on the mean results. Then the proportion of model replications for which that was the optimal solution was found and plotted. By definition, only the strategies that have been shown not to be dominated can appear on the CEAF. For the cost per case of survival achieved outcome, it can be seen that IFL is the optimal treatment strategy for all values of the WTP up to (and beyond) £100,000.

Probabilistic sensitivity analysis: morbidity-free survival outcome

Figure 17 illustrates the overall uncertainty in the model results for the outcome measure of morbidity-free survival at 2 years. Once again it is clear from the degree of overlap in the results obtained from the different treatment strategies that there is uncertainty regarding which one might be considered most cost-effective when a range of values is sampled from the distributions that describe the data values. This output is therefore used to examine the overall uncertainty related to the optimal decision across a range of plausible WTP values, in which, for this outcome, the WTP is measured in pounds per additional case of morbidity-free survival achieved.

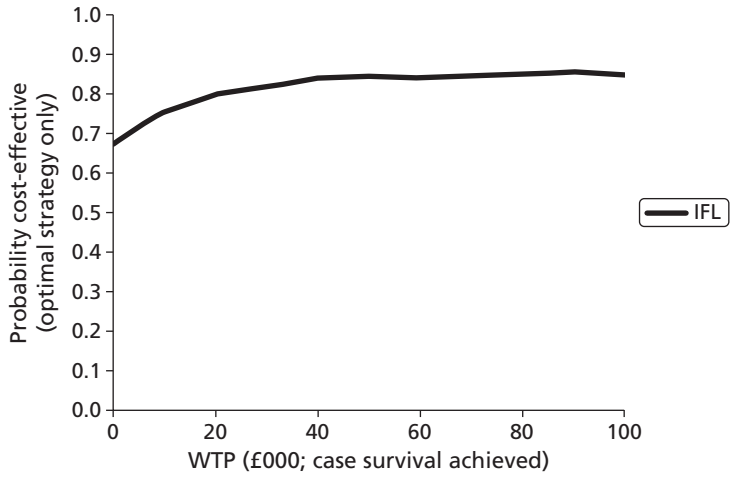


FIGURE 16 Cost-effectiveness acceptability frontier showing the results of the sensitivity analysis examining the optimal investigative strategy across a range of WTP thresholds for the outcome of case of survival at 2 years achieved.

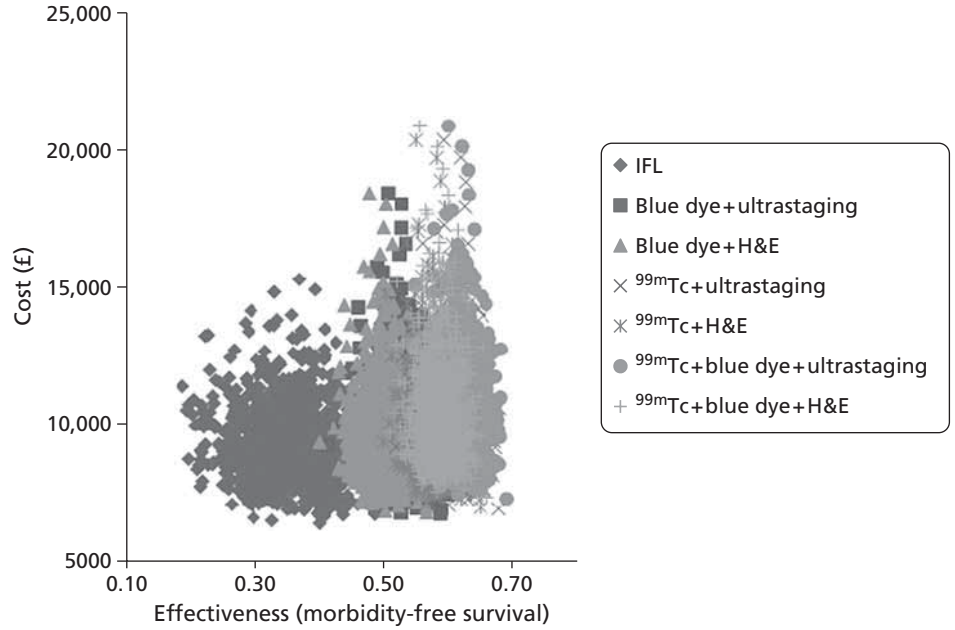


FIGURE 17 Scatterplot showing the uncertainty in costs and effectiveness within the model for each of the seven strategies for 1000 runs with morbidity-free survival at 2 years as the outcome measure.

Using the results shown in *Figure 17*, which illustrate the overall uncertainty in the model results with the outcome measure being morbidity-free survival at 2 years, *Figure 18* shows the CEAF for the outcome of additional case of morbidity averted. It can be seen that as the WTP crosses the ICER between two non-dominated strategies, the choice of optimal strategy changes, with a discontinuity in the curve being seen. Up to a WTP of £3500 the IFL strategy is the most cost-effective and, then, from £3500 to approximately £42,000 the ^{99m}Tc + ultrastaging strategy is the most cost-effective and, finally, for a WTP greater than approximately £42,000 the ^{99m}Tc + blue dye + ultrastaging strategy becomes the most cost-effective.

Probabilistic sensitivity analysis: long-term morbidity-free survival outcome

Using the results shown in *Figure 19*, which illustrates the overall uncertainty in the model results with the outcome measure being long-term morbidity-free survival at 2 years, *Figure 20* shows the CEAF for the outcome of an additional case of long-term morbidity averted. It can be seen that for a WTP of less than approximately £5000 the IFL strategy is the most cost-effective. For a WTP in the range from £5000 to £77,500 the ^{99m}Tc + ultrastaging strategy is the most cost-effective and, then, for WTP values greater than £77,500 the ^{99m}Tc + blue dye + ultrastaging strategy is the most cost-effective.

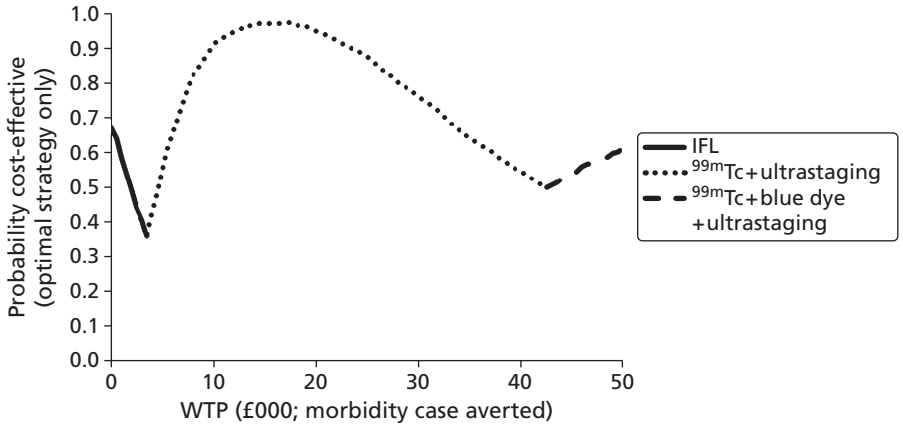


FIGURE 18 Cost-effectiveness acceptability frontier showing the results of the sensitivity analysis examining the optimal treatment strategy across a range of WTP thresholds for the outcome of additional case of morbidity averted at 2 years.

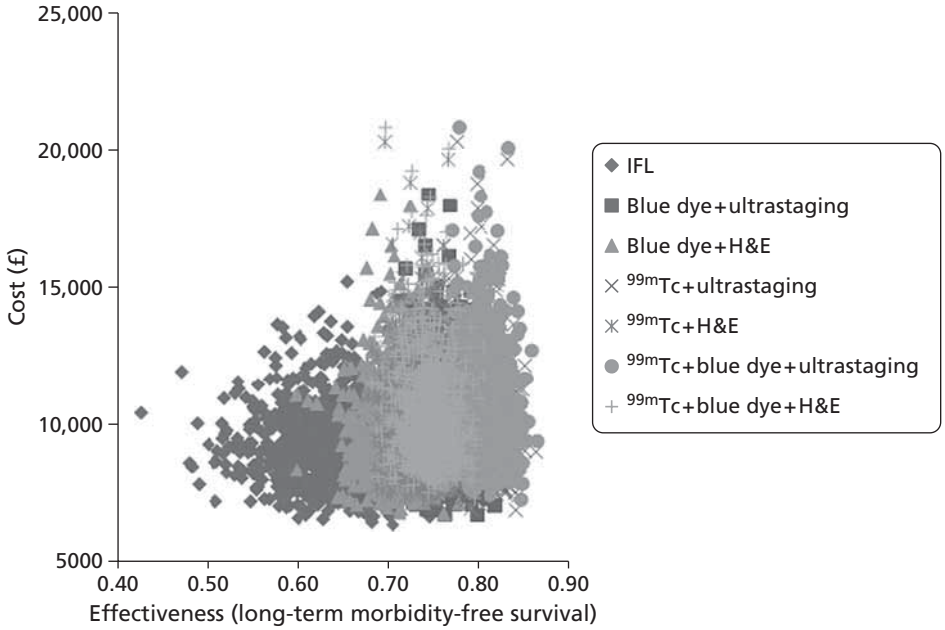


FIGURE 19 Scatterplot showing the uncertainty in costs and effectiveness within the model for each of the seven strategies for 1000 runs with long-term morbidity-free survival at 2 years as the outcome measure.

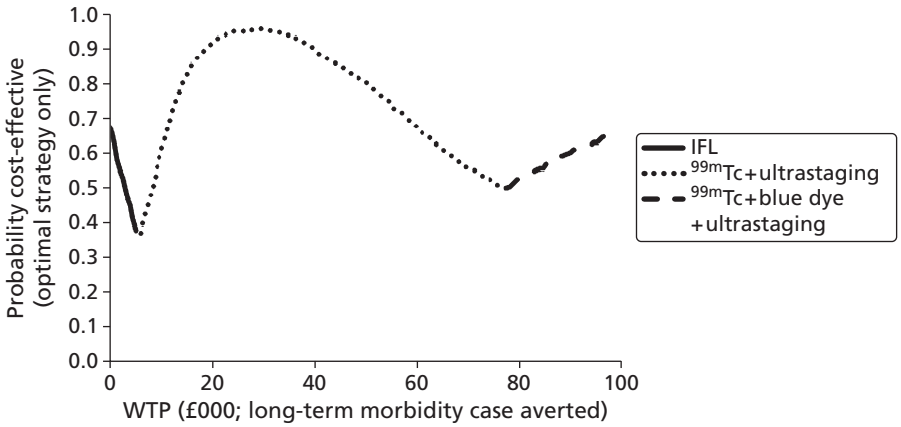


FIGURE 20 Cost-effectiveness acceptability frontier showing the results of the sensitivity analysis examining the optimal investigative strategy across a range of WTP thresholds for the outcome of additional case of long-term morbidity averted at 2 years.

Incremental analysis

Bilateral comparisons were carried out for the two morbidity outcomes, as these were thought to be the most informative, with the outcome for overall survival being excluded from this incremental analysis.

1. Inguinofemoral lymphadenectomy versus blue dye + ultrastaging

Figure 21 shows the modelled uncertainty in the differences in the costs and effectiveness between IFL and blue dye + ultrastaging for morbidity-free survival at 2 years. It shows that blue dye + ultrastaging may or may not increase the cost but will also certainly increase the effectiveness in terms of morbidity-free survival. *Figure 22* shows the proportion of model replications for which blue dye + ultrastaging is preferred to IFL. Blue dye+ ultrastaging is the preferred option at any WTP over £2000 although there is considerable uncertainty at any WTP around this figure. However, by the time the WTP exceeds £12,000, it is almost certain that blue dye + ultrastaging is preferred to IFL.

Figure 23 shows the modelled uncertainty in the differences in the costs and effectiveness between IFL and blue dye + ultrastaging for the long-term morbidity-free survival at 2 years outcome measure.

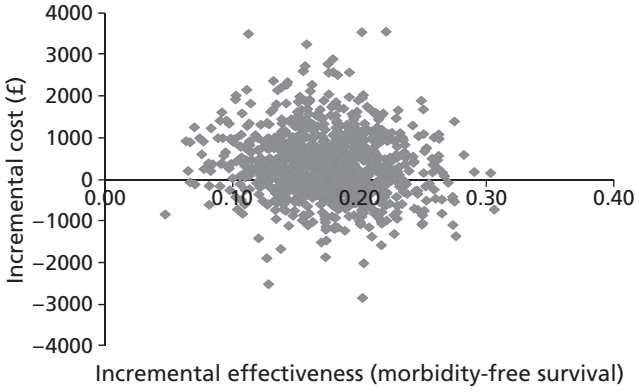


FIGURE 21 Cost-effectiveness plane: IFL vs. blue dye + ultrastaging for the morbidity-free survival at 2 years outcome measure.

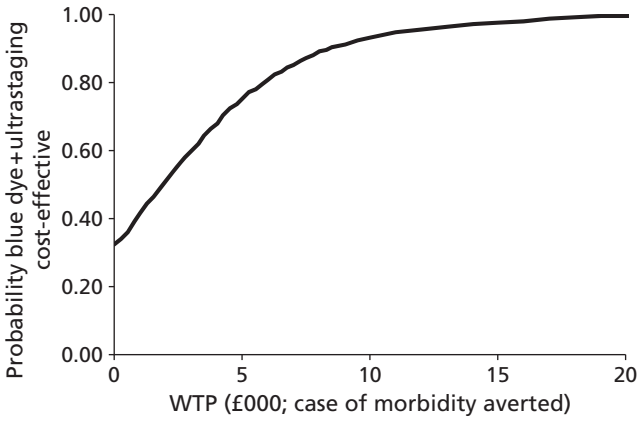


FIGURE 22 Cost-effectiveness acceptability curve: IFL vs. blue dye + ultrastaging for the case of morbidity averted at 2 years outcome measure.

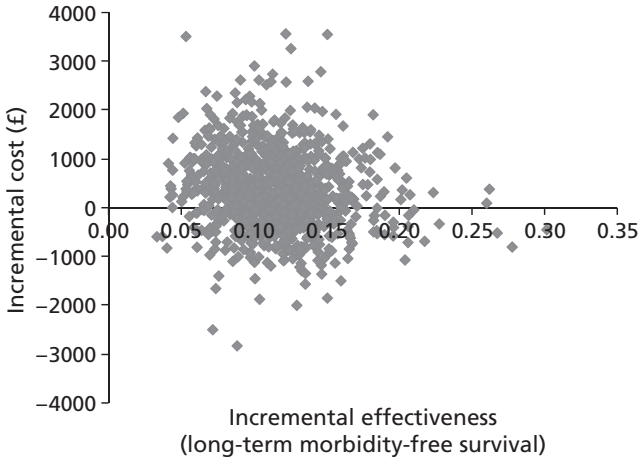


FIGURE 23 Cost-effectiveness plane: IFL vs. blue dye + ultrastaging for the long-term morbidity-free survival at 2 years outcome measure.

It shows that blue dye + ultrastaging may or may not increase the cost but will also certainly increase the effectiveness in terms of long-term morbidity-free survival.

Figure 24 shows the proportion of model replications for which blue dye + ultrastaging is preferred to IFL for the case of long-term morbidity averted outcome measure. Blue dye + ultrastaging is the preferred option at any WTP over £3000, although there is considerable uncertainty around this figure. However, by the time the WTP exceeds £18,000, it is almost certain that ^{99m}Tc + ultrastaging is preferred to IFL.

2. Blue dye + ultrastaging versus ^{99m}Tc + ultrastaging

Figure 25 shows the modelled uncertainty in the difference in costs between blue dye + ultrastaging and ^{99m}Tc + ultrastaging for the outcome measure of morbidity-free survival at 2 years. This shows that ^{99m}Tc + ultrastaging when compared with blue dye + ultrastaging will, on most occasions, increase the cost and will certainly increase the effectiveness in averting cases of morbidity. *Figure 26* shows the proportion of model replications for which ^{99m}Tc + ultrastaging is preferred to blue dye + ultrastaging at any given WTP per case of morbidity averted. It is more likely than not that ^{99m}Tc + ultrastaging is cost-effective compared with blue dye + ultrastaging above a WTP threshold of around £5000. At a WTP greater than £12,000 it is almost certain that ^{99m}Tc + ultrastaging will be preferred to blue dye + ultrastaging.

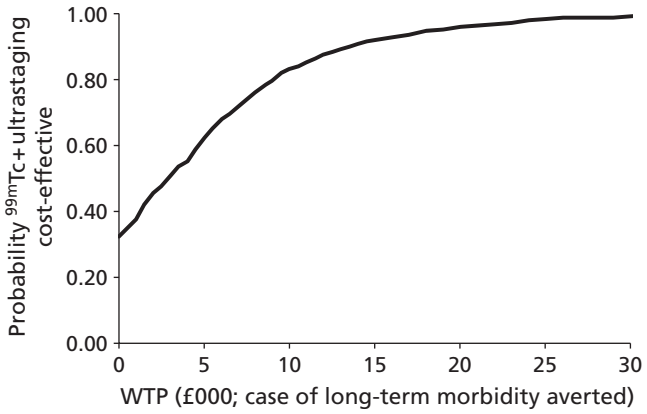


FIGURE 24 Cost-effectiveness acceptability curve: IFL vs. blue dye + ultrastaging for the case of long-term morbidity averted at 2 years outcome measure.

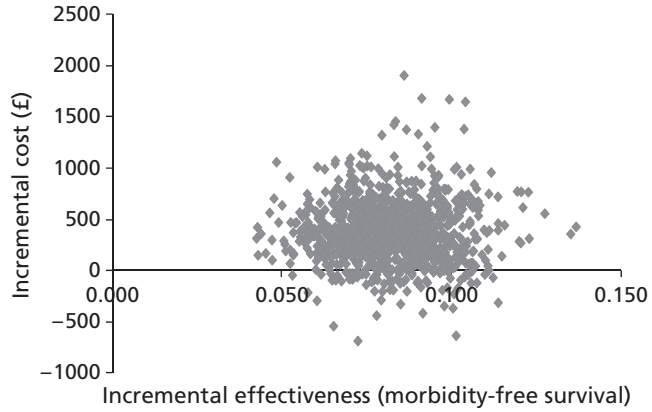


FIGURE 25 Cost-effectiveness plane: blue dye + ultrastaging vs. ^{99m}Tc + ultrastaging with morbidity-free survival at 2 years as the outcome measure.

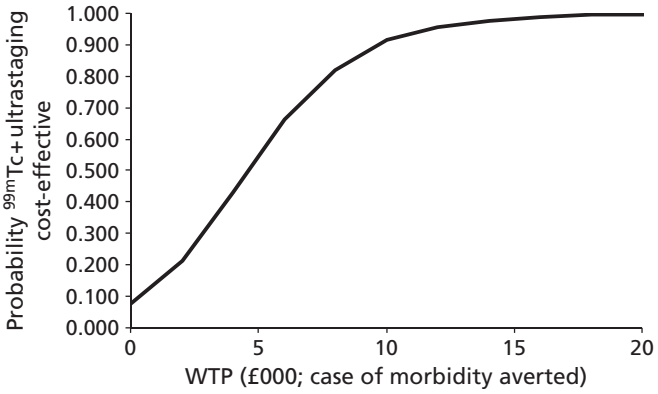


FIGURE 26 Cost-effectiveness acceptability curve: blue dye + ultrastaging vs. ^{99m}Tc + ultrastaging with case of morbidity averted at 2 years as the outcome measure.

Figure 27 shows the modelled uncertainty in the difference in costs between blue dye + ultrastaging and ^{99m}Tc + ultrastaging for the outcome measure of long-term morbidity-free survival at 2 years. It shows that the ^{99m}Tc + ultrastaging when compared with the blue dye + ultrastaging is more likely than not to increase the cost and will certainly increase the effectiveness in averting cases of long-term morbidity. *Figure 28* shows the proportion of model replications for which ^{99m}Tc + ultrastaging is preferred to blue dye + ultrastaging at any given WTP per case of long-term morbidity averted. It is more likely than not that ^{99m}Tc + ultrastaging is cost-effective compared with blue dye + ultrastaging above a WTP threshold of around £9000. At a WTP greater than £23,000 it is almost certain that ^{99m}Tc + ultrastaging will be preferred to blue dye + ultrastaging.

3. ^{99m}Tc + ultrastaging versus ^{99m}Tc + blue dye + ultrastaging

Figure 29 shows the modelled uncertainty in the differences in the costs and effectiveness between ^{99m}Tc + ultrastaging and ^{99m}Tc + blue dye + ultrastaging for morbidity-free survival at 2 years. It shows that ^{99m}Tc + blue dye + ultrastaging will always increase the cost and will almost always increase the effectiveness in terms of morbidity-free survival. *Figure 30* shows the proportion of model replications for which ^{99m}Tc + blue dye + ultrastaging is preferred to ^{99m}Tc + ultrastaging. ^{99m}Tc + blue dye+ ultrastaging is the preferred option at any WTP over £45,000, although there is considerable uncertainty at any WTP around this figure. Even when the WTP reaches £100,000, it still cannot be said that ^{99m}Tc + blue dye + ultrastaging is certainly preferred to ^{99m}Tc + ultrastaging.

Figure 31 shows the modelled uncertainty in the differences in the costs and effectiveness between ^{99m}Tc + ultrastaging and ^{99m}Tc + blue dye + ultrastaging for long-term morbidity-free survival. It shows that ^{99m}Tc + blue dye + ultrastaging will always increase the cost and will almost certainly increase the

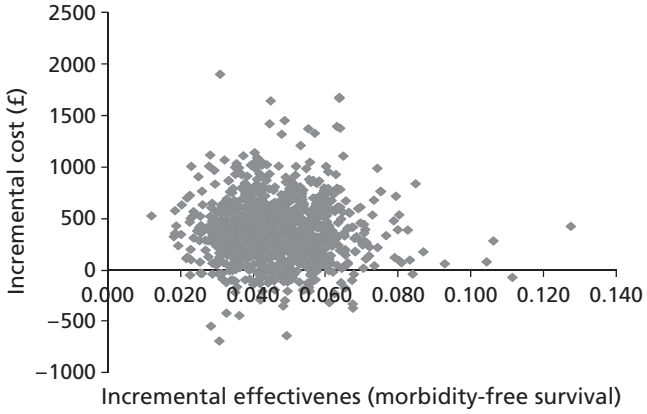


FIGURE 27 Cost-effectiveness plane: blue dye + ultrastaging vs. ^{99m}Tc + ultrastaging with long-term morbidity-free survival at 2 years as the outcome measure.

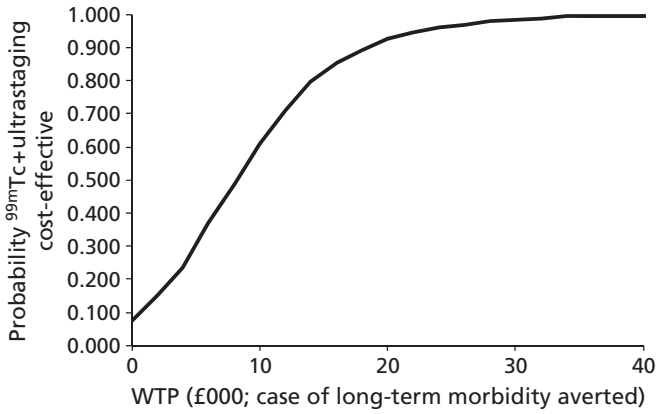


FIGURE 28 Cost-effectiveness acceptability curve: blue dye + ultrastaging vs. ^{99m}Tc + ultrastaging with case of long-term morbidity averted at 2 years as the outcome measure.

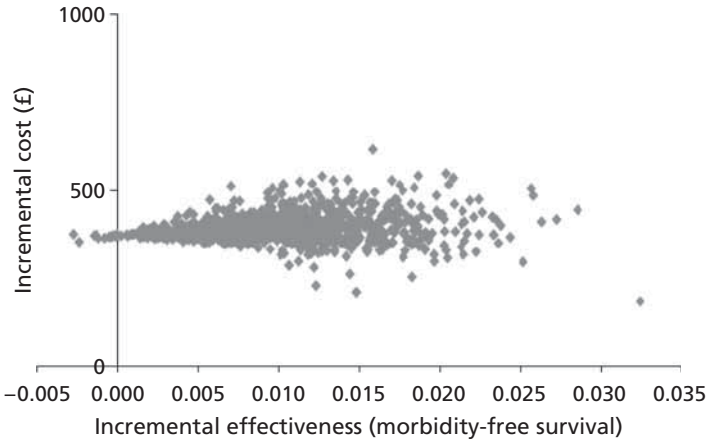


FIGURE 29 Cost-effectiveness plane: ^{99m}Tc + ultrastaging vs. ^{99m}Tc blue dye + ultrastaging with morbidity-free survival at 2 years as the outcome measure.

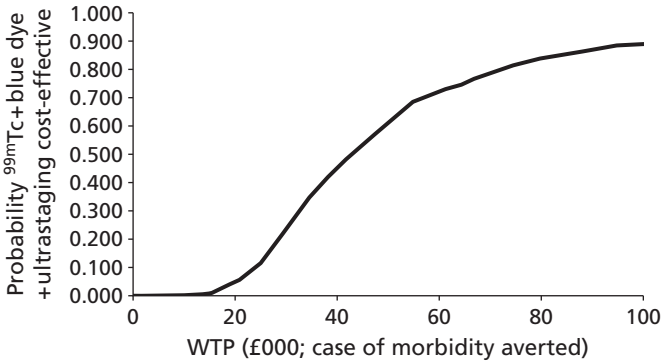


FIGURE 30 Cost-effectiveness acceptability curve: ^{99m}Tc + ultrastaging vs. ^{99m}Tc + blue dye + ultrastaging with case of morbidity averted at 2 years as the outcome measure.

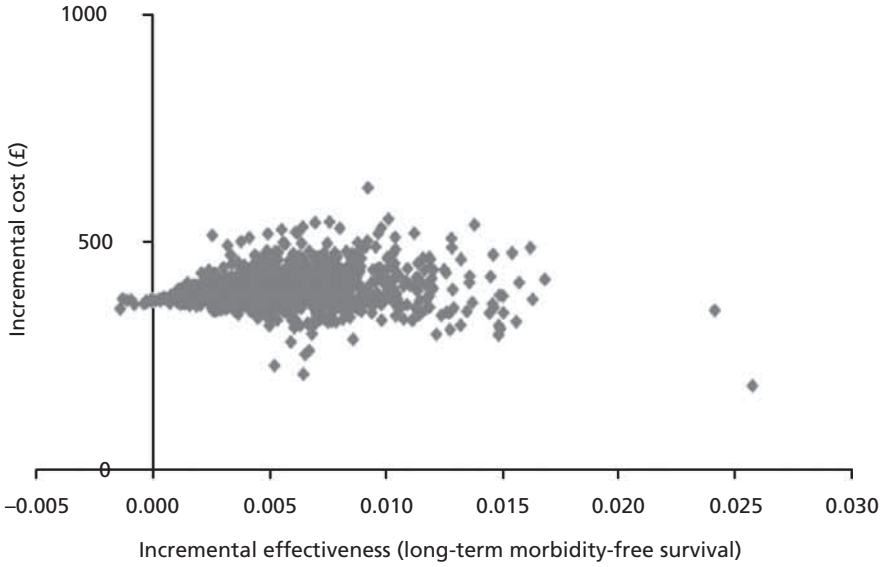


FIGURE 31 Cost-effectiveness plane: ^{99m}Tc + ultrastaging vs. ^{99m}Tc + blue dye + ultrastaging with long-term morbidity-free survival at 2 years as the outcome measure.

effectiveness in terms of long-term morbidity-free survival at 2 years. *Figure 32* shows the proportion of model replications for which ^{99m}Tc + blue dye + ultrastaging is preferred to ^{99m}Tc + ultrastaging in terms of long-term morbidity-free survival. ^{99m}Tc + blue dye+ ultrastaging is the preferred option at any WTP over £75,000, although there is considerable uncertainty at any WTP around this figure. Even when the WTP reaches £100,000, it still cannot be said that ^{99m}Tc + blue dye + ultrastaging is certainly preferred to ^{99m}Tc + ultrastaging.

4. Blue dye + H&E versus blue dye + ultrastaging

Figures 33 and *34* shows the modelled uncertainty in the difference in costs between the strategies of blue dye + H&E and blue dye + ultrastaging for the outcome measures of morbidity and long-term morbidity-free survival at 2 years. This shows that the blue dye + ultrastaging strategy when compared with blue dye + H&E will almost always reduce the costs and will always increase the effectiveness at averting cases of morbidity and long-term morbidity. This is reflected in *Figures 35* and *36*, which show that across all values of the WTP for a case of morbidity and long-term morbidity averted blue dye + ultrastaging is preferred to blue dye + H&E.

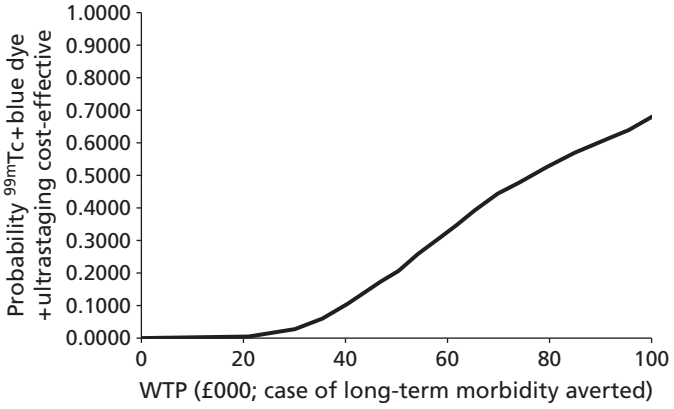


FIGURE 32 Cost-effectiveness acceptability curve: ^{99m}Tc + ultrastaging vs. ^{99m}Tc+ blue dye + ultrastaging with case of long-term morbidity averted at 2 years as the outcome measure.

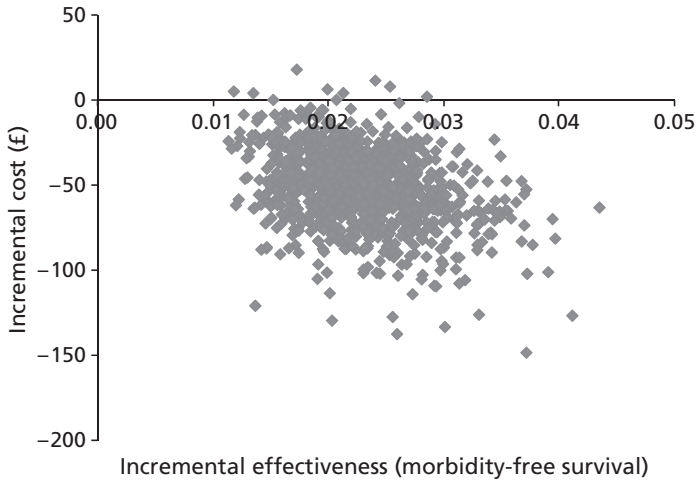


FIGURE 33 Cost-effectiveness plane: blue dye + H&E vs. blue dye + ultrastaging with morbidity-free survival at 2 years as the outcome measure.

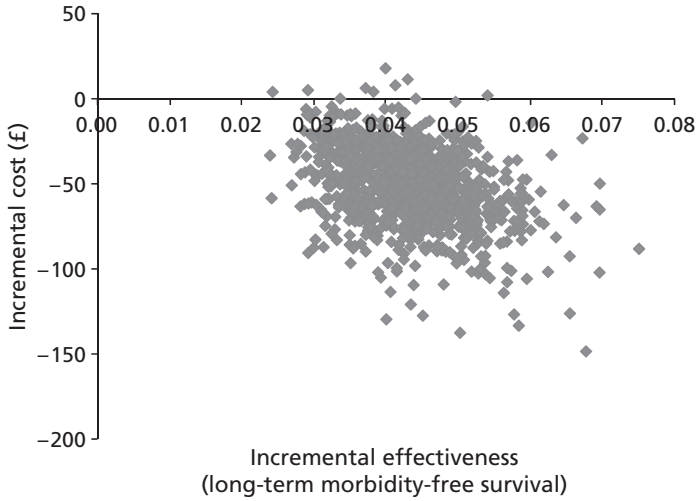


FIGURE 34 Cost-effectiveness plane: blue dye + H&E vs. blue dye + ultrastaging with long-term morbidity case averted at 2 years as the outcome measure.

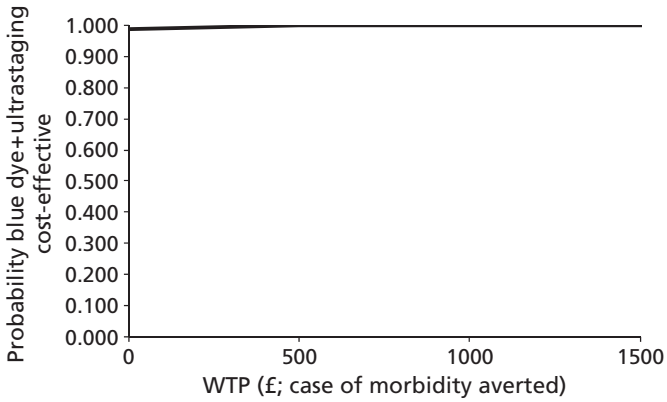


FIGURE 35 Cost-effectiveness acceptability curve: blue dye + H&E vs. blue dye + ultrastaging with case of morbidity averted at 2 years as the outcome measure.

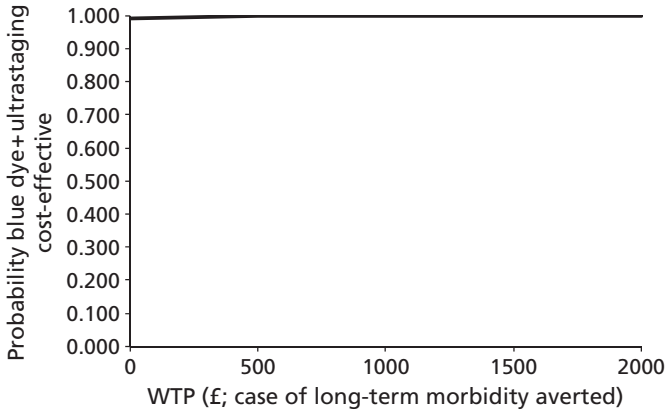


FIGURE 36 Cost-effectiveness acceptability curve: blue dye + H&E vs. blue dye + ultrastaging with case of long-term morbidity averted at 2 years as the outcome measure.

5. ^{99m}Tc + H&E versus ^{99m}Tc + ultrastaging

Figures 37 and *38* show the modelled uncertainty in the difference in costs between the strategies of ^{99m}Tc + H&E and ^{99m}Tc + ultrastaging for the outcome measures of morbidity and long-term morbidity-free survival at 2 years. This shows that the ^{99m}Tc + ultrastaging strategy when compared with ^{99m}Tc + H&E will almost always reduce the costs and will always increase the effectiveness at averting cases of morbidity and long-term morbidity. This is reflected in *Figures 39* and *40*, which show that across all values of the WTP for a case of morbidity and long-term morbidity averted, ^{99m}Tc + ultrastaging is preferred to ^{99m}Tc + H&E.

6. ^{99m}Tc + blue dye + H&E versus ^{99m}Tc + blue dye + ultrastaging

Figures 41 and *42* show the modelled uncertainty in the difference in costs between the strategies of ^{99m}Tc + blue dye + H&E and ^{99m}Tc + blue dye + ultrastaging for the outcome measures of morbidity and long-term morbidity-free survival at 2 years. This shows that the ^{99m}Tc + blue dye + ultrastaging strategy, when compared with ^{99m}Tc + blue dye + H&E, will almost always reduce the costs and will always increase the effectiveness at averting cases of morbidity and long-term morbidity. This is reflected in *Figures 43* and *44*, which show that across all values of the WTP for a case of morbidity and long-term morbidity averted, ^{99m}Tc + blue dye ultrastaging is preferred to ^{99m}Tc + blue dye + H&E.

One-way sensitivity analysis

In order to examine the impact of a selection of model assumptions on the model results, the following sensitivity analysis was undertaken.

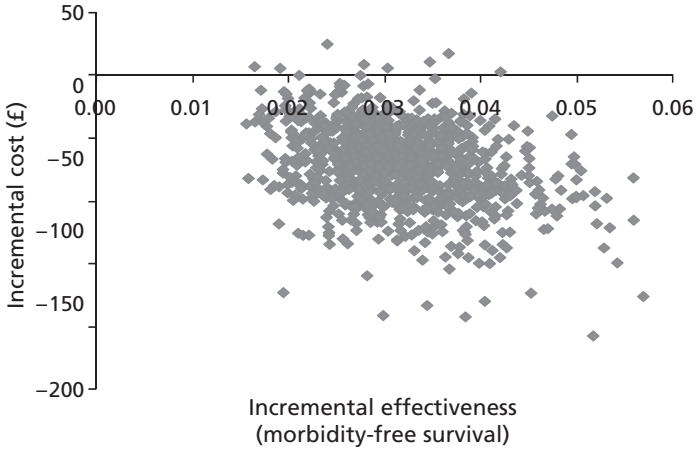


FIGURE 37 Cost-effectiveness plane: ^{99m}Tc + H&E vs. ^{99m}Tc + ultrastaging with morbidity-free survival at 2 years as the outcome measure.

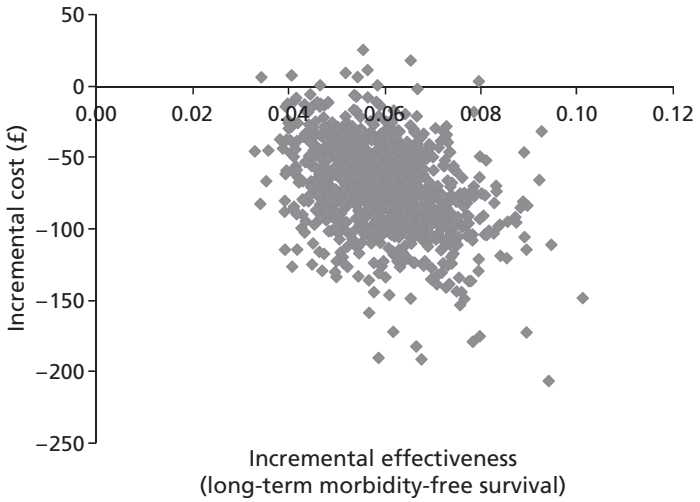


FIGURE 38 Cost-effectiveness plane: ^{99m}Tc + H&E vs. ^{99m}Tc + ultrastaging with long-term morbidity-free survival at 2 years as the outcome measure.

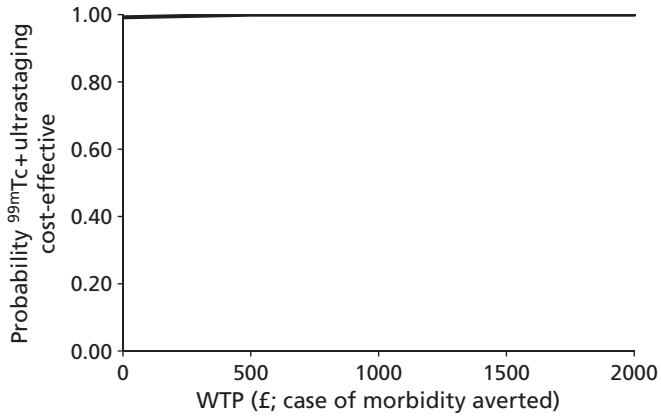


FIGURE 39 Cost-effectiveness acceptability curve: ^{99m}Tc + H&E vs. ^{99m}Tc + ultrastaging with case of morbidity averted at 2 years as the outcome measure.

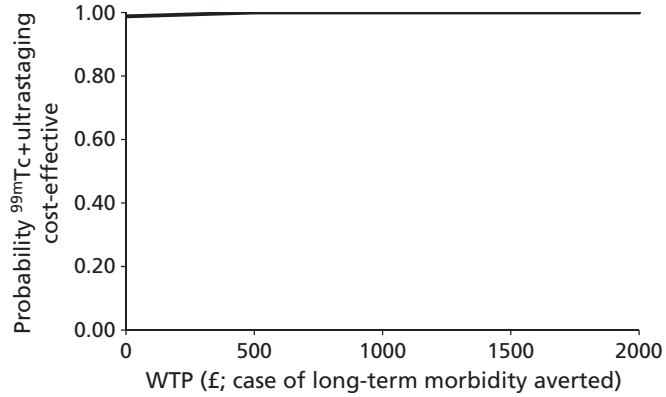


FIGURE 40 Cost-effectiveness acceptability curve: ^{99m}Tc + H&E vs. ^{99m}Tc + ultrastaging with case of long-term morbidity averted at 2 years as the outcome measure.

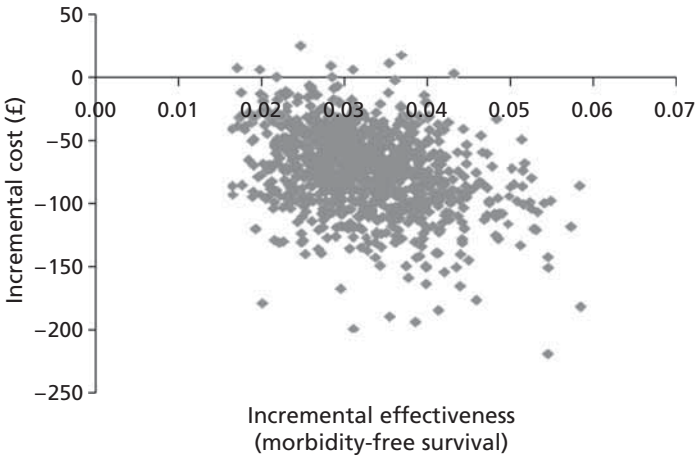


FIGURE 41 Cost-effectiveness plane: ^{99m}Tc + blue dye + H&E vs. ^{99m}Tc + blue dye + ultrastaging with morbidity-free survival at 2 years as the outcome measure.

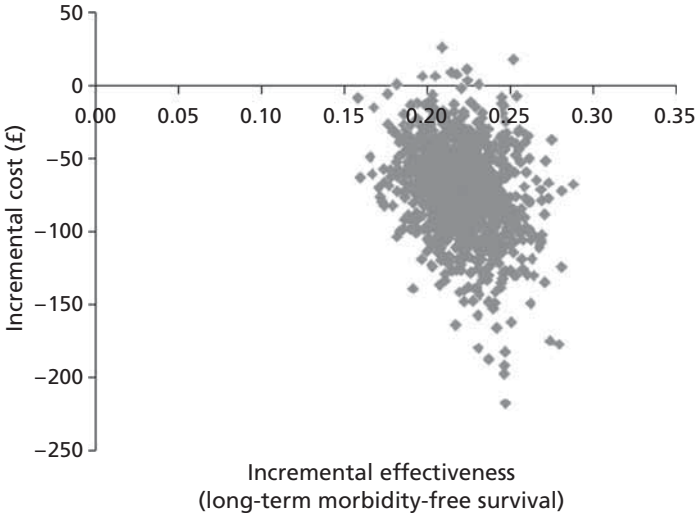


FIGURE 42 Cost-effectiveness plane: ^{99m}Tc + blue dye + H&E vs. ^{99m}Tc + blue dye + ultrastaging] with long-term morbidity-free survival at 2 years as the outcome measure.

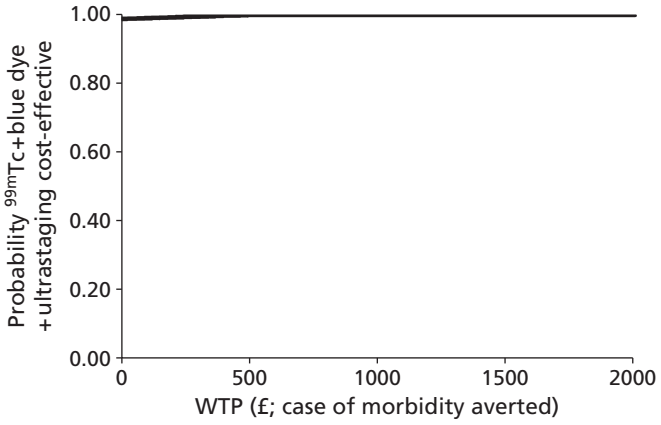


FIGURE 43 Cost-effectiveness acceptability curve: ^{99m}Tc + blue dye + H&E vs. ^{99m}Tc + blue dye + ultrastaging with case of morbidity averted at 2 years as the outcome measure.

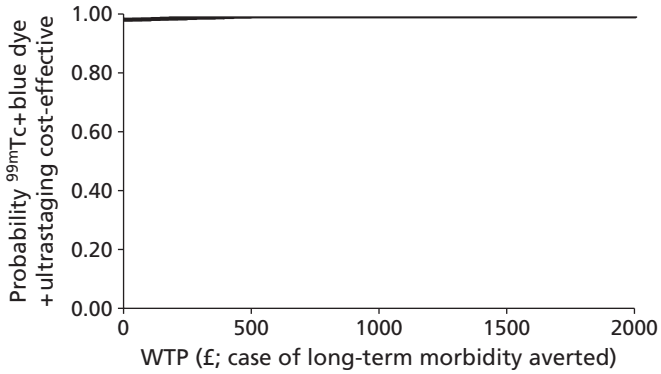


FIGURE 44 Cost-effectiveness acceptability curve: ^{99m}Tc + blue dye + H&E vs. ^{99m}Tc + blue dye + ultrastaging with case of long-term morbidity averted at 2 years as the outcome measure.

Assumed age of the cohort

The impact of the assumed age of the patient cohort on the costs and outcomes of each of the clinical pathways is shown in *Table 51* for ages 55 and 75 years, respectively. It can be seen that for the deterministic results in all cases, the assumed age of the cohort has no impact on the conclusions from the model. IFL continues to be the option that is the least costly and sees that highest survival among the patients. The strategies of blue dye + ultrastaging, ^{99m}Tc + ultrastaging and ^{99m}Tc + blue dye + ultrastaging show ICER values for morbidity and long-term morbidity that are very similar to those obtained at baseline.

TABLE 51 Results of one-way sensitivity analysis

Scenario	ICER – overall survival at 2 years	ICER – morbidity-free survival at 2 years (£)	ICER – long-term morbidity-free survival at 2 years (£)
Baseline			
IFL	Dominates		
Blue dye + ultrastaging		2400	3700
^{99m}Tc + ultrastaging		4900	8900
^{99m}Tc + blue dye + ultrastaging		41,200	74,300
Age = 55 years			
IFL	Dominates		
Blue dye + ultrastaging		2300	3600
^{99m}Tc + ultrastaging		4900	8800
^{99m}Tc + blue dye + ultrastaging		40,900	73,800
Age = 75 years			
IFL	Dominates		
Blue dye + ultrastaging		2400	4200
^{99m}Tc + ultrastaging		5100	9000
^{99m}Tc + blue dye + ultrastaging		42,600	76,000

continued

TABLE 51 Results of one-way sensitivity analysis (*continued*)

Scenario	ICER – overall survival at 2 years	ICER – morbidity-free survival at 2 years (£)	ICER – long-term morbidity-free survival at 2 years (£)
Increased mortality due to patient morbidity			
IFL	Dominates		
Blue dye + ultrastaging		2300	3700
^{99m}Tc + ultrastaging		4900	8700
^{99m}Tc + blue dye + ultrastaging		40,700	72,000
Cost of implementing ^{99m}Tc + blue dye together			
Cost of ^{99m}Tc + blue dye = cost ^{99m}Tc (£3836)			
IFL	Dominates		
Blue dye + ultrastaging		2400	£3700
^{99m}Tc + blue dye + ultrastaging		4700	8400
Cost of ^{99m}Tc + blue dye = 50% more than the cost of ^{99m}Tc (£5754)			
IFL	Dominates		
Blue dye + ultrastaging		2400	3700
^{99m}Tc + ultrastaging		4900	8900
^{99m}Tc + blue dye + ultrastaging		195,700	352,600
Groin recurrence rate following negative result for metastasis			
IFL	Dominates		
Blue dye + ultrastaging		2600	4300
^{99m}Tc + ultrastaging		5200	9800
^{99m}Tc + blue dye + ultrastaging		42,400	81,300

Increasing the death rate for morbidity

It has been assumed in this study at baseline that the morbidity experienced by the patients has no impact on their overall survival. It can be seen from the results obtained in *Table 51* that adding this factor to the model parameterisation has very little impact on the conclusions drawn from the model. IFL continues to be the option that sees that highest survival among the patients, while the strategies of blue dye + ultrastaging, ^{99m}Tc + ultrastaging and ^{99m}Tc + blue dye + ultrastaging show ICER values for morbidity and long-term morbidity that are very similar to those obtained at baseline.

The cost of implementing ^{99m}Tc and blue dye together

It was assumed in this study at baseline that the cost of implementing ^{99m}Tc and blue dye together is 10% more than the cost of implementing ^{99m}Tc alone. In this sensitivity analysis the impact of the assumption on the model results is examined by varying this cost across a number of plausible values. In the first case, it is assumed that implementing ^{99m}Tc and blue dye together costs the same as implementing ^{99m}Tc on its own (£3836) and in the second case, it is assumed that ^{99m}Tc and blue dye together cost 50% more than ^{99m}Tc alone (£5754).

It can be seen from the results obtained in *Table 51* in the case of overall survival that this change to the model parameterisation has very little impact on the conclusions drawn from the model. In terms of overall survival at 2 years, the IFL strategy continues to dominate all the other strategies. However, with respect to morbidity-free survival and long-term morbidity-free survival at 2 years, the cost of ^{99m}Tc and blue dye implemented together has the potential to have a significant impact on the conclusions drawn from the model. When the cost of ^{99m}Tc and blue dye implemented together is equal to the cost of ^{99m}Tc, this leads the strategy ^{99m}Tc + blue dye + ultrastaging to become far more acceptable in terms of its ICER value in comparison with the other strategies. However, when the cost of implementing ^{99m}Tc and blue dye together is increased to 50% greater, then the strategy of ^{99m}Tc + blue dye + ultrastaging becomes far less acceptable with an ICER value for morbidity and long-term morbidity survival of £195,700 and £352,600, respectively.

Groin recurrence rate following negative result for metastasis

It was assumed at baseline that the groin recurrence rate following a negative result from a SLN biopsy and histopathology is 6 out of 259 (2.3%).[73] To investigate the impact of this assumption on the model results, an alternative higher value of 2 out of 32 (6.5%), taken from a much smaller study,[65] was used.

It can be seen from the results obtained in *Table 51* that this change to the model parameterisation has very little impact on the conclusions drawn from the model. In terms of overall survival, the IFL strategy continues to dominate all the other strategies. In terms of morbidity-free survival and long-term morbidity-free survival at 2 years, the strategy of ^{99m}Tc + blue dye + ultrastaging is, again, the most effective option.

Chapter 8 Discussion

Principal findings

Test accuracy systematic review

Twenty-six studies gave information on test accuracy of ^{99m}Tc and/or blue dye identification of SLN biopsy with reference standard of either IFL for all or IFL for SLN-positive nodes (containing metastases) and clinical follow-up for SLN-negative nodes (see *Table 7*). Most of the studies included SCCs only (when reported), but 10 included up to eight patients with other forms of vulval cancer including melanoma and adenocarcinoma.[39,53,57,58,61,63,66,71,72,74] Most of the studies were small, with fewer than 50 patients, but one study involved 127 patients[57] and one 403 patients.[73] The histopathological techniques used varied between studies and included frozen sections, ultrastaging, H&E staining and immunohistochemical techniques. Not all studies had the same histochemical techniques for the SLN as for the remaining lymph nodes. There were, in effect, three index tests (^{99m}Tc, blue dye, and ^{99m}Tc and blue dye) and five reference standard groups (H&E only or insufficient details to determine histopathological techniques used, frozen sections only, immunohistochemistry, ultrastaging and ultrastaging with immunohistochemistry). Therefore, calculating the sensitivity and specificity of finding metastases in a SLN biopsy compared with the reference standard was not straightforward and no meta-analysis of all 26 studies was appropriate. In addition, a SLN was not found in all patients, so information also needed to be collected on whether or not the patient had malignancy at IFL. Ultimately, each patient could be categorised as follows:

- SLN found.
    - Has malignancy in SLN.
    - Has no malignancy in SLN.
        - Has malignancy in other lymph nodes found at IFL.
        - Has no malignancy in other lymph nodes found at IFL.
- No SLN found.
    - Has malignancy in other lymph nodes found at IFL.
    - Has no malignancy in other lymph nodes found at IFL.
    - Has no IFL.
        - Has groin metastases at follow-up.
        - Has no groin metastases at follow-up.

Teasing out the results for all of these different categories of patients was extremely difficult, particularly when the studies were small and poorly reported. In some studies, it was obvious that these different possibilities were not clearly thought through when either designing or reporting the study; therefore, some of the categories were combined, for example patients with negative SLN and malignancy in other nodes, and patients with no SLN found and malignancy in other nodes.

If malignancy is found in a SLN, then the patient will have groin metastases, so no FPs would be possible and all point estimates of specificity were 100%. CIs were wide, usually because of small sample sizes. However, FN results maybe more important clinically because of the risk of groin metastases to the patient. Sensitivities approached 100% in most categories of index tests and reference standards. In the largest group of studies (11 studies), using ^{99m}Tc and blue dye, ultrastaging and using immunohistochemistry, the pooled sensitivity was 95.6% (95% CI 91.5% to 98.1%) with some heterogeneity and an estimate of the

negative predictive value was 97.8% (see *Tables 16* and *17*). It is unclear whether or not the number of patients with metastases missed by biopsying the SLN is offset by the morbidity caused by IFL such as wound breakdown, cellulitis and lymphoedema. Where clinical follow-up is used, there will be more FNs at the time of the original operation, because of the longer time to develop a metastasis, so the sensitivity will be lower.

The SLN detection rate varied between studies: 76–100% for ^{99m}Tc only, 53–88% for blue dye only and 84–100% for both ^{99m}Tc and blue dye (see *Table 24*). The results from the 95% CIs of the combined rates suggest that, if SLN biopsy is going to be used, it is important that both tests are performed in every patient because not to do so risks missing the SLN in some patients. Protocols in more recent studies such as Levenback *et al.*[106] have tended to reflect this.

Recurrences can occur at the vulva, the groin and in distant organs. IFL is used not only to detect malignancies in groin nodes but also to remove them. Therefore, it is reasonable to assume that patients undergoing IFL are less likely to have groin recurrences than patients undergoing SLN biopsy only or no groin surgery (but this will have no impact on vulval recurrences). If metastases were found in the SLN, it would not be clinically acceptable to not perform an IFL (unless the patient refuses this procedure), so it is very difficult to demonstrate whether or not IFL does reduce groin recurrence rates because of lack of a comparator. Analysis of studies with clinical follow-up showed that groin recurrences do occur in patients who have had IFL and in patients with SLN-negative biopsies. In general, there was a higher rate of groin recurrences in SLN-positive patients than SLN-negative patients, but the numbers in the studies were small.

Survival rates were available in nine studies[50,54,60,63,65,70,71,73,74] and showed that survival was worse in SLN-positive compared with SLN-negative patients. However, vulval cancer is largely a disease of older women so, of the deaths in the studies, approximately half were from vulval cancer and half were from other causes.

Quality of life was reported for a subset of patients[82] in the largest included study.[73] QoL was lower in IFL patients than in SLN biopsy patients, at least according to some subscales of the FACT-V questionnaire, including discomfort in groins/vulva/legs, contentment, oedema, complaints and use of stockings. AEs were generally more frequent and more severe for the IFL patients than for the SLN biopsy patients.

Clinical effectiveness systematic review

All included studies had to report results for patients in whom more than 75% were FIGO stages I and II. Included in the systematic review was one RCT, three case–control studies and 13 case series (see *Table 31* and Stehamn *et al.*[87]). The RCT compared groin RT with bilateral IFL in patients who had undergone radical vulvectomy or modified radical vulvectomy and showed that IFL was associated with better survival and fewer recurrences than groin RT. The case–control studies and case series evaluated a variety of treatments, including surgery to the vulva (radical vulvectomy, modified radical vulvectomy, hemivulvectomy or wide local excision) alone or in combination with unilateral or bilateral IFL, RT or all three treatments. Two case series also evaluated RT only in some patients.[91,99] The smallest study enrolled 47 patients[91] and the largest enrolled 6965 patients;[99] patients from Kumar *et al.*[99] were from the SEER database. Follow-up of patients varied between 2 months[92] and 30 years,[94] but some studies did not report median follow-up times. The majority of patients had SCC, but three[88,89,98] also had a small proportion with other tumour types such as melanoma and adenocarcinoma. The mean or median ages of patients were all over 60 years, but some younger patients were included in each of the case series as lower age ranges included 20 years,[91] 23 years,[94] 27 years,[100] etc. In many of the studies, the rates of recurrence and survival in the different treatment groups were combined so that it was difficult to gain many insights into the effectiveness of different types of treatments.

The general trend of results was that there was more deaths in node-positive patients (but not for the cases in Stehman *et al.*[89]) and in patients treated with RT. There were fewer deaths in younger patients. The approximate relative rates of recurrence in vulva, groin and distant were around 4 : 2 : 1, respectively. In addition, the numbers dying from vulval cancer and other causes were approximately equal. No studies reported QoL. It was very difficult to pick out any trends in AEs owing to the variability in methods and categories of reporting, as most compared different types of surgery to the vulva (radical vulvectomy, modified radical vulvectomy, hemivulvectomy) rather than comparing IFL and no IFL, for example.

Economic evaluation

No studies were identified that have previously considered the cost-effectiveness of the available technologies for the treatment of vulval cancer among presumed FIGO stage I and stage II patients and, therefore, appropriate comparisons with other existing studies were not possible.

The results of the base-case deterministic analyses based on the outcome of cost per death averted at 2 years showed that, for patients with presumed stage I and stage II vulval cancer, the treatment strategy of IFL is both less costly and more effective than any of the strategies that used SLN biopsy.

When considering the outcome measures of morbidity-free survival and long-term morbidity-free survival at 2 years, it was found that the strategy of ^{99m}Tc + ultrastaging, for which ultrastaging is administered in the case of a negative H&E test, was likely to be cost-effective. Note that ultrastaging here is used as a proxy for more involved histopathological techniques such as immunohistochemistry. Moreover, it was noted that the strategies that included blue dye only as the approach to the SLN biopsy and H&E only for the histopathology were never found to be cost-effective and were always dominated by other strategies (other strategies being less costly and more effective). This finding emphasises that using blue dye and H&E for the identification of the SLN and the identification of metastasis, respectively, are not sensitive enough to be used on their own.

The ICER based on the outcome of morbidity-free survival at 2 years for the strategy of blue dye + ultrastaging compared with IFL was £2400 per case of morbidity-free survival, the ICER for ^{99m}Tc + ultrastaging compared with blue dye + ultrastaging was £4900 per case of morbidity-free survival and the ICER for ^{99m}Tc + blue dye + ultrastaging compared with ^{99m}Tc + ultrastaging was £41,200 per case of morbidity-free survival. Similarly, the outcome measure of long-term morbidity-free survival at 2 years was £3700 per case of long-term morbidity-free survival, the ICER for ^{99m}Tc + ultrastaging compared with blue dye + ultrastaging was £8900 per case of long-term morbidity-free survival and the ICER for ^{99m}Tc + blue dye + ultrastaging compared with ^{99m}Tc + ultrastaging was £74,300 per case of long-term morbidity-free survival.

The PSA suggests that at a WTP threshold of less than £3500 for a case of morbidity averted the IFL strategy is the most cost-effective. Then, from £3500 to £42,000, the ^{99m}Tc + ultrastaging strategy is the preferred option, given the current model, and then for a WTP of greater than £42,000 the ^{99m}Tc + blue dye + ultrastaging strategy becomes most cost-effective. In the case of the long-term morbidity averted at 2 years outcome measure, the ^{99m}Tc + ultrastaging strategy is the preferred option for a WTP of greater than £5000 compared with the IFL strategy. If the WTP for a case of long-term morbidity averted exceeds £77,500, then the ^{99m}Tc + blue dye + ultrastaging strategy becomes most cost-effective compared with the ^{99m}Tc + ultrastaging strategy. These findings provide further evidence that in terms of morbidity and long-term morbidity averted at 2 years, ^{99m}Tc + ultrastaging is likely to be the most cost-effective approach to the treatment of early-stage vulval cancer patients.

In order to examine some of the assumptions made, further one-way deterministic sensitivity analysis was undertaken. It was found that changing the assumed age of the patients in the model made no difference to the conclusions drawn from the model. This was also the case when examining the potential impact of increased mortality due to morbidity experienced by the patients. In addition, varying the groin recurrence rate after a negative SLN test across plausible values did not impact on the conclusions drawn from the

model. However, the cost of administering ^{99m}Tc and blue dye together has the potential to have a significant impact on the results obtained from this economic evaluation. During the parameterisation of this economic model, a cost for this specific procedure was not available and, so, at baseline it was assumed that this procedure would cost 10% more than the cost of ^{99m}Tc alone. When this cost was increased as part of sensitivity analysis, it was found that the ^{99m}Tc + blue dye + ultrastaging strategy was even less cost-effective with an ICER of £195,700 for the morbidity outcome compared with ^{99m}Tc + ultrastaging and £352,600 for the long-term morbidity outcome also compared with ^{99m}Tc + ultrastaging. The ICER values for the ^{99m}Tc + ultrastaging strategy remained unchanged. Although, conversely, when the cost of ^{99m}Tc and blue dye implemented together was set equal to ^{99m}Tc alone, then it was found that the strategy of ^{99m}Tc + blue dye + ultrastaging had an ICER of £4700 for morbidity averted and £8400 for long-term morbidity averted compared with blue dye + ultrastaging. However, it is probably not a reasonable assumption to assume that ^{99m}Tc and blue dye implemented together costs the same as when ^{99m}Tc is administered alone and, therefore, it can be argued that this provides further evidence of the robustness of the conclusion that ^{99m}Tc + ultrastaging is the most cost-effective strategy for the morbidity outcomes.

Strengths of the project

The strength of the test accuracy systematic review was the rigour of its conduct and the focus on comparing and contrasting the different versions of index test and reference standards. As much information as possible was extracted from each study including SLN detection rates, recurrence rates, survival, QoL and AEs in addition to test accuracy.

The strength of the clinical effectiveness systematic review included the rigorous efforts made to find as much relevant evidence as possible, demonstrated by the number of full papers examined and subsequently excluded (listed in *Appendix 11*). The focus was on FIGO stages I and II evidence to ensure compatibility with the test accuracy systematic review and also separation of the different treatment types.

The economic evaluation has had the advantage of being able to use the best available data in the model established in the systematic reviews of the evidence, particularly the sensitivity and specificity of the procedures used to identify a SLN and metastases. All assumptions used in the model were agreed by a panel of experts a priori, with key assumptions being examined through the use of sensitivity analysis. Owing to the scarcity of vulval cancer, many of the data points used in this modelling study were based on small samples. However, the resultant uncertainly in these parameter values was examined through the use of PSA and the conclusions were mainly robust.

Limitations of the project

In the test accuracy systematic review, some of the included studies had considerable methodological limitations, including lack an adequate description of inclusion criteria, population (especially stages of disease) and reference standard used. Other problems included the omission of results or failure to separate results for the different categories of patients. Areas for which information was often not reported included the blinding of test results and surgeons' experience. With histological examination, the thickness of sections taken was often not well reported so that some studies may have used much thicker sections than others. Thinner sections are more likely to find micrometastases. There was no information on the therapeutic impact of SLN biopsy.

The main limitation of the clinical effectiveness systematic review was the lack of good-quality information on the effectiveness of the different categories of treatment. Many of the case series were comparing different types of surgical techniques but tended not to compare surgery only with surgery plus IFL and, so, only combined results were available for many of the studies. This meant that the number of hospital

days per treatment type was also unavailable. In the original protocol for this work, it was estimated that eliciting subjective probabilities from clinicians would be useful in determining diagnostic and therapeutic impact of SLN biopsy compared with IFL. During the project, and from our work in a previous project on positron emission tomography, computed tomography (PET-CT) for recurrent cervical cancer,[107] it became obvious that the missing information was QoL from patients, and clinicians' estimates would probably not have been helpful.

Any systematic review must have a cut-off date for the searches, and it is unfortunate when a major study is published after this date. A large, good-quality diagnostic study of 452 patients comparing SLN biopsy with IFL has recently been published.[106] It used blue dye and ^{99m}Tc in most patients and ultrastaging with immunohistochemistry. The results from this paper are similar to those of the included studies. The sensitivity was 92.3% – very similar to that found by Klat *et al.*,[60] which had similar study characteristics.

The major limitation of the economic evaluation was the absence of overall QoL estimates that differentiate between the impact of the SLN biopsy and IFL on the morbidity of the patients. The only global QoL measure available was in a relatively small sample and showed no difference between SLN and IFL strategies.[82] Instead, three outcome measures were considered in this study – overall mortality, morbidity-free survival and long-term morbidity-free survival – all over a 2-year time horizon. The inevitable difficulty with this selection is that it is very difficult to know which should be the primary outcome measure. Intuitively, overall survival would seem to be best choice; however, in this setting, this outcome provided little insight into the clinical dilemma facing clinicians treating vulval cancer: is the extra morbidity from IFL worth it when compared with the slight increase in mortality and less morbidity from a SLN procedure? IFL, while the cheapest strategy considered in this analysis, is a highly morbid procedure that is most effective at reducing the probability of a groin recurrence in the future and hence patient mortality. The option of using SLN biopsy has been introduced into practice with the aim of reducing patient morbidity, but at the expense of the increased possibility of patients being diagnosed FN for metastasis. Therefore, it was inevitable that, in this study, IFL has been shown to be the most effective procedure in terms of overall survival. Perhaps the best outcome measure in this study is morbidity-free survival as this also incorporates all types of morbidity and the impact of overall survival into the outcome, although long-term morbidity-free survival has also been considered, since lymphoedema, which is a long-term complication, has been found to have some significant impacts on patients treated with an IFL compared with those receiving a SLN biopsy.[82]

Cost data for the SLN biopsies were particularly difficult to identify as the NHS reference costs do not differentiate between the types of biopsy implemented. All the costs used in this study were checked by experts in the field to ensure that reasonable values had been adopted.

Uncertainties

The GROINSS-V study[73] and a more recent diagnostic study of SLN biopsy in vulval cancer[106] both discuss pathological characteristics of women in whom they consider SLN biopsy would be appropriate (in particular primary tumour size < 4 cm) and biopsy technique (SLN detected using ^{99m}Tc and blue dye, malignancy checked for with ultrastaging and immunohistochemistry, use of a high-volume centre of expertise) but it is uncertain whether or not patients would rather risk unremoved groin metastases by forgoing IFL if they are SLN negative at biopsy. The one published survey available shows that 100% of gynaecologists would accept a 1% FN rate from SLN biopsy, but only 33% of patients would accept a SLN biopsy in which the risk was 1 in 80.[41] It is clear that some metastases will be missed, but it is unclear whether or not the relative benefit of not having the morbidity of IFL is outweighed by the risk of missing metastases. This may depend on a number of factors, including the age of the patient, the stage of disease and the aggressiveness of the malignancy found in the vulval specimen.

The most effective treatment strategy for early vulval cancer is currently unclear, as is the relative effectiveness of vulval surgery only compared with vulval surgery with IFL. Much work has focused on the relative benefits of different types of surgical procedure, rather than the impact on mortality of IFL versus no IFL. From the RCT comparing RT with IFL, it is likely that IFL enhances survival but the sample sizes were small and the results were only just statistically significant so replication of this study would be useful to reduce some uncertainty.

The most important uncertainty in this project is the cost–utility of SLN biopsy compared with IFL, but this could not be calculated owing to lack of information on generic QoL.

Chapter 9 Overall conclusions

Recommendation for practice/implications for service provision

There is insufficient evidence to suggest that SLN biopsy should be used in routine clinical practice on health economic grounds. The strategy of IFL for all was found to be less costly and more effective when considering cost per death averted. Based on the findings of the current model and acknowledging the limitations that have been highlighted in terms of the inability to apply QALYs in this economic evaluation, the results of this analysis suggest that ^{99m}Tc + ultrastaging in the treatment of early-stage vulval cancer is likely to be cost-effective in terms of case of morbidity averted and long-term morbidity averted at 2 years. Note that ultrastaging has been used here as a proxy for more in-depth histopathological techniques such as immunohistochemistry. There is some uncertainty regarding the acceptability of the ^{99m}Tc + blue dye + ultrastaging strategy in terms of the outcome measures of case of morbidity and long-term morbidity averted at 2 years, as there is difficulty in attempting to apply these outcome measures to any acceptability threshold.

Recommendation for research

There needs to be further evaluation of patient preferences regarding the circumstance in which patients would rather risk unremoved groin metastases by forgoing IFL if they are SLN-negative at biopsy, incorporating a number of factors including the age of the patient, the stage of disease, persistency of side effects and the aggressiveness of the malignancy found in the vulval specimen.

There needs to be a robust prospective evaluation of the relative effectiveness of the different treatment strategies for vulval cancer, taking into account the uncertainty around the need for IFL in early-stage vulval cancer. As vulval cancer is uncommon, a multicentre RCT involving several countries will likely be needed to enrol sufficient patients in order to deal with the uncertainty.

There needs to be some information on the QoL in vulval cancer, using a generic QoL measure such as EQ-5D. This analysis has highlighted the importance of obtaining overall QoL values that describe the impact of the SLN biopsy and IFL and their related complications on patients over time. A previous study has attempted to identify these values but did not find a difference in the QoL estimates between 62 patients who received either a SLN biopsy or an IFL.[82] Intuitively, there would need to be a difference in QoL between these two groups because, if this were not the case, IFL – with its increased effectiveness at reducing the risk of a further groin recurrence and, therefore, patient mortality, but with its much higher risk of morbidity – would always be preferred. Therefore, future work should be undertaken to examine the QoL in these treatment groups perhaps by using an alternative type of questionnaire and through a larger study that includes more patients so would have better power to determine a small difference.

Good-quality, larger studies are required to more accurately estimate the test accuracy of SLN biopsy with histopathology. These studies need to distinguish carefully between finding the SLN and finding malignancy within the node.

Acknowledgements

Anne Fry-Smith for devising original searches, Shakila Thangaratinam for writing original bid and protocol, Marian Ellis for providing patient input to the protocol, Tara Selman for the original PhD work leading up to the development of this research and Kottekkattu Balan for providing clinical input to the protocol.

Contributions of authors

Catherine Meads supervised and co-ordinated the project, rewrote and edited the entire manuscript.

Andrew Sutton conducted the economic evaluation, wrote the economic evaluation chapter and the parts of the discussion chapter related to the economic evaluation.

Sylwia Małysiak conducted the original systematic review of test accuracy.

Monika Kowalska conducted the original systematic review of effectiveness.

Anna Zapalska conducted the original systematic review of test accuracy.

Ewelina Rogozińska wrote the original background section.

Peter Baldwin provided clinical advice for the project.

Adam Rosenthal provided clinical advice for the project.

Raji Ganesan provided pathology advice for the project and wrote the section on VIN.

Ewa Borowiack managed the original systematic reviews.

Pelham Barton provided economic evaluation supervision and edited the economic chapter.

Tracy Roberts managed the economic evaluation and edited the economic chapter.

Sudha Sundar provided clinical input for the project and the final report.

Khalid Khan provided the design of the HTA and overall management of project.

References

1. Anon. *Management of vulval cancer*. London, UK: Royal College of Obstetricians and Gynaecologists; 2006.
2. Tyring SK. Vulvar squamous cell carcinoma: guidelines for early diagnosis and treatment. *Am J Obstet Gynecol* 2003;**189**(Suppl. 3):S17–23. http://dx.doi.org/10.1067/S0002-9378(03)00792-0
3. de Hullu JA, van der Avoort IA, Oonk MH, van der Zee AG. Management of vulvar cancers. *Eur J Surg Oncol* 2006;**32**:825–31. http://dx.doi.org/10.1016/j.ejso.2006.03.035
4. Gottleib WH. The assessment and surgical management of early-stage vulvar cancer. *Best Pract Res Clin Obstet Gynaecol* 2003;**17**:557–69. http://dx.doi.org/10.1016/S1521-6934(03)00066-X
5. Magrina JF, Gonzalez-Bosquet J, Weaver AL, Gaffey TA, Leslie KO, Webb MJ, *et al.* Squamous cell carcinoma of the vulva stage IA: long-term results. *Gynecol Oncol* 2000;**76**:24–7. http://dx.doi.org/10.1006/gyno.1999.5638
6. Reynolds RK, Loar PV. Gynecology. In: Doherty G, editor. *Current Diagnosis and Treatment: Surgery*. New York, NY: McGraw-Hill; 2010. pp. 966–84.
7. van Doorn HC, Ansink A, Verhaar-Langereis MM, Stalpers LL. Neoadjuvant chemoradiation for advanced primary vulvar cancer. *Cochrane Database Syst Rev* 2009;**4**:1–20.
8. Oonk MH, de Hullu JA, van der Zee AG. current controversies in the management of patients with early-stage vulvar cancer. *Curr Opin Oncol* 2010;**22**:481–6. http://dx.doi.org/10.1097/CCO.0b013e32833c06da
9. Hart WR. Vulvar intraepithelial neoplasia: Historical aspects and current status. *Int J Gynecol Pathol* 2001;**20**:16–30. http://dx.doi.org/10.1097/00004347-200101000-00003
10. Sideri M, Wilkinson EJ. Squamous vulvar intraepithelial neoplasia: 2004 modified terminology, ISSVD Vulvar Oncology Subcommittee. *J Reprod Med* 2005;**25**:807–10.
11. Yang B, Hart WR. Vulvar intraepithelial neoplasia of the simplex (differentiated) type: a clinicopathologic study including analysis of HPV and p53 expression. *Am J Surg Pathol* 2000;**24**:429–41. http://dx.doi.org/10.1097/00000478-200003000-00013
12. de Bie RP, van de Niewenhof HP, Bekkers RL, Melchers WJ, Siebers AG, Bulten J, *et al.* Patients with usual vulvar intraepithelial neoplasia-related vulvar cancer have an increased risk of cervical abnormalities. *Br J Cancer* 2009;**101**:27–31. http://dx.doi.org/10.1038/sj.bjc.6605124
13. Eva LJ, Ganesan R, Chan KK, Honest LDM. Differentiated-type vulval intraepithelial neoplasia has a high-risk association with vulval squamous cell carcinoma. *Int J Gynecol Cancer* 2009;**19**:741–4. http://dx.doi.org/10.1111/IGC.0b013e3181a12fa2
14. Kokka F, Singh N, Faruqi A, Gibbon K, Rosenthal AN. Is differentiated vulval intraepithelial neoplasia the precursor lesion of human papillomavirus-negative vulval squamous cell carcinoma? *Int J Gynecol Cancer* 2011;**21**:1297–305. http://dx.doi.org/10.1097/IGC.0b013e31822dbe26
15. Daling JR, Sherman K, Hislop TG, Maden C, Mandelson MT, Beckmann AM. Cigarette smoking and the risk of anogenital cancer. *Am J Epidemiol* 1992;**135**:180–9.
16. Madeleine MM, Daling JR, Schwarts SM. Cofactors with human papillomavirus in a population-based study of vulval cancer. *J Natl Cancer Inst* 1997;**89**:1516–23. http://dx.doi.org/10.1093/jnci/89.20.1516
17. Brinton LA, Nasca PC, Mallin K, Baptiste MS, Wilbanks GD, Richart RM. Case–control study of cancer of the vulva. *Obstet Gynecol* 1990;**75**:859–66.

18. Hording U, Krigsholme B, Andreasson B, Visfeldt J, Daugaard S, Bock JE. Human papillomavirus in vulval squamous-cell carcinoma and in normal vulval tissues: a search for a possible impact of HPV on vulval cancer prognosis . *Int J Cancer* 1993;**55**:396. http://dx.doi.org/10.1002/ijc.2910550310

19. MacNab JC, Walkinshaw SA, Cordiner JW, Clements JB. Human papillomavirus in clinically and histologically normal tissue of patients with genital cancer. *N Engl J Med* 1986;**315**:1052–8. http://dx.doi.org/10.1056/NEJM198610233151703

20. Frisch M, Biggar RJ, Goedert JJ. Human papillomavirus-associated cancers in patients with human immunodeficiency virus infection and acquired immunodeficiency syndrome. *J Natl Cancer Inst* 2000;**92**:1500–10. http://dx.doi.org/10.1093/jnci/92.18.1500

21. Cancer Research UK. *Vulval Cancer Incidence Statistics* . URL: http://info.cancerresearchuk.org/cancerstats/types/vulva/incidence/uk-vulva-cancer-incidence-statistics (accessed 1 April 2012).

22. Mabuchi K, Bross DS, Kessler II. Epidemiology of cancer of the vulva. A case-control study. *Cancer* 1985;**55**:1843–8. http://dx.doi.org/10.1002/1097-0142(19850415)55:8<1843::AID-CNCR2820550833>3.0.CO;2-M

23. Boffetta P, Grindley G, Lindelof B. Cancer risk in a population-based cohort of patients hospitalized for psoriasis in Sweden. *J Invest Dermatol* 2001;**117**:1531–7.

24. Penn I. Cancers of the anogenital region in renal transplant recipients. Analysis of 65 cases. *Cancer* 1986;**58**:611–16. http://dx.doi.org/10.1002/1097-0142(19860801)58:3<611::AID-CNCR2820580303>3.0.CO;2-M

25. Hacker NF. Vulvar cancer. In: Berek JS, Hacker NF, editors. *Practical Gynaecologic Oncology*. Philadelphia, PA: Williams & Wilkins; 2005. pp. 585–602.

26. Anon. *Handbook of Gynaecologic Oncology for Specialists and Trainees*. Kuala Lumpur: Globalcrest Sdn Bhd; 2011.

27. Cancer Research UK. *Vulval Cancer Mortality Statistics*. URL: http://info.cancerresearchuk.org/cancerstats/types/vulva/mortality/ (accessed 1st April 2012)

28. National Cancer Institute. *Vulvar cancer treatment (PDQ)*. URL: www.cancer.gov/cancertopics/pdq/treatment/vulvar/HealthProfessional/page5 (accessed 5 April 2011)

29. Homesley HD, Bundy BN, Sedlis A, Yordan E, Berek JS, Jahshan A, *et al.* Assessment of current International Federation of Gynecology and Obstetrics staging of vulvar carcinoma relative to prognostic factors for survival (a Gynecologic Oncology Group study). *Am J Obstet Gynecol* 1991;**164**:997–1003. http://dx.doi.org/10.1016/0002-9378(91)90573-A

30. Homesley H, Bundy B, Sedlis A. Prognostic factors for groin node metastasis in squamous cell carcinoma of the vulva. *Gynecol Oncol* 1993;**49**:279–83. http://dx.doi.org/10.1006/gyno.1993.1127

31. Lucas WE. Die stadienangepasste Behandlung des Vulvakarzinoms. *Gynakologe* 1981;**14**:150.

32. Weijmar Schultz WC, van der Wiel HB, Bouma J, Janssens J, Littlewood J. Psychosexual functioning after the treatment of cancer of the vulva. *Cancer* 1990;**66**:402–7.

33. Gould N, Kamelle S, Tillmanns T, Scribner D, Gold M, Walker J. Predictors of complications after inguinal lymphadenectomy. *Gynecol Oncol* 2001;**82**:329–32. http://dx.doi.org/10.1006/gyno.2001.6266

34. Gaarenstroom KN, Kenter GG, Trimbos JB, Agous I, Amant F, Peters AA, *et al.* Postoperative complications after vulvectomy and inguinofemoral lymphadenectomy using separate groin incisions. *Int J Gynecol Cancer* 2003;**13**:522–7. http://dx.doi.org/10.1046/j.1525-1438.2003.13304.x

35. Beesley V, Janda M, Eakin E, Obermair A, Battistutta D. Lymphedema after gynecological cancer treatment: prevalence, correlates, and supportive care needs. *Cancer* 2007;**109**:2607–14. http://dx.doi.org/10.1002/cncr.22684

36. Pereira de Godoy BN, Braile DM, de Fatima Godoy M, Longo J. Quality of life and peripheral lymphoedema. *Lymphology* 2002;**35**:72–5.

37. Helm CW, Hatch K, Austin JM, Partridge EE, Soong SJ, Elder JE, *et al.* A matched comparison of single and triple incision techniques for the surgical treatment of carcinoma of the vulva. *Gynecol Oncol* 1992;**46**:150–6. http://dx.doi.org/10.1016/0090-8258(92)90247-G

38. Sedlis A, Homesley H, Bundy B. Positive groin lymph nodes in superficial squamous vulvar cancer. *Am J Obstet Gynecol* 1987;**156**:1159–64. http://dx.doi.org/10.1016/0002-9378(87)90132-3

39. Lindell G, Jonsson C, Ehrsson RJ, Jacobsson H, Danielsson KG, Kallstrom BN, *et al.* Evaluation of preoperative lymphoscintigraphy and sentinel node procedure in vulvar cancer. *Eur J Obstet Gynecol Reprod Biol* 2010;**152**:91–5. http://dx.doi.org/10.1016/j.ejogrb.2010.05.011

40. Terada KY, Shimizu DM, Wong JH. Sentinel node dissection and ultrastaging in squamous cell Cancer of the vulva. *Gynecol Oncol* 2000;**76**:40–4.

41. de Hullu JA, Ansink AC, Tymstra T, van der Zee AG. What doctors and patients think about false-negative sentinel lymph nodes in vulvar cancer. *J Psychosom Obstet Gynecol* 2001;**22**:199–203. http://dx.doi.org/10.3109/01674820109049974

42. Anon. *Systematic reviews, CRD's Guidance for undertaking reviews in health care*. York: Centre for Reviews and Dissemination, University of York; 2009.

43. Higgins JT, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester: Wiley-Blackwell; 2008.

44. Diagnostic Test Accuracy Working Group. *Handbook for DTA reviews*; 2012. URL: http://srdta cochrane org/handbook-dta-reviews (accessed January 2012).

45. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol* 2009;**62**:1006–12. http://dx.doi.org/10.1016/j.jclinepi.2009.06.005

46. Whiting P, Rutjes AWS, 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* 2003;**3**:25.

47. 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. http://dx.doi.org/10.1186/1471-2288-9-3

48. Wells G, Shea B, O'Coonell D, Robertson J, Peterson J, Welch V, *et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analysis*. URL: www.ohri.ca/programs/clinical_epidemiology/oxford.asp (accessed 5 April 2012).

49. Dalziel K, Round A, Stein K, Garside R, Castelnuovo E, Payne L. Do findings of case series studies vary significantly according to methodological characteristics? *Health Technol Assess* 2005;**9**(2).

50. Achimas-Cadariu P, Harter P, Fisseler-Eckhoff A, Beutel B, Traut A, Du BA. Assessment of the sentinel lymph node in patients with invasive squamous carcinoma of the vulva. *Acta Obstet Gynecol Scand* 2009;**88**:1209–14. http://dx.doi.org/10.3109/00016340903317982

51. Basta A, Pitynski K, Basta P, Hubalewska-Hola A, Oplawski M, Przesziakowski D. Sentinel node in gynaecological oncology. *Rep Pract Oncol Radiother* 2005;**10**:1–4.

52. Brunner AH, Polterauer S, Tempfer C, Joura E, Reinthaller A, Horvat R, *et al.* The accuracy of intraoperative frozen section of the inguinal sentinel lymph node in vulvar cancer. *Anticancer Res* 2008;**28**:4091–4.

53. Camara O, Gonnert H, Herrmann J, Egbe A, Diebolder H, Gajda M, *et al.* Sentinel lymph node biopsy in vulvar cancer: a pilot study. *Eur J Gynaecol Oncol* 2009;**30**:622–4.

54. Crosbie EJ, Winter-Roach B, Sengupta P, Sikand KA, Carrington B, Murby B, *et al.* The accuracy of the sentinel node procedure after excision biopsy in squamous cell carcinoma of the vulva. *Surg Oncol* 2010;**19**:e150–4. http://dx.doi.org/10.1016/j.suronc.2010.08.003

55. De Cicco C, Sideri M, Bartolomei M, Grana C, Cremonesi M, Fiorenza M, *et al.* Sentinel node biopsy in early vulvar cancer. *Br J Cancer* 2000;**82**:295–9.

56. de Hullu JA, Hollema H, Piers DA, Verheijen RH, van Diest PJ, Mourits MJ, *et al.* Sentinel lymph node procedure is highly accurate in squamous cell carcinoma of the vulva. *J Clin Oncol* 2000;**18**:2811–16.

57. Hampl M, Hantschmann P, Michels W, Hillemanns P. Validation of the accuracy of the sentinel lymph node procedure in patients with vulvar cancer: results of a multicenter study in Germany. *Gynecol Oncol* 2008;**111**:282–8. http://dx.doi.org/10.1016/j.ygyno.2008.08.007

58. Hauspy J, Beiner M, Harley I, Ehrlich L, Rasty G, Covens A. Sentinel lymph node in vulvar cancer. *Cancer* 2007;**110**:1015–23. http://dx.doi.org/10.1002/cncr.22874

59. Johann S, Klaeser B, Krause T, Mueller MD. Comparison of outcome and recurrence-free survival after sentinel lymph node biopsy and lymphadenectomy in vulvar cancer. *Gynecol Oncol* 2008;**110**:324–8. http://dx.doi.org/10.1016/j.ygyno.2008.04.004

60. Klat J, Sevcik L, Simetka O, Graf P, Waloschek T, Kraft O, *et al.* Characteristics of sentinel lymph nodes' metastatic involvement in early stage of vulvar cancer. *Aust N Z J Obstet Gynaecol* 2009;**49**:672–6.

61. Levenback C, Coleman RL, Burke TW, Bodurka-Bevers D, Wolf JK, Gershenson DM. Intraoperative lymphatic mapping and sentinel node identification with blue dye in patients with vulvar cancer. *Gynecol Oncol* 2001;**83**:276–81. http://dx.doi.org/10.1006/gyno.2001.6374

62. Louis-Sylvestre C, Evangelista E, Leonard F, Itti E, Meignan M, Paniel BJ. Interpretation of sentinel node identification in vulvar cancer. *Gynecol Obstet Fertil* 2006;**34**:706–10.

63. Martinez-Palonez JM, Perez-Benavente MA, Gil-Moreno A, Diaz-Feijoo B, Roca I, Garcia-Jimenez A, *et al.* Comparison of recurrence after vulvectomy and lymphadenectomy with and without sentinel node biopsy in early stage vulvar cancer. *Gynecol Oncol* 2006;**103**:865–70. http://dx.doi.org/10.1016/j.ygyno.2006.05.024

64. Merisio C, Berretta R, Gualdi M, Pultrone DC, Anfuso S, Agnese G, *et al.* Radioguided sentinel lymph node detection in vulvar cancer. *Int J Gynecol Cancer* 2005;**15**:493–7. http://dx.doi.org/10.1111/j.1525-1438.2005.15314.x

65. Moore RG, Robison K, Brown AK, DiSilvestro P, Steinhoff M, Noto R, *et al.* Isolated sentinel lymph node dissection with conservative management in patients with squamous cell carcinoma of the vulva: a prospective trial. *Gynecol Oncol* 2008;**109**:65–70. http://dx.doi.org/10.1016/j.ygyno.2007.12.027

66. Nyberg RH, Iivonen M, Parkkinen J, Kuoppala T, Maenpaa JU. Sentinel node and vulvar cancer: a series of 47 patients. *Acta Obstet Gynecol Scand* 2007;**86**:615–19. http://dx.doi.org/10.1080/00016340701286793

67. Pitynski K, Basta A, Oplawski M, Przeszlakowski D, Hubalewska-Hola A, Krysztopowicz W. Lymph node mapping and sentinel node detection in carcinoma of the cervix, endometrium and vulva. *Ginekol Pol* 2003;**74**:830–5.

68. Radziszewski J, Kowalewska M, Jedrzejczak T, Kozlowicz-Gudzinska I, Nasierowska-Guttmejer A, Bidzinski M, *et al.* The accuracy of the sentinel lymph node concept in early stage squamous cell vulvar carcinoma. *Gynecol Oncol* 2010;**116**:473–7. http://dx.doi.org/10.1016/j.ygyno.2009.10.072

69. Rob L, Robova H, Pluta M, Strnad P, Kacirek J, Skapa P, *et al.* Further data on sentinel lymph node mapping in vulvar cancer by blue dye and radiocolloid Tc99. *Int J Gynecol Cancer* 2007;**17**:147–53. http://dx.doi.org/10.1111/j.1525-1438.2007.00806.x

70. Terada KY, Shimizu DM, Jiang CS, Wong JH. Outcomes for patients with T1 squamous cell cancer of the vulva undergoing sentinel node biopsy. *Gynecol Oncol* 2006;**102**:200–3. http://dx.doi.org/10.1016/j.ygyno.2005.11.042

71. Vakselj A, Bebar S. The role of sentinel lymph node detection in vulvar carcinoma and the experiences at the Institute of Oncology Ljubljana. *Radiol Oncol* 2007;**41**:167–73. http://dx.doi.org/10.2478/v10019-007-0027-4

72. Van den Eynden J, Lannoo L, Amant F, Stroobants S, Vergote I. Sentinel node biopsy in the treatment of vulvar cancer. *Tijdschr voor Geneeskunde* 2003;**59**:1169–78.

73. Van der Zee AG, Oonk MH, de Hullu JA, Ansink AC, Vergote I, Verheijen RH, *et al.* Sentinel node dissection is safe in the treatment of early-stage vulvar cancer. *J Clin Oncol* 2008;**26**:884–9. http://dx.doi.org/10.1200/JCO.2007.14.0566

74. Vidal-Sicart S, Puig-Tintore LM, Lejarcegui JA, Paredes P, Ortega ML, Munoz A, *et al.* Validation and application of the sentinel lymph node concept in malignant vulvar tumours. *Eur J Nucl Med Mol Imaging* 2007;**34**:384–91. http://dx.doi.org/10.1007/s00259-006-0237-9

75. Sliutz G, Reinthaller A, Lantzsch T, Mende T, Sinzinger H, Kainz C, *et al.* Lymphatic mapping of sentinel nodes in early vulvar cancer. *Gynecol Oncol* 2002;**84**:449–52. http://dx.doi.org/10.1006/gyno.2001.6572

76. Hefler LA, Grimm C, Six L, Seebacher V, Polterauer S, Joura E, *et al.* Inguinal sentinel lymph node dissection vs. complete inguinal lymph node dissection in patients with vulvar cancer. *Anticancer Res* 2008;**28**:515–17.

77. de Hullu JA, Doting E, Piers DA, Hollema H, Aalders JG, Koops HS, *et al.* Sentinel lymph node identification with technetium-^{99m}-labeled nanocolloid in squamous cell cancer of the vulva. *J Nucl Med* 1998;**39**:1381–5.

78. Levenback C, Burke TW, Gershenson DM, Morris M, Malpica A, Ross MI. Intraoperative lymphatic mapping for vulvar cancer. *Obstet Gynecol* 1994;**84**:163–7.

79. Levenback C, Burke TW, Morris M, Malpica A, Lucas KR, Gershenson DM. Potential applications of intraoperative lymphatic mapping in vulvar cancer. *Gynecol Oncol* 1995;**59**:216–20.

80. Frumovitz M, Ramirez PT, Tortolero-Luna G, Malpica A, Eifel P, Burke TW, *et al.* Characteristics of recurrence in patients who underwent lymphatic mapping for vulvar cancer. *Gynecol Oncol* 2004;**92**:205–10. http://dx.doi.org/10.1016/j.ygyno.2003.09.022

81. Rob L, Robova H, Pluta M, Strnad P, Kacirek J, Chmel R, *et al.* Sentinel lymph nodes identification in vulvar cancer–methods and technique. *Ceska Gynekol* 2006;**71**:298–301.

82. Oonk MH, van Os MA, de Bock GH, de Hullu JA, Ansink AC, van der Zee AG. A comparison of quality of life between vulvar cancer patients after sentinel lymph node procedure only and inguinofemoral lymphadenectomy. *Gynecol Oncol* 2009;**113**:301–5. http://dx.doi.org/10.1016/j.ygyno.2008.12.006

83. Oonk MH, van Hemel BM, Hollema H, de Hullu JA, Ansink AC, Vergote I, *et al.* Size of sentinel-node metastasis and chances of non-sentinel-node involvement and survival in early stage vulvar cancer: results from GROINSS-V, a multicentre observational study. *Lancet Oncol* 2010;**11**:646–52. http://dx.doi.org/10.1016/S1470-2045(10)70104-2

84. Vidal-Sicart S, Puig-Tintore LM, Ordi J, Lejarcegui JA. Localisation of sentinel lymph nodes in cancer of the vulva. *Ginecol Clin Quir* 2002;**3**:67–71.

85. Vidal-Sicart S, Domenech B, Lujan B, Pahisa J, Torne A, Martinez-Roman S, *et al.* Sentinel node in gynaecological cancers. Our experience. *Rev Esp Med Nucl* 2009;**28**:221–8. http://dx.doi.org/10.1016/S1578-200X(09)70022-1

86. Puig-Tintore LM, Ordi J, Vidal-Sicart S, Lejarcegui JA, Torne A, Pahisa J, *et al.* Further data on the usefulness of sentinel lymph node identification and ultrastaging in vulvar squamous cell carcinoma. *Gynecol Oncol* 2003;**88**:29–34. http://dx.doi.org/10.1006/gyno.2002.6857

87. Stehman FB, Bundy BN, Thomas G, Varia M, Okagaki T, Roberts J, *et al.* Groin dissection versus groin radiation in carcinoma of the vulva: a Gynecologic Oncology Group study. *Int J Radiat Oncol Biol Phys* 1992;**24**:389–96. http://dx.doi.org/10.1016/0360-3016(92)90699-I

88. Manavi M, Berger A, Kucera E, Vavra N, Kucera H. Does T1, N0-1 vulvar cancer treated by vulvectomy but not lymphadenectomy need inguinofemoral radiation? *Int J Radiat Oncol Biol Phys* 1997;**38**:749–53. http://dx.doi.org/10.1016/S0360-3016(97)00060-6

89. Stehman FB, Bundy BN, Dvoretsky PM, Creasman WT. Early stage I carcinoma of the vulva treated with ipsilateral superficial inguinal lymphadenectomy and modified radical hemivulvectomy: a prospective study of the Gynecologic Oncology Group. *Obstet Gynecol* 1992;**79**:490–7.

90. Katz A, Eifel PJ, Jhingran A, Levenback CF. The role of radiation therapy in preventing regional recurrences of invasive squamous cell carcinoma of the vulva. *Int J Radiat Oncol Biol Phys* 2003;**57**:409–18. http://dx.doi.org/10.1016/S0360-3016(03)00591-1

91. Anderson JM, Cassady JR, Shimm DS, Stea B. Vulvar carcinoma. *Int J Radiat Oncol Biol Phys* 1995;**32**:1351–7. http://dx.doi.org/10.1016/0360-3016(95)00090-L

92. Andrews SJ, Williams BT, DePriest PD, Gallion HH, Hunter JE, Buckley SL, *et al.* Therapeutic implications of lymph nodal spread in lateral T1 and T2 squamous cell carcinoma of the vulva. *Gynecol Oncol* 1994;**55**:41–6. http://dx.doi.org/10.1006/gyno.1994.1244

93. Burke TW, Levenback C, Coleman RL, Morris M, Silva EG, Gershenson DM. Surgical therapy of T1 and T2 vulvar carcinoma: further experience with radical wide excision and selective inguinal lymphadenectomy. *Gynecol Oncol* 1995;**57**:215–20. http://dx.doi.org/10.1006/gyno.1995.1128

94. Busch M, Wagener B, Schaffer M, Duhmke E. Long-term impact of postoperative radiotherapy in carcinoma of the vulva FIGO I/II. *Int J Radiat Oncol Biol Phys* 2000;**48**:213–18. http://dx.doi.org/10.1016/S0360-3016(00)00586-1

95. de Hullu JA, Hollema H, Lolkema S, Boezen M, Boonstra H, Burger MP, *et al.* Vulvar carcinoma. The price of less radical surgery. *Cancer* 2002;**95**:2331–8. http://dx.doi.org/10.1002/cncr.10969

96. DeSimone CP, Van Ness JS, Cooper AL, Modesitt SC, DePriest PD, Ueland FR, *et al.* The treatment of lateral T1 and T2 squamous cell carcinomas of the vulva confined to the labium majus or minus. *Gynecol Oncol* 2007;**104**:390–5. http://dx.doi.org/10.1016/j.ygyno.2006.08.035

97. Farias-Eisner R, Cirisano FD, Grouse D, Leuchter RS, Karlan BY, Lagasse LD, *et al.* Conservative and individualized surgery for early squamous carcinoma of the vulva: the treatment of choice for stage I and II (T1-2N0-1M0) disease. *Gynecol Oncol* 1994;**53**:55–8.

98. Hallak S, Ladi L, Sorbe B. Prophylactic inguinal-femoral irradiation as an alternative to primary lymphadenectomy in treatment of vulvar carcinoma. *Int J Oncol* 2007;**31**:1077–85.

99. Kumar S, Shah JP, Bryant CS, Imudia AN, Morris RT, Malone JM Jr. A comparison of younger vs. older women with vulvar cancer in the United States. *Am J Obstet Gynecol* 2009;**200**:e52–5. http://dx.doi.org/10.1016/j.ajog.2008.09.869

100. Scheistroen M, Nesland JM, Trope C. Have patients with early squamous carcinoma of the vulva been overtreated in the past? The Norwegian experience 1977–1991. *Eur J Gynaecol Oncol* 2002;**23**:93–103.

101. Tantipalakorn C, Robertson G, Marsden DE, Gebski V, Hacker NF. Outcome and patterns of recurrence for International Federation of Gynecology and Obstetrics (FIGO) stages I and II squamous cell vulvar cancer. *Obstet Gynecol* 2009;**113**:895–901.

102. Vavra N, Kucera H, Egarter C, Weghaupt K. Effect of inguinal lymph node irradiation on the treatment result of vulvar cancer with different risk factors. *Geburtshilfe Frauenheilkd* 1990;**50**:470–6.

103. Office for National Statistics. *Natural Death Rates. 2010 Mid-Year Estimates*. URL: www.ons.gov.uk/ons/rel/vsob1/death-reg-sum-tables/2010/index.html (accessed 1 November 2011).

104. Fonseca-Moutinho JA, Coelho MC, Silva DP. Vulvar squamous cell carcinoma. Prognostic factors for local recurrence after primary en bloc radical vulvectomy and bilateral groin dissection. *J Reprod Med* 2000;**45**:672–8.

105. Anon. *NHS Reference Costs 2009/2010*. London: Department of Health; 2010.

106. Levenback CF, Ali S, Coleman RL, Gold M, Fowler JM, Judson PL, *et al.* Lymphatic mapping and sentinel lymph node biopsy in women with squamous cell carcinoma of the vulva: a Gynecologic Oncology Group study. *J Clin Oncol* 2012;**30**:3786–91. http://dx.doi.org/10.1200/JCO.2011.41.2528

107. Meads C, Auguste P, Davenport C, Malysiak S, Sundhar 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**(12).

Appendix 1 Protocol

Sentinel lymph node (SLN) status in vulval cancer: Systematic quantitative reviews and decision-analytic model based economic evaluation

1 Clinical background

Vulval cancer accounts for approximately 3–5% of all gynaecological malignancies and 1% of all cancers in women, with an incidence rate of 1–2/100,000.[1] In the UK, it affects approximately 1,063 women every year with a 1 in 316 lifetime risk of developing vulval cancer.[2] Mortality data from 2007 shows 384 deaths in the UK.[2] Squamous cell carcinomas (SCC) account for more than 90% of vulval cancers;[3] the other 10% include melanomas, sarcomas, basal cell carcinomas and adenocarcinomas.[4] Although its peak incidence is in the 7th decade, there has been a significant increase in rates of vulval cancer in younger women. The proportion of women diagnosed with this cancer under the age of 50 rose from 6% in 1975 to 15% in 2006.[2] This trend has been observed in many countries, and has been linked to the rising incidence of vulval intraepithelial neoplasia (VIN) in young women caused by infection with HPV.[5,6]

Vulval cancer is curable when diagnosed at an early stage. The standard treatment for squamous cell carcinoma of the vulva is radical surgery, which in all but stage Ia or superficially invasive disease includes inguinofemoral lymphadenectomy.[7] The inguinal lymph node status has been identified as the single most important factor in predicting mortality attributable to vulval cancer.[8] Overall, about a third of patients with operable disease have nodal spread.[7] Those patients with primary lesions not more than 2 cm, who are inguinal node negative have a 98% 5-year survival rate, while those with any size lesion and three or more unilateral nodes or two or more bilateral nodes associated have a 29% 5-year survival rate.[9] Morbidity from lymphadenectomy is high with significant negative impact on the Quality of Life (QoL). Nodal assessment with biopsy is currently not routinely performed in practice.

The likelihood of metastasis is related to the size and the depth of the primary tumour. In stage Ia, this likelihood is almost zero, and rises once invasion extends beyond 1 mm depth. At the time of presentation, up to 25% of patients with vulval cancer are stage I, and of these, 30% are stage Ia. In absolute terms, this means that in any one year there will be a requirement for 700–750 groin lymph node dissections in the UK.[10] However, only around 30% of these operated cases will have evidence of nodal involvement;[7] the rest being node negative. This project will assess if nodal biopsy can be accurately and efficiently performed to direct the need for further lymphadenectomy.

1.1 Current clinical practice

Traditionally, the management of vulval cancer involves radical surgery which includes the excision of the primary lesion and unilateral or bilateral superficial and deep inguinofemoral lymphadenectomy.[7] The efficacy of this treatment is good, with reported groin recurrence rates varying between 1% and 10%.[11] However, as only 25% – 35% of patients with early-stage disease will have lymph node metastases,[7,12] and the remaining 65% – 75% possibly do not benefit from elective inguinofemoral lymphadenectomy while risking significant morbidity.[13] In the short term, wound healing in the groin is compromised by infection and breakdown in 20% to 40% of patients.[13] In the long term, lymphoedema of the legs with increased risk for erysipelas occurs in 30% to 70% of patients.[13] These complications can be incapacitating with major impact on sexual and psychological function.[14] Patients are also subjected to groin radiotherapy if cancer metastasis is detected on histopathological examination of lymph nodes. Patients treated with both complete inguinofemoral lymph node dissection and external beam radiotherapy to groin nodes suffer the morbidity of both treatments with a higher risk of lymphoedema and cellulitis.

Despite significant surgical morbidity and a low frequency of lymph node metastases, an elective inguinofemoral lymphadenectomy is regarded as standard of care. This is because unrecognised disease in

the inguinofemoral lymph nodes is nearly always fatal.[13] An accurate test is needed that could identify those patients in whom the risk of metastases is low. Such a test could help exclude the need for lymphadenectomy and would be extremely valuable. A minimally invasive technique to detect metastasis to the groin lymph nodes has a huge potential to reduce unnecessary morbidity.

1.2 Tests for nodal involvement

Assessment of the nodal status by clinical palpation of the groins is inadequate; of patients with clinically normal lymph nodes, 16–24% has metastases, while 24–41% of those with clinically involved nodes are negative at histological examination.[15,16] There are several minimally and non-invasive tests available for the status of groin nodes in vulval cancer, but none are routinely used in clinical practice. These include ultrasonography with or without fine-needle aspiration, computerised tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET) and SLN identification (using blue dye or technetium-99m-labelled nanocolloid—99mTc).[7] This call for proposals focuses on the value of sentinel node biopsy using 99mTc or blue dye.

1.3 Sentinel lymph node (SLN) identification by lymphoscintigraphy

SLN refers to any lymph node that receives drainage directly from the primary tumour[17] (Fig 1). The SLN will be the lymph node with the highest probability of containing metastasis. Removal of SLN should therefore allow assessment of the status of lymphatic basin without the need to remove all the lymph nodes, providing an opportunity to avoid the morbidity associated with formal lymphadenectomy. If SLN is negative the rest of the groin should be at least risk of having subclinical metastasis.[18] The removal of fewer nodes (typically 1-3/groin) also permits more focussed pathological assessment of the SLN and direct pathological resources compared with 10-20 nodes removed by inguinofemoral lymphadenectomy.[18] SLNs are identified by isosulfan blue dye or 99mTc enhanced lymphoscintigraphy alone or in combination.[18]

Pre operatively 99mTc nanocolloid is injected around the primary tumour and dynamic images are obtained. The first site of focal accumulation is considered to be the sentinel node. During surgery a gamma probe (typically a caesium iodide scintillator) is used to confirm the location of the SLN. Once the SLN have been removed the groin is rescanned to ensure removal of all SLNs. A background count after SLN removal of < 10% of the initial value is typically used to confirm that all relevant nodes have been sampled. The signal arising from the SLN will depend on the dose of isotope used, the time lapse since injection, the number of and rate of flow within lymphatic channels and the distance between the probe and the target SLN. Where the signal from the SLN is weak, removal of the primary injection site on the vulval may be performed first, to facilitate localisation of the SLN in the absence of background signal.

99mTc enhanced lymphoscintigraphy has the advantage that in comparison to blue dye it facilitates localisation of the SLN even prior to the skin incision, potentially enabling a smaller groin incision (2 cm) to be used. The use of isotope also permits detection of sentinel nodes outside the usual basin, with the identification of aberrant drainage in those cases of clitoral involvement. Use of radiocolloid alone in vulvar carcinoma can avoid complications from blue dye (allergic reactions, permanent staining, and false oximeter readings.[19] This technique has already become well established in breast cancer where it was found to be better than blue dye in the identification of the sentinel node[20] and available data in vulval cancer would support the increased sensitivity afforded by the use of isotope.

Blue dye is injected intraoperatively around the site of primary tumour. This test is done in isolation or in combination with 99mTc enhanced lymphoscintigraphy. The use of blue dye to identify the sentinel node has the advantage that the entire test is performed under general anaesthetic at the time of the operation; hence additional preoperative assessment is not required. This technique does require a sizable skin incision to allow the surgeon dissection down to Camper's fascia and identification of the blue afferent channel to the blue/sentinel node.

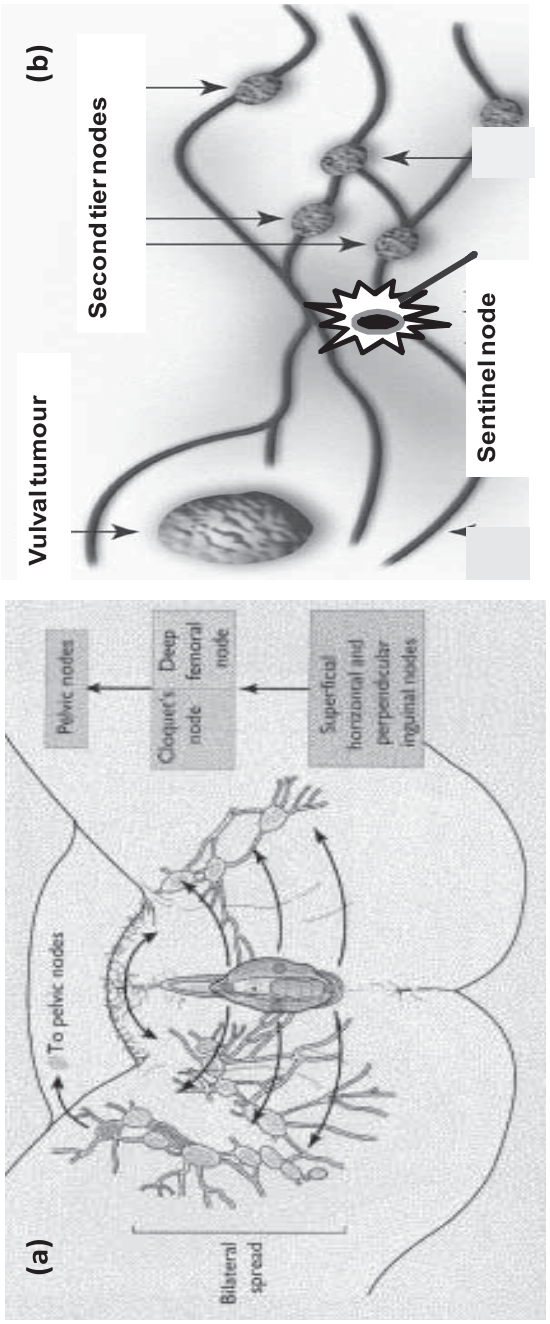


Fig 1. Lymphatic drainage of vulva (a) and radioactivity in sentinel node in vulval cancer (b)

(Adapted from Nature Biotech 2004; 22: 38-39 and Eur J Surg Oncol 2006; 32: 825-831)

1.4 Histopathological examination method of SLN

The accuracy of SLN biopsy in staging vulval cancer depends on the histopathological technique used for the examination of the sentinel nodes. Post-operative pathological assessment of the SLN may be combined with intra operative frozen section analysis.[18] Frozen section offers the advantage of performing synchronous complete lymphadenectomy and SLN biopsy if the SLNs are positive. The sensitivity of frozen section usually by single haematoxylin and eosin (H/E) staining is around 80%.[21] Apart from the concerns of accuracy with high false negative rates, there is a risk of loss of diagnostic tissue with intra operative methods. Standard pathologic examination of lymph nodes, i.e., H&E staining of a bivalved node, will sample only a fraction of the resected tissue. This sampling methodology could potentially fail to detect lymph node metastasis in the SLN producing a false negative result. Detailed analysis with serial sectioning and immunohistochemical staining can identify micrometastases that are not otherwise apparent. Enhanced pathologic analysis, termed "ultrastaging," generally involves serial sections through the node and application of specific immunohistochemical staining for epithelial antigens. Post-operative ultrastaging is labour intensive and could not be performed on the large numbers of nodes removed at formal lymphadenectomy. It has been shown to improve the sensitivity in detecting micro and macro metastatic deposits compared with routinely stained sections.[22] Whilst the biological relevance of metastases detected by ultrastaging remains controversial, groin recurrence has been identified in patients with micrometastasis only in the SLN.[18] Reverse transcriptase-polymerase chain reaction (RT-PCR) has the potential to accurately detect micrometastasis.[20] Their use is limited for SLN biopsy in other tumours with very little data in vulval cancer. The therapeutic impact of stage migration due to detection of micrometastasis by the use of techniques with ever increasing sensitivity compared with standard histopathology needs further evaluation.

1.5 Role of SLN biopsy in clinical management of vulval cancer

Early vulval cancer may be treated with radical excision of the primary tumour in combination with SLN biopsy. Most protocols utilise ultrastaging for assessment of the SLN. Where the SLNs are negative no further treatment is employed and the patient is observed. Where such a protocol is followed for patients with small (< 4 cm) tumours and no obvious preoperative metastases, the groin recurrence rate is low (3% including multifocal disease; 2.3% in unifocal vulvar disease), there is excellent disease-specific survival rate of 97% at 3 years and minimal treatment-related morbidity.[13] If SLNs are positive for micro or macrometastasis, the patients typically undergo inguinofemoral lymphadenectomy. For those with a metastasis > 5 mm, more than one intranodal metastasis and/or extranodal spread, postoperative external-beam radiotherapy (50 Gy) to the groin/pelvis is recommended, possibly combined with chemotherapy. There are no randomised controlled trials (RCT) in vulval cancer that compare the effect of SLN biopsy on treatment and outcome due to the rarity of the condition. In fact, the EORTC (European Organisation for Research and Treatment of Cancer) withheld an attempt to assess SLN biopsy by an RCT as the power calculation estimated a prohibitively large sample size.[21] An ongoing study (GROINSS-V II GROningen INternational Study on Sentinel nodes in Vulvar cancer) is investigating the safety of omitting further surgery in such SLN positive patients but this is not yet within standard practice.

The performance of SLN biopsy is known to be associated with a learning curve.[23] Due to the consistent pattern of lymphatic drainage of the vulva the learning curve is considered to be much steeper for vulval cancer than other tumours. Introduction of SLN biopsy into routine practice will require quality control at each step. This multidisciplinary procedure, includes injection of radioactive tracer by either the surgeon or a nuclear medicine physician familiar with vulval anatomy, careful interpretation of lymphoscintigram, a surgeon with successful experience (sentinel node procedure followed by full lymphadenectomy) in at least 10 patients, and a pathology department experienced in ultrastaging (laterally sectioning of the nodes at 3-mm intervals and then each block cut at 400-μm intervals) of the sentinel nodes.[13]

There is a need to systematically review the comparative accuracy of 99mTc and blue dye for SLN biopsy in staging for vulval cancer. Moreover it is important to review how testing of sentinel nodes will impact on staging, therapeutic options and outcomes. Fig 2 conceptualises the role of SLN biopsy in the management of vulval cancer to be evaluated in this project.

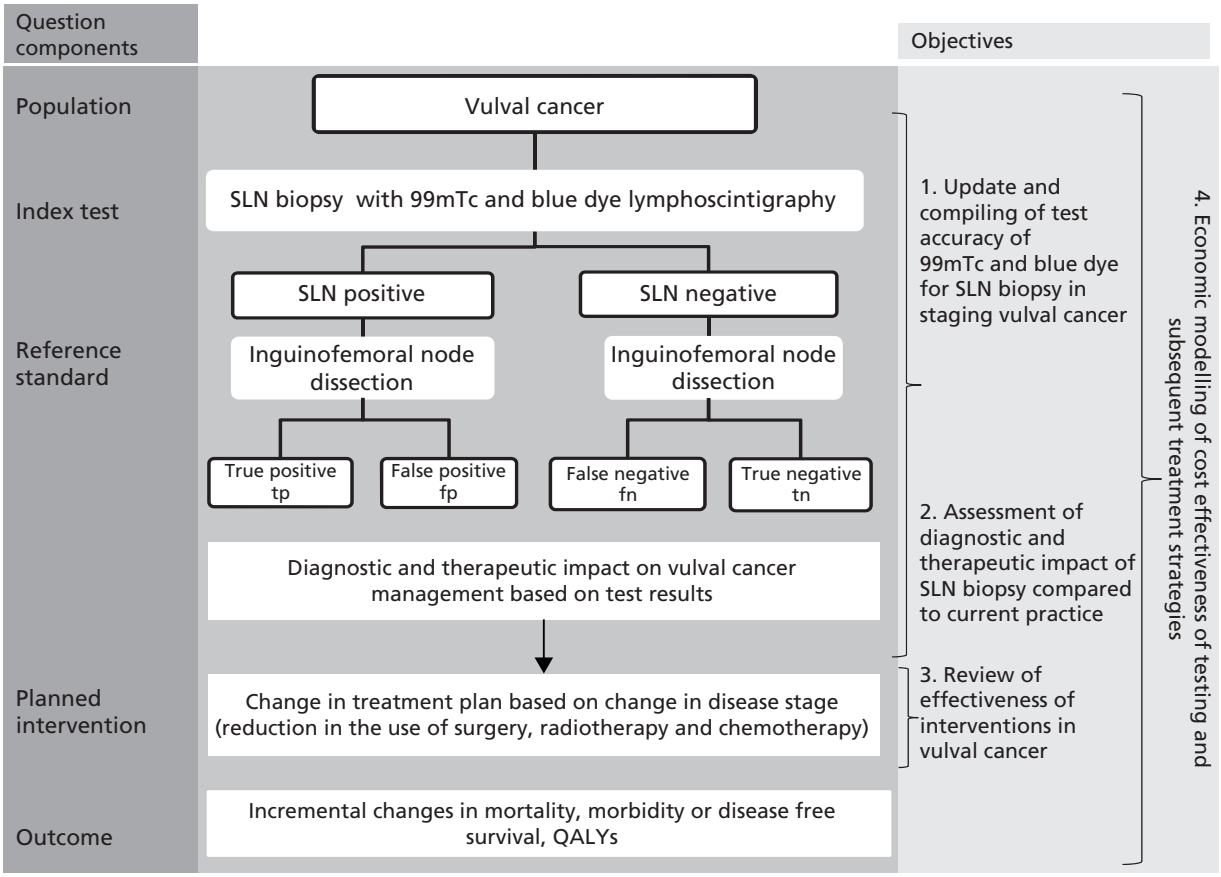


Fig 2 SLN biopsy and treatment strategies in women with vulval cancer.

2 Work leading to the proposal

An MRC Fellowship awarded to TS supervised by KSK and TR conducted systematic reviews of accuracy of tests for lymph node metastasis in gynaecologic oncology[24] (PhD awarded at University of Birmingham in 2009). We published a systematic review of accuracy of tests that evaluate node status in vulval cancer in 2005. We have also initiated a decision-analytic model based economic evaluation to determine the relative costs and effectiveness of a range of alternative preoperative tests and subsequent management strategies for inguinofemoral lymph nodes.[24] This HTA call for proposals gives us the opportunity to update the accuracy review, to use more robust statistical methods for meta analysis,[25] to complete the economic evaluation initiated, and to undertake a probabilistic sensitivity analysis on the model.

2.1 Systematic review of accuracy of tests for sentinel node status

We identified relevant literature from 1974–2005 to conduct the review using a prospective protocol and widely recommended methods.[26,27] The initial search generated 1154 citations from which 82 articles were potentially relevant. After assessment of the full manuscripts, a total of 24 articles that reported 29 tests were selected.[10] Studies included in the review were those that compared the index test to the histological evaluation of inguinofemoral lymphadenectomy specimen. Eleven studies evaluated the accuracy of 99mTc and 8 studied blue dye for SLN biopsy in staging. SLN biopsy using 99mTc had a pooled sensitivity and negative Likelihood ratio (LR-) of 97% (91–100 95% CI) and 0.12 (0.053–0.28 95% CI), respectively, and was the most accurate of the tests reviewed.[10] Blue dye alone for identification of SLN had a pooled sensitivity of 95% (82–99 95% CI) and LR- of 0.16 (0.07–0.32 95% CI).[10] This review needs updating with reassessment of the study quality, use of bivariate meta analysis and metaregression to compare 99mTc vs blue dye.

Table 1 Accuracy of studies evaluating sentinel nodes with 99mTc enhanced lymphoscintigraphy and blue dye in vulval cancer

Study (year)	TP	FP	TN	FN	Sensitivity (95% CI)	Specificity (95% CI)	LR + (95% CI)	LR– (95% CI)
Isosulfan blue dye								
Levenback 1994	2	0	5	0	100% (63–100)	100% (48–100)	10 (0.67–149)	0.18 (0.01–2.31)
Levenback 1995	3	0	15	0	100% (29–100)	100% (78–100)	28 (1.78–439.41)	0.13 (0.01–1.73)
Ansink et al 1999	9	0	41	2	81% (48–98)	100% (91–100)	66.5 (4.17–1061)	0.21 (0.07–0.64)
Echt 1999	2	0	7	0	100% (16–100)	100% (59–100)	13.3 (0.87–204.7)	0.18 (0.01–2.25)
Molups 2001	2	0	6	0	100% (16–100)	100% (54–100)	11.6 (0.77–176.84)	0.18 (0.01–2.27)
Levenback 2001	10	0	45	0	100% (69–100)	100% (92–100)	88 (5.6–1387)	0.05 (0.003–0.69)
Puig-Tintore 2003	5	0	24	0	100% (48–100)	100% (86–100)	45.8 (2.92–720.2)	0.08 (0.006–1.21)
Moore 2003	3	0	16	0	100% (29–100)	100% (79–100)	29.75 (1.89–456.0)	0.13 (0.01–1.72)
Summary	36	0	159	2	95% (82–99)	100% (98–100)	27.4 (10.4–72.2)	0.16 (0.07–0.32)
99mTc identification								
DeCesart 1997	10	0	0	0	96% (68–100)	50% (6–95)	–	–
de Hullu 1998	2	0	8	0	100% (16–100)	100% (63–100)	15 (.96–232)	0.18 (0.014–2.23)
De Cicco 2000	8	0	29	0	100% (16–100))	100% (88–100)	50 (3.06–817.98)	0.17 (0.013–2.13)
De Hullu 2000	27	0	68	0	100% (87–100)	100% (95–100)	135.54 (8.6–2146)	0.018 (0.001–0.28)
Sideri 2000	13	0	31	0	100% (75–100)	100% (89–100)	61.7 (3.94–867.23)	0.04 (0.002–0.52)
Molpus 2001	2	0	6	0	100% (16–100)	100% (54–100)	11.7 (0.78–176.84)	0.18 (0.01–2.27)
Tavares 2002	3	0	12	0	100% (29–100)	100% (73–1000)	22.8 (1.46–353.43)	0.13 (0.01–1.74)
Boran 2002	4	0	11	2	67% (22–96)	100% (71–100)	15.4 (0.97–245.88)	0.37 (0.14–1.01)
Sliutz 2002	9	0	17	0	100% (66–100)	100% (80–100)	34.2 (2.22–527.99)	0.05 (0.003–0.76)
Moore 2003	9	0	22	0	100% (66–100)	100% (85–100)	43.7 (2.81–680.31)	0.05 (0.003–0.76)
Puig-Tintore 2003	5	0	25	0	100% (48–100)	100% (86–100)	47.7 (3.03–749.83)	0.08 (0.006–1.21)
Summary	91	0	229	2	97% (91–100)	100% (98–100)	33.4 (14.–79.8)	0.12 (0.053–0.28)

2.2 Prospective observational study of SLN biopsy in early vulval cancer (GROINSS-V GROningen INternational Study on Sentinel nodes in Vulvar cancer)

PB was the principal UK investigator of the GROINSS-V study and is now the chief investigator for the UK in the follow-on GROINSS-V II study. GROINSS-V was a large, prospective, multicentre observational study on SLN detection using radioactive tracer and blue dye in patients with early vulval cancer.[13] The study demonstrated that for appropriately selected patients, sentinel node dissection appears to be safe, reliable and associated with reduced morbidity as compared with formal inguinofemoral lymphadenectomy. The accuracy of SLN biopsy has been verified by two different reference standards. When SLN was found to be negative, inguinofemoral lymphadenectomy was omitted, and the patient was observed with follow-up for 2 years at intervals of every 2 months for groin recurrences. There is a risk of bias with this differential verification. Nevertheless GROINSS-V is the largest single well conducted study (n = 403) to date compared with other studies evaluating SLN biopsy. Furthermore it offers prospective data to evaluate the actual therapeutic impact of the alternate strategy of performing SLN biopsy without routine inguinofemoral lymphadenectomy.

2.3 Review of outcomes following inguinofemoral lymphadenectomy

We systematically reviewed the literature (1974–2005) for relevant clinical outcomes, Quality Adjusted Life Years (QALYs) and early and late complication rates following inguinal femoral lymphadenectomy.[24] The 5 year survival rate was estimated from published studies for node negative and node positive patients, the range of survival depending on the size of primary vulval lesion and the number of positive lymph nodes. For those with negative lymph nodes as point estimate for five year survival was 84% (70–98%) and for node positive patients 42.5% (25–60%). QALYs are the preferred outcome taking into account both the quantity and quality of life.[24] We failed to identify any studies that had used QALY data. As patients suffering for breast cancer also suffer similar lymphoedema a review of that literature was undertaken, but this also failed to provide relevant data. The data on complication rates came from studies specifically reporting on complications of lymphadenectomy excluding the radical vulvectomy procedure and from those that used a triple incision approach to the management of vulval cancer.[28] Two studies reported immediate post operative complication rates ranging from 44% to 66%.[29] A literature review of long-term complications found variation in rates from 12–51% with an average point estimate for the model of 34%.[30–32] The duration of inpatient stay was taken from local hospital statistics (Birmingham Women's Hospital NHS foundation Trust). We will obtain updated data from current inpatient statistics from the Pan Birmingham gynaecological cancer centre and the Addenbrookes NHS trust for this project. This review needs updating with recent searches and where published data are not available, we will contact the individual specialist centres to provide more information. This will enable us to evaluate the diagnostic and therapeutic impact of performing SLN biopsy over inguinofemoral lymphadenectomy.

2.4 Model based cost-effectiveness analysis of testing for sentinel nodes

We initiated an economic evaluation using a decision analytic model to compare strategies that used the results from various pre operative tests to determine the need for performing inguinofemoral lymphadenopathy in women over 70 years of age with vulval cancer.[24] This work is not prepared for publication as there are several areas for improvement identified below. Fig 3 shows a subset of the model to illustrate the approach developed.

The outcomes were cost per morbidity-free 5 year survival and cost per death avoided at 5 year. Absence of relevant data prevented the estimation of cost per QALY. The strategy of inguinal femoral lymphadenectomy for all patients with vulval cancer without pre operative lymph node testing was the most expensive and not the most effective in avoiding death. It was the least effective in providing morbidity-free survival. The most effective strategy for both was that of pre operative SLN biopsy. This had an annual Incremental cost-effectiveness ratio (ICER) of £33,079 for the outcome of death avoided and an ICER of £100,888 for morbidity-free survival. Both ICERs were above the willingness to pay threshold set by NICE of £30,000. Through this exploratory work we have identified the following areas of additional work for a robust evaluation through this project. We will construct the model for women with vulval cancer of all age groups. We will evaluate the development of late complications in those women suffering early complications and incorporate data on inpatient stay. We will perform probabilistic sensitivity analysis. With this exhaustive evaluation we hope to have reliable answers to guide practice.

3 Research Objectives

The commissioning brief is for an evidence synthesis of the added value of 99mTc enhanced lymphoscintigraphy for SLN biopsy in staging for women with vulval cancer in comparison to the current practice of inguinofemoral lymphadenectomy. Our project will follow the key steps involved in health technology assessment of tests[33] and will meet the commissioned brief by fulfilling the following objectives:

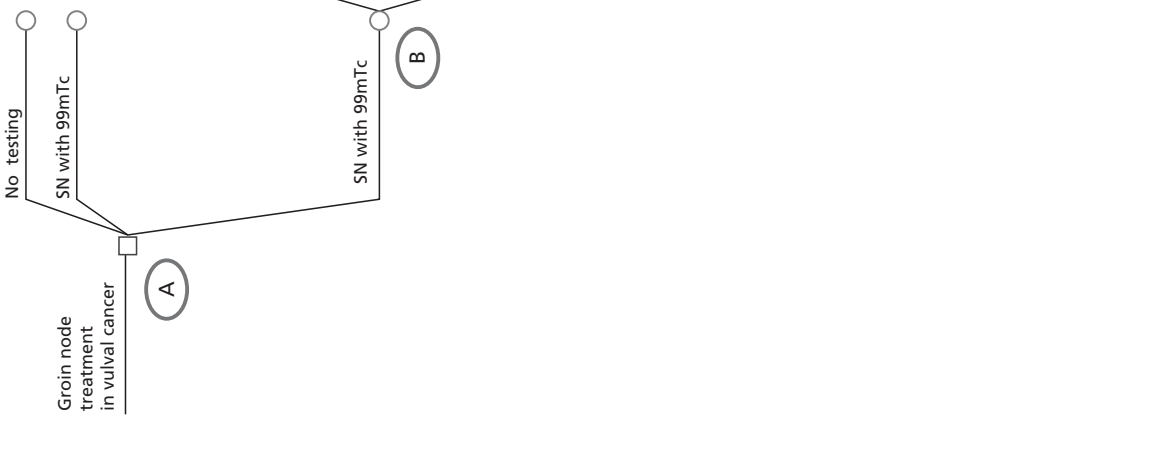


Fig 3 Branch of the decision tree model used in cost-effectiveness analysis of treatment strategies for inguinal femoral lymph nodes in women with vulval cancer.

Objectives	Plan of Research
To determine the accuracy of SLN biopsy with 99mTc enhanced and blue dye lymphoscintigraphy compared with the histopathology of inguinofemoral lymphadenectomy specimen in vulval cancer through systematic review, bivariate meta analysis and metaregression analysis.	Section 5.1
To assess through systematic review the diagnostic and therapeutic impact of lymphoscintigraphy for SLN biopsy in	Section 5.1
changing disease staging	Section 5.2
changing planned treatment	
reducing complications associated with lymphadenectomy	
improving morbidity and disease free survival	
To determine the effectiveness of various interventions (surgery, radiotherapy, chemotherapy) in the management of vulval cancer	
To evaluate the cost-effectiveness of lymphoscintigraphy directed treatment vs current treatment strategy in terms of both human and financial costs using decision-analytic modelling.	Section 5.4

The relationship of our objectives to the clinical process is shown in Fig 2.

4 Relevance to commissioning brief

The title of the HTA commissioning brief (09/112) refers to 'The value of adding 99mTc enhanced lymphoscintigraphy for SLN biopsy to current methods of staging of vulval cancer'. It goes on to include the following in the scope of the work to be carried out: effect of staging on treatment planned, reduction in the need for lymphadenectomy, decision analysis and cost effectiveness of added value of 99mTc enhanced lymphoscintigraphy compared with blue dye and current practice of inguinofemoral lymphadenopathy. From this, we take it that the scope of the work is to be broad.

We have published systematic review of the diagnostic accuracy of 99mTc enhanced and blue dye lymphoscintigraphy directed SLN biopsy in the staging of vulval cancer in 2006.[10] We shall update the search to incorporate the findings of the primary studies published in the last 5 years. The brief has specified inguinofemoral lymph node dissection as the gold standard. It also asks researchers to 'ensure data is included from the latest clinical trials and identify to what extent this will lead to a change in staging of the cancer and the subsequent treatment and quality of life of patients'. The largest and recent trial in this area, GROINSS-V, uses different reference standards - inguinofemoral lymph node dissection (if SLN positive) and follow up for groin recurrence (if SLN negative).[13] We have broadened our search and selection criteria to include studies that confirm SLN status by either inguinofemoral lymph node dissection or follow up and accounted for this variation in study design in our analysis and modelling.

In order to determine the value of SLN biopsy with 99mTc, information on diagnostic accuracy alone will not be sufficient. We will evaluate the potential of SLN biopsy in correctly identifying the sentinel node (localisation or mapping failure rates), the impact of surgeon and team's experience and skills on the accuracy, the impact of variation in disease characteristics, index test protocols and histopathological examination methods on the accuracy. In addition information on diagnostic impact, therapeutic and patient outcomes will be needed. Thus, it is crucial to review effectiveness of various interventions in patients with vulval cancer in addition to accuracy of 99mTc enhanced and blue dye lymphoscintigraphy directed SLN biopsy compared with inguinofemoral lymphadenopathy in staging of vulval cancer to inform decision analytic modelling. Through this project we will update the structure of our existing decision analytic model, will update the probability and cost input data, and will perform probabilistic sensitivity analysis.

We believe that it is feasible to undertake this work within the time scale with the resources we have requested. Our team has the necessary experience and expertise for fulfilling all the requirements in the HTA brief. We have a very strong, internationally renowned, group knowledgeable in systematic reviews of diagnostic and effectiveness data and in economic modelling. Through our MRC training fellowship project, we have the expertise to undertake and update systematic reviews on SLN status in vulval cancer.[24]

5 Plan of research

The plan of research will be to update systematic reviews of the accuracy of 99mTc enhanced and blue dye lymphoscintigraphy for sentinel lymph node biopsy in vulval cancer and to undertake systematic reviews of the effectiveness of treatments for vulval cancer. Simultaneously a previously developed decision analytic model will be refined and additional rapid systematic reviews will be undertaken as necessary to populate the emerging model.

We will address the following structured question:

Population: Women with early-stage vulval cancer

Index Tests: 99mTc enhanced and blue dye lymphoscintigraphy for SLN biopsy

Reference standard: Histopathology of inguinofemoral node dissection

Follow up for groin recurrence

Interventions: Current practice of surgery with routine inguinofemoral lymphadenectomy compared with interventions based on SLN status with or without groin node dissection; radiation, or chemotherapy

Outcomes:

- Test accuracy: Accuracy of 99mTc enhanced lymphoscintigraphy compared with blue dye in identifying potentially curable disease
- Diagnostic impact: change in staging after 99mTc enhanced lymphoscintigraphy compared with blue dye or current practice
- Therapeutic impact: change in treatment plan including avoidance of full inguinofemoral lymphadenectomy after 99mTc enhanced lymphoscintigraphy compared with blue dye and current practice by response to treatment that permits continuation or alteration of treatment or decision on clinical follow up only
- Patient outcomes: mortality, morbidity-free survival, Quality of Life
- Economic outcome: Use of resources, cost per death avoided, cost per complication free survival, cost per quality adjusted life years (costs per QALY)

Study design:

- Test accuracy studies
- Prospective cohort studies of outcomes of patients tested
- Studies investigating diagnostic and therapeutic impact with or without concurrent assessment of test accuracy.
- Randomised controlled trials and non randomised controlled studies assessing effectiveness of interventions.

- Economic evaluations

Exclusions:

- Advanced stage vulval cancer, inoperable tumours, tumours with diameter > 4 cm or those unsuitable for primary surgery
- Clinical suspicion of metastases with palpable inguinofemoral lymph nodes, enlarged lymph nodes (> 1.5 cm) on imaging or cytologically proven inguinofemoral lymph node metastases.
- Patients with multifocal tumours

Systematic reviews of test accuracy, diagnostic and therapeutic impact, and effectiveness will be updated/carried out using established methodology in line with the recommendations of the NHS Centre for Reviews and Dissemination and the Cochrane Collaboration including those of Cochrane Methods Working Group on Screening and Diagnostic tests.[27] Inclusion, data extraction and quality assessment will be carried out in duplicate with differences resolved by consensus and/or arbitration involving a third reviewer.

5.1 Reviews of test accuracy and impact of testing

Evidence on the accuracy of SLN biopsy with 99mTc and blue dye lymphoscintigraphy will be reviewed. Alongside this we will review the impact of SLN biopsy on staging and treatment in vulval cancer. Studies will be identified from a database of published and unpublished literature which will be assembled. We have published a systematic review of literature on diagnostic accuracy of SLN biopsy and have identified 24 relevant studies.[21] We will rerun our search strategy and update the accuracy review, seeking studies on diagnostic and therapeutic impact in addition.

5.1.1 Study identification and selection

Evidence on the accuracy of sentinel node biopsy using 99mTc and blue dye and their diagnostic and therapeutic impact in early vulval cancer will be identified from sensitive searches of published and unpublished sources. Language restrictions will not be applied to electronic searches. The following databases will be searched: MEDLINE, EMBASE, Science Citation Index, MEDION and Cochrane Database of Systematic Reviews (CDSR), Cochrane Central Register of Controlled Trials (CENTRAL), Database of Abstracts of Reviews of Effects (DARE) and Health Technology Assessment Database (HTA). Information on studies in progress, unpublished research or research reported in the grey literature will be sought by searching a range of relevant databases including Clinical Trials.com and UK Clinical Research Network Portfolio. A draft MEDLINE strategy is included in appendix A. Electronic searches will be supplemented by hand searching, contacting manufacturers and consultation with experts in the area. In addition authors of included studies will be contacted for information on relevant published or unpublished studies. The preliminary search undertaken to update our published accuracy review has yielded additional 142 citations with 85 on test accuracy. Citations identified by the search will be selected for inclusion in the review in a two-stage process using predefined and explicit criteria regarding populations, index tests, reference standard, outcomes and study design. These criteria have been piloted in our previous review.

5.1.2 Study quality assessment and data extraction

Methodological quality of the selected primary studies of test accuracy will be assessed based on elements of study design, conduct and analysis included in a validated assessment tool, QUADAS, which will be adapted to the topic area.[35] Existing studies of diagnostic and therapeutic impact are likely to be concurrent test accuracy studies evaluating SLN biopsy with either groin dissection or clinical follow up[13] as reference standards. There are no validated assessment tools for studies that evaluate the impact of testing. We shall adapt the QUADAS criteria to evaluate the studies on impact adjusting for test, reference standard, treatment and outcome characteristics. Data extraction will be performed using pre-designed, piloted data extraction forms, drawing on existing pro-formas used by the project team in previous, completed reviews in the topic area.[10] Missing information will be obtained from investigators if is crucial to subsequent stages of analysis and modelling. To avoid introducing bias, unpublished information will be treated in the same fashion as published information. In addition to using double data extraction to ensure the reproducibility of the overview, sensitivity analyses around important or questionable judgements regarding quality assessment and data extraction will be performed.

5.1.3 Data synthesis

Sensitivity, specificity and LRs for individual studies comparing 99mTc and blue dye with inguinofemoral lymphadenopathy will be derived.

It is anticipated that the following will be important sources of variation in test accuracy estimates:

- Population characteristics: Stage of vulval cancer, size of lesion, location of lesion, method of diagnosis of vulval cancer (clinical diagnosis, excision or punch biopsy)
- Index test characteristics: Type of sentinel node biopsy, reporting of test execution and interpretation, number, training and expertise of the persons reading and executing the test, healthcare setting (secondary or tertiary)
- Reference test: readers of histopathology of nodes from inguinofemoral lymphadenectomy blind to the index results, clinical follow up for groin recurrence in test negative patients
- Study quality: study design (prospective or case-control) and study quality (high: meeting all assessment criteria; medium: meeting at least one assessment criteria; low: meeting no quality criteria). High-quality studies will be used as the reference category to determine whether or not medium- and low-quality studies have biased estimates of test accuracy.

Based on an investigation of heterogeneity summary estimates of sensitivity, specificity and summary ROC curves will be derived using bivariate method for meta-analysis.[25,36–38] LRs are considered more clinically meaningful as measures of test accuracy and they allow estimation of probabilities for economic modelling. Post test probabilities can be used to tailor the absolute effectiveness estimates according to test results. Presence of a threshold effect will be examined by plotting sensitivity against 1- specificity in a ROC analysis and by calculating Spearman correlation coefficients.[39]

Heterogeneity of results between studies will be investigated qualitatively by examining the distribution of sensitivities and specificities in (ROC) space and variability of estimates of diagnostic odds ratios (DOR) across studies using the forest plot.[39] In addition, heterogeneity will be investigated quantitatively using meta-regression and subgroup analyses. Quantitative investigation will be undertaken based on variables defined a priori and including population characteristics, index and reference test characteristics and study quality.[40] Metaregression will allow us to test the hypothesis as to whether or not 99mTc is more accurate than blue dye. We will perform sensitivity analysis to assess the effect of study design and quality including those with differential reference standard on the overall accuracy, diagnostic and therapeutic impact. This data will be utilised in the development of the decision analytic model.

The risk of publication and related biases is expected to be high in reviews of test accuracy. Publication bias will be investigated using funnel plots of DOR against corresponding variances.[41,42] Qualitative investigation will be based on the premise that large gaps in the funnel indicate possible 'missing' publications. These omissions are usually due to small studies showing limited accuracy and are unlikely to be missing at random. Statistical investigation of publication bias will be undertaken in STATA based on templates of commands and instructions already developed by the project team.

5.2 Review of effectiveness of interventions

For evidence on the effectiveness of treatments for vulvar cancer we will begin by searching for existing systematic reviews. Any existing reviews will be examined for relevance and currency in order to inform further searching for primary studies. Existing reviews will be assessed for their quality and currency follow existing guidelines QUOROM and PRISMA.[43] Through this process we will identify gaps where reviews do not exist and where they need updating.

5.2.1 Study identification and selection

Where necessary effectiveness reviews of RCTs of treatments for vulval cancer will be undertaken following existing guidelines[27] ensuring the output complies with the QUOROM statement.[43] Searches for further primary studies will be performed. The following databases will be searched: MEDLINE, EMBASE,

Science Citation Index and the The Cochrane Library (all databases). On-going studies will be sought by searching Clinical Trials.com and the UK Clinical Research Network portfolio. Draft searches for MEDLINE are included in Appendix A. Studies will be selected for inclusion in the review in a two-stage process using predefined and explicit criteria regarding populations, interventions and outcomes using procedures similar to the ones outlined in the previous section 5.1.1.

5.2.2 Study quality assessment and data extraction

The quality of included reviews will be assessed against a validated tool and a reporting checklist, QUOROM.[43] Methodological quality of randomised and non-randomised trials will be assessed based on accepted criteria. Information on the adequacy of randomisation, sequence generation, concealment, blinding, description of withdrawals, and follow-up rates would be sought as these are elements most likely to have a direct relationship to bias in a RCT.[44] Procedures for obtaining missing information and resolving disagreements will be similar to the ones outlined in section 5.1.2.

5.2.3 Data synthesis

RevMan and STATA softwares will be used to conduct analyses. Heterogeneity of results between studies and investigation for publication bias will be statistically and graphically assessed using established methods. The decision to proceed to meta-analysis will depend on the degree of heterogeneity in the data set. It is anticipated that the following will be important sources of variation in the estimates of effectiveness:

- Population characteristics: Stage of vulval cancer, age of patient, number of lymph node metastases, morphology of the nodes (size, extracapsular involvement)
- Treatment characteristics: Type of intervention (surgery, radiotherapy or chemo radiotherapy), duration of therapy, healthcare setting (secondary or tertiary), timing of intervention
- Outcome measures: Mortality, morbidity, Quality of life

Conclusions regarding the typical estimate of an effect size of the intervention will be interpreted cautiously if there is significant heterogeneity. Where uncertainty exists, the output from data syntheses will be employed following triangulation against subjective probability estimates, judiciously in decision analytic modelling.

5.3 Eliciting subjective probabilities

In anticipation of small numbers of effectiveness studies subjective probabilities will be elicited, using a group interview, from between 10 and 15 clinical experts in the fields of gynaecological cancer and oncology with no conflict of interest in the area, identified by clinicians in the project team and project advisors. The aim of the elicitation process will be to gather subjective views about the size and probability of diagnostic and therapeutic impact of SLN biopsy using 99mTc to current practice in the staging of vulval cancer and reduction in the need for lymphadenectomy.

A face-face group interview (behavioural aggregation) will be used in preference to individual interviews as this facilitates a common understanding of the problem and task from experts and will allow us to benefit from group discussion and interaction leading to a consensus of opinion. The expert group will be facilitated by both a clinical and non-clinical expert drawn from the project team with sufficient statistical expertise to provide probabilistic training to experts, validate their results and provide feedback. The interview will take place over two half days and will, briefly, comprise:

- Training of experts (probability, probability distributions, judgement heuristics and biases)
- Practicing elicitations
- Eliciting probabilities
- Presentation of results back to experts
- Repeat elicitation of probabilities to check face validity and if necessary ensure a joint probability distribution.

Findings from the elicitation process will be triangulated with findings from the systematic reviews and probability distributions will assist with populating the decision analytic model. As well as expertise within the project team[45] we have access to experts in the field, based at the University of Birmingham.[46] Furthermore the use of this method in the HTA funded project PET-CT imaging in restaging recurrent cervical cancer (HTA No 09/29/02) will help us to develop and refine the questionnaire for this project.

5.4 Model based economic evaluation

The objective of the economic evaluation will be to compare the relative cost effectiveness of undertaking SLN biopsy compared with current practice of intervention involving inguinofemoral lymphadenectomy without testing.

5.4.1 Perspective and data collection

If SLN biopsy is shown to be an accurate and effective alternative to the standard practice in staging vulval cancer then it is likely that important cost implications will be seen for the health care sector. For example, SLN biopsy may detect additional evidence of the extent of metastasis compared with standard investigations which could increase the number and extent of subsequent tests and treatment required by the individual. But the additional costs associated with more accurate staging of the cancer may lead to a reduction in costs associated with unnecessary or ineffective subsequent treatments and also prolong the life of the woman. Thus, if available data allow, the economic evaluation will be based on an outcome of cost per QALY and/or Cost per morbidity-free 5 year survival (this latter is an outcome we have used in our previous analysis24 due to the paucity of quality of life data)/or cost per 'death due to recurrent cancer' avoided. The analysis will adopt the perspective of the NHS.

Therefore data collection required for the model based economic evaluation will at least include:

- The equipment, other resource use and costs associated with SLN biopsy
- knock on costs associated with additional further tests and treatments that are required as a result of the staging
- equipment, resource use and costs associated with current practice
- Accuracy of the SLN biopsy and current practice package compared with the accuracy of current practice tests alone
- Effectiveness of alternative intervention pathways that are followed as a result of the diagnosis
- Outcomes such as quality of life associated with vulval cancer at various disease stages

Cost data will be collected from two principal sources. First, once the clinical evidence has been synthesised into the main strategies of diagnosis and treatment, relevant studies will be examined for their data on costs and resource use. These data will be subject to relevant quality criteria. Additional cost data will be available from other sources such as the National Schedule for Reference Costs. If necessary primary cost and resource data will be collected from Pan Birmingham Cancer network and Addenbrookes Hospital to complete any gaps in the information required for the modelling process.

Additional searches will be undertaken to help populate the decision model. The Information Officer will work in close liaison with the health economist to identify the model questions. Information to answer these questions will be provided by focused searching of appropriate databases, including reference cost databases, statistical sources and other sources of relevant information. The evidence found in the clinical accuracy and effectiveness reviews will provide the majority of the parameters required to carry out the economic evaluations of alternative test and treat packages. Additional data on early complications predisposing to late complications, and psychosexual problems will be systematically obtained by searching the relevant literature. Where there is paucity of evidence, we will elicit subjective probabilities as detailed above. The costs for lymphoedema and district nurse input will be gathered by liaising with the lymphoedema service and the centres.

5.4.2 Model and analysis

The economic evaluation will involve the development of an existing decision analytic simulation model as a framework for conducting cost-effectiveness analyses.[24] The economic evaluations will inform current treatment policy in this clinical area. A modelling framework is ideally suited to demonstrate and explore the importance of the inherent uncertainty. We will develop the model including women with vulval cancer of all age groups. This will be a development of the existing model which focuses only on women 70 years and older.

An incremental approach will be adopted with a focus on additional costs and gain in benefits associated with a move away from current practice to alternative test and treatment strategies. Using discounting, adjustments will be made to reflect the differential timing of costs and outcomes in terms of the extension to the length of life extend associated with the test and treat strategies. The base-case analysis will follow Treasury recommendations for public sector projects.

5.4.3 Presentation of results and sensitivity analysis

The results of these economic analyses will be presented using cost-effectiveness acceptability curves to reflect sampling variation and uncertainties in the appropriate threshold cost-effectiveness value. We shall also use both simple and probabilistic sensitivity analyses to explore the robustness of these results to plausible variations in key assumptions and variations in the analytical methods used, and to consider the broader issue of the generalisability of the results.

In addition to probabilistic sensitivity analysis on our base-case model, we shall include a range of alternative analyses to explore the robustness of these results to plausible variations in key assumptions and variations in analysis, and to consider generalisability of the results.

6 Expertise in the team

The applicants have a wide and appropriate range of expertise in systematic reviews, gynaecological oncology, clinical pathology, clinical epidemiology, health measurement, economic evaluation, medical statistics, information science and health technology assessment.

The team (KSK, ST, SS, TR, AF) has recently been awarded a HTA grant to undertake systematic review and economic modelling of clinical effectiveness of PET-CT imaging in restaging recurrent cervical cancer (HTA No 09/29/02). KSK and TR have successfully completed many HTA projects on systematic reviews of test and treatments including systematic reviews of tests for pre-eclampsia, intrapartum rapid tests for Group B streptococcus infection and preterm labour.[47–49] In addition KSK has experience of the process of eliciting subjective probabilities. His former student and current colleague TS was awarded MRC research training fellowship to undertake systematic reviews of accuracy of tests and treatment in gynaecologic cancer including vulval cancer and for undertaking decision analytic modelling and economic evaluation.[24] KSK has also led a grant on the methodology of evaluation of tests without gold standards by the NHS Research Methodology Programme. ST has undertaken many systematic reviews on tests and treatment in with pre-eclampsia, preterm labour and epilepsy. TS has conducted systematic review of accuracy of sentinel node biopsy with tests including 99mTc and blue dye in vulval cancer. She has been awarded PhD by the University of Birmingham for her work 'Non invasive and minimally invasive diagnosis and treatment of lymphadenopathy in gynaecologic cancer. Systematic review of evidence'[24]

SS and PB are both gynaecologic oncologists involved in managing women with vulval cancer. They are members of the gynaecological cancer clinical studies group of the NCRI (National Cancer Research Institute) – the national group responsible for selecting national trials for inclusion in the NCRI portfolio and supporting and directing clinical research in gynaecological cancer. PB is a former trial group member of the large GROINSS-V prospective study investigating the role of SLN biopsy in early vulval cancer. His is also the Chief Investigator in the UK in the ongoing prospective multicentre international study GROINSS-V

II which is evaluating the safety of omitting surgery in selected SLN positive patients. PB represents the study at the vulval subgroup of the NCRI. RG is a Consultant Histopathologist involved with Pan Birmingham Cancer network and Associate Director of Birmingham Cytology Training Centre (CTC). KB is a Consultant in Nuclear Medicine and has expertise in 99m Tc and blue dye lymphoscintigraphy for SLN biopsy. AF (information specialist) has extensive experience as an information specialist in providing support to a diagnostic and effectiveness technology assessments as a member of the West Midlands Health Technology Assessment Collaboration and the Aggressive Research Intelligence Facility (ARIF) based at the University of Birmingham. She is currently working on an HTA assessing the value of PET-CT for recurrent breast cancer and her expertise in devising the search strategy and database management will be of benefit to this proposal. Hilary Jeffries is a retired Lead cancer nurse and McMillan community nurse specialist with extensive experience of interaction with patients with vulval cancer. ME is a member of the Consumer liaison group, NCRI.

7 Contribution to Collective Research Effort

This systematic review on the value of SLN biopsy with 99mTc enhanced lymphoscintigraphy compared with blue dye and the current management of groin node dissection in vulval cancer and the cost effectiveness analysis of SLN biopsy using the above methods in comparison to current management fits comfortably with previously published HTA evaluations of sentinel node biopsy in other cancers. This research application complements existing National cancer research network portfolio research in gynaecological cancer. The ongoing GROINSS-V II study evaluates if inguinofemoral lymph node dissection can be omitted in the presence of a positive SLN and the treatment of groins with radiotherapy or chemoradiotherapy instead. This project will augment the current published evidence acquired through the MRC training fellowship on SLN biopsy in vulval cancer through update of the review and further comprehensive development of the decision analytic model.[24]

Due to the multiple methods employed by the proposed evidence synthesis the project team expect that the outputs of the work would be of interest to a broad research and clinical community including experts in the areas of evidence synthesis and in particular synthesis of test accuracy, gynaecological cancer, and decision making. Outputs would be submitted for presentation at national and international conferences such as Health Technology Assessment international, Medical Decision Making, European Society of Gynaecological Oncology (EGSO) and Society of Gynaecological Oncology (SGO). Similarly the outputs of this work would be of interest to a variety of peer reviewed journals and the project team would aim for a minimum of 3 peer reviewed publications in addition to publication as an HTA monograph. The project team have involved members of the NCRI consumer liaison group and VACO (Vulva Awareness Campaign Organisation), an international support group dedicated to women with vulval cancer. Users will be represented in study conduct and planning of dissemination strategies. The team will benefit from the HJ in an advisory role who has worked closely with women with vulval cancer. She has recently published a qualitative study on the experiences of women with vulval cancer as part of her PhD.[50] Experience from previous research conducted by the team has already indicated that publication and dissemination needs careful consideration from the outset.[47–49] Publication strategy will also need to anticipate early the need for versions of the report, which can be, used by women themselves. For this we will seek input from relevant consumers.

8 Details about any related (planned or active) grants held by members of the research team

KSK (as supervisor) and TS were awarded MRC research training fellowship to undertake systematic reviews of accuracy of tests in gynaecologic cancer including vulval cancer. The resulting PhD has recently been awarded.[24] Information from the accuracy of tests in vulval cancer will be updated and the analysis will be refined. We have also developed a decision analytic modelling structure for tests in vulval cancer

that will be improved upon. ST, KK, TR and SS have been successful in obtaining HTA grant for conducting systematic review and developing an economic model to evaluate the clinical effectiveness of PET CT in recurrent cervical cancer. (HTA No. 09/29/02) SS has a PhD student funded by the department of Health investigating the epigenetic changes induced by HPV in cervical cancer. PB holds a CRUK award to support the ongoing GROINSS-V II study. The expertise of SS and PB in gynaecologic oncology will be of use in providing subjective probabilistic estimates for test accuracy and effectiveness. AF (information specialist) is currently working on an HTA assessing the value of PET-CT for recurrent breast cancer and her expertise in devising the search strategy and database management will be of benefit to this proposal.

9 Summary for the non expert

Vulval cancer accounts for approximately 3–5% of all gynaecological cancers. In the UK, the lifetime risk of developing vulval cancer is 1 in 316. Although the peak incidence of this cancer is in the 7th decade, the proportion of women diagnosed with vulval cancer under the age of 50 has risen from 6% in 1975 to 15% in 2006. Vulval cancer is curable when diagnosed in an early stage. The current treatment for early-stage vulval cancer is extensive removal of the vulval tumour and excision of the groin nodes to check for spread of cancer. Cancer in the groin nodes has been identified as the single most important factor in predicting survival. Removal of the groin nodes is associated with complications in the short term (infection, wound breakdown) and long term (lymphoedema, cellulitis, sexual dysfunction) with significant negative impact on the Quality of Life. Only about a third of patients with operable disease have nodal spread and the rest are unlikely to benefit from routine removal of groin nodes. Despite the risk of significant complications and low probability of cancer spread to the nodes, groin nodes are routinely removed as missed cancer in the groin nodes is nearly always fatal. A test that could accurately identify those patients in whom cancer has spread to the groin nodes without extensive removal of all groin would be extremely valuable. There are several methods to check for involvement of the groin nodes, but none are routinely used in clinical practice. This HTA proposal focuses on the value of testing the groin node (sentinel node) with biopsy by locating them with radioactive substance (99m Technetium) or blue dye.

The sentinel node (SLN) refers to any node that receives lymphatic drainage directly from the vulval tumour and therefore has the highest probability of containing cancer cells. If the sentinel node is free from cancer, the rest of the groin should be at least risk of having spread of cancer. Identification and removal of the SLN(s) avoids the significant complications associated with complete groin dissection. In those patients negative for cancer in the sentinel nodes where extensive groin dissection was omitted, studies suggest that the risk of future disease in the groin is low, the survival rate is excellent (97% at 3 years) and that there are few complications associated with this smaller operation. If the SLN is negative no further treatment is therefore required and the patient will be followed up in the clinic.

There is a need to systematically review the accuracy of SLN biopsy with 99mTc and blue dye in identifying the spread of cancer to the groins. Moreover it is important to review how testing of SLN will have an impact on the extent of cancer spread, treatment decisions, clinical and cost outcomes.

For the proposed project our objectives are as follows:

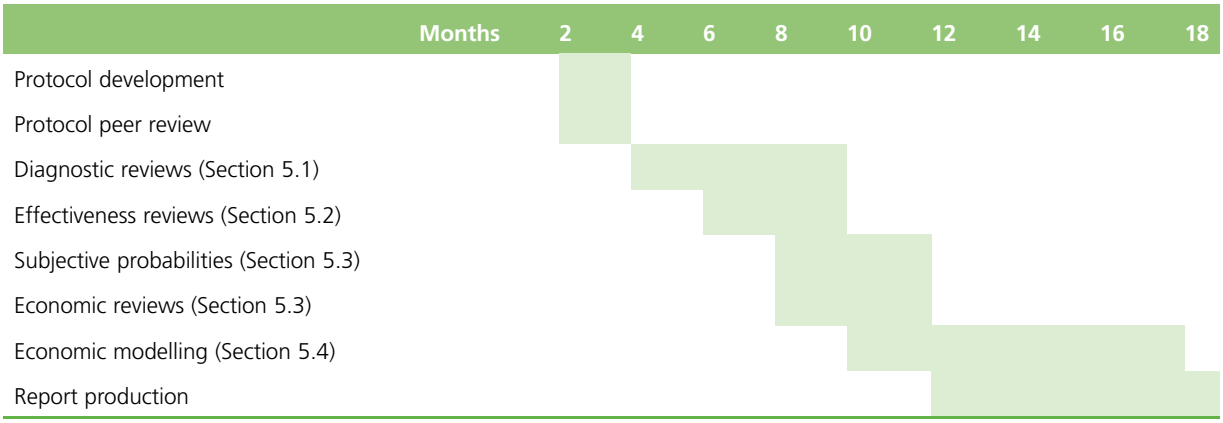
- In women who have been diagnosed to have vulval cancer, to systematically review the literature
- To assess if SLN biopsy with 99mTc or blue dye can accurately diagnose spread of cancer to the groins compared with current practice of routine extensive removal of all groin nodes
- To evaluate if the use of SLN biopsy results in change in (re)staging i.e. extent of disease compared with current practice
- To assess the impact of performing SLN biopsy on the typical standard treatment
- To summarise the effectiveness of available treatments in women with vulval cancer
- To estimate the impact of SLN biopsy results on patient outcomes and the costs associated with its routine use in this patient group.

We plan to fulfil the above objectives by systematically identifying the available evidence on the diagnostic accuracy of SLN biopsy with 99mTc and blue dye in vulval cancer compared with the accuracy of existing practice of groin dissection used in this patients group and the effectiveness of treatments for vulval cancer. The evidence found will be used in an economic evaluation comparing existing testing and treatment strategies with SLN biopsy guided treatment strategies. This evaluation will inform current treatment policy in this clinical area and highlight future research need.

10 Project Timetable and Milestones

Fig 3 shows the project timetable and milestones for the accuracy and effectiveness reviews and economic modelling. We have carefully evaluated the ongoing work and the level of staffing within our departments and feel that we would be able to commence the work in Sep 2010 for a period of 18 months, if funded.

Fig 3 Timetable



11 Justification for the support required

Staff:
Supervisor, also providing support for researchers, for example: double data extraction, assisting with inclusion decisions and being the lead for producing the final report – 1 day per week for the duration of the project.

Researcher to perform systematic review of accuracy and effectiveness studies and to identify additional epidemiological and background information for input into the modelling exercise – 1 wte for 18 months.

Health economist to perform systematic review of cost-effectiveness literature and modelling – 1wte for 12 months.

Information support for searching and document retrieval –20 days.

Equipment and consumables:
two standard specification computers, printing cartridges, paper and photocopying,

telephone and fax calls, postage,

estimated 200 interlibrary loans.

Support:

- Meeting room, refreshments and travel for the project team and consultants based on 4 face to face meetings over 12 months.
- Meeting room, refreshments and travel for consumer group representatives
- Administrative support, for steering group and preparation of final report –10 days over 12 months.

We are in an excellent position to gauge the level of resources required to deliver this type of project (systematic review and cost-effectiveness analysis) with several years experience in their delivery. We are able to draw on additional in-house expertise if necessary. Travel costs have included the cost of travel of experts and non experts for obtaining probabilistic estimates.

Appendix A Search strategy for electronic database identification of diagnostic studies for preoperative tests of lymph node status and therapeutic studies of interventions in vulval cancer

Test accuracy search – proposed MEDLINE strategy

Ovid MEDLINE(R) 1950 to November Week 3 2009

1. technetium.tw. (12014)
2. (radionuclide adj imag$).tw. (1508)
3. technetium/ (18834)
4. radionuclide imaging/ (23714)
5. sentinel lymph node biopsy/ (5545)
6. 99m tc$.tw. (3979)
7. 99mtc$.tw. (16283)
8. (sentinel adj2 lymph adj2 node$).tw. (4422)
9. or/1-8 (62808)
10. vulvar neoplasms/ (6493)
11. ((vulva or vulval or vulvar) adj5 (cancer$ or carcinoma$ or adenocarcinoma$ or carcinogen$ or sarcoma$ or malignan$ or tumo?r$ or neoplas$)).tw. (4197)
12. or/10-11 (7219)
13. 9 and 12 (192)

Effectiveness search (systematic reviews) – proposed MEDLINE strategy

Ovid MEDLINE(R) 1950 to November Week 3 2009>

1. vulvar neoplasms/ (6493)
2. (vulva or vulval or vulvar) adj5 (cancer$ or carcinoma$ or adenocarcinoma$ or carcinogen$ or sarcoma$ or malignan$ or tumo?r$ or neoplas$)).tw. (4197)
3. or/1-2 (7219)
4. limit 3 to "reviews (specificity)" (32)

Effectiveness search (RCTs) – proposed MEDLINE strategy

Ovid MEDLINE(R) 1950 to November Week 3 2009>

1. vulvar neoplasms/ (6493)
2. ((vulva or vulval or vulvar) adj5 (cancer$ or carcinoma$ or adenocarcinoma$ or carcinogen$ or sarcoma $ or malignan$ or tumo?r$ or neoplas$)).tw. (4197)
3. or/1-2 (7219)
4. limit 3 to "therapy (optimized)" (91)

Reference List

1. Hacker NF. Vulvar Cancer. In Berek JS, Hacker NF. *Practical Gynecologic Oncology 2005*. Philadelphia, PA; Williams and Wilkins; 2005.
2. URL: Http://cancerresearchuk.org/cancerstats/types/vulva/incidence/index.htm (accessed 2009).
3. Darling JR, Sherman JH. Cancers of the vulva and vagina. In Schottenfeld D, Fraumeni J Jr. (eds) *Cancer epidemiology and prevention*; 1996.
4. Woolas RP, Shepherd JH. Current developments in the management of vulval carcinoma. In O'Brien PMS (ed). *The yearbook of obstetrics and gynecology*; 1999.
5. Jones RW, Baranyai J, Stables S. Trends in squamous cell carcinoma of the vulva: the influence of vulvar intraepithelial neoplasia. *Obstet Gynecol* 1997;**90**:448–52.
6. Joura EA, Losch A, Haider-Angeler MG, *et al.* Trends in vulvar neoplasia. Increased incidence of vulvar intraepithelial neoplasia and squamous cell carcinoma of the vulva in young women. *J Reprod Med* 2000;**45**:613–5.
7. de Hullu JA, van der Avoort IA, Oonk MH, van der Zee AG. Management of vulvar cancers. *Eur J Surg Oncol* 2006;**32**:825–31.
8. RCOG Working Party report. *Management of vulval cancer* ;1999.
9. Homesley HD, Bundy BN, Sedlis A, *et al.* Assessment of current International Federation of Gynecology and Obstetrics staging of vulvar carcinoma relative to prognostic factors for survival (a Gynecologic Oncology Group Study). *Am J OG* 1991;**164**:997–1003.
10. Selman TJ, Luesley DM, Acheson N, Khan KS, Mann CH. A systematic review of the accuracy of diagnostic tests for inguinal lymph node status in vulvar cancer. *Gynecol Oncol* 2005;**99**:206–14.
11. Stehman FB, Ali S, DiSaia PJ. Node count and groin recurrence in early vulvar cancer: A Gynecologic Oncology group study. *Gynecol Oncol* 2009;**113**:52–6.
12. de hullu JA, Doting E, Piers DA, *et al.* Sentinel lymph node identification with technetium 99m labelled nanocolloid in squamous cell cancer of the vulva. *J Nucl Med* 1998;**39**:1381–5.
13. van der Zee AG, Oonk MH, de Hullu JA, *et al.* Sentinel node dissection is safe in the treatment of early stage vulvar cancer. *J Clin Onc* 2008;**26**:884–9.
14. Barton DP. The prevention and management of treatment related morbidity in vulvar cancer. *Best Pract Res Clin Obstet Gynaecol* 2003;**17**(4):683–701.
15. Homesley H, Bundy B, Sedlis A. Prognostic factors for groin node metastasis in squamous cell carcinoma of the vulva. *Gynecol Oncol* 1993;**49**:279–83.
16. Sedlis A, Homesley H, Bundy B. Positive groin lymph nodes in superficial squamous vulvar cancer. *Am J OG* 1987;**156**:1159–64.

17. Thompson JF, Uren RF. What is a 'sentinel' lymph node? *Eur J Surg Oncol* 2000;**26**:103–4.

18. Terada KY, Shimizu DM, Wong JH. Sentinel node dissection and ultrastaging in squamous cell cancer of the vulva. *Gynecol Oncol* 2000;**76**:40–4.

19. Blake Cady. Sentinel lymph node procedure in squamous cell carcinoma of the vulva. *J Clin Onc* 2000;**18**:2795–7.

20. Medical Services Advisory Committee assessment report. *Sentinel lymph node biopsy in breast cancer*. MSAC reference 1065; 2005.

21. de Hullu JA. Van der Zee AG. Groin surgery and the sentinel lymph node. *Best Prac Res Clin Obstet Gynaecol* 2003;**17**:571–89.

22. de Hullu JA, Hollema H, Piers DA, *et al.* Sentinel lymph node procedure is highly accurate in squamous cell carcinoma of the vulva. *J Clin Onc* 2000;**18**:2811–6.

23. Morton DL, Thompson JF, Essner R, *et al.* Validation of the accuracy of intraoperative lymphatic mapping and sentinel lymphadenectomy for early stage melanoma: A multicenter trial. – Multicenter selective lymphadenectomy trial group. *Ann surg* 1999;**230**:453–63.

24. Selman TJ. *Non-invasive and minimally invasive diagnosis and therapy of lymphadenopathy in gynaecological cancers. Systematic reviews of the evidence*. PhD thesis. University of Birmingham; 2009.

25. Rietsma JB, Glas AS, Rutjes AW, *et al.* Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J clin Epidemiol* 2005;**58**:982–90.

26. Cochrane methods. *Working group on systematic reviews of screening and diagnostic tests: recommended methods*. 6-6; 1996.

27. Khan KS, ter Riet G, Glanville J, Soweden AJ, Kleijnen J. *Undertaking systematic reviews of research on effectiveness. CRD's guidance for carrying out or commissioning reviews*. CRD Report Number 4. 2nd edn. York: NHS Centre for Reviews and dissemination, University of York; 2001.

28. Fotiou SK, Tserkezoglou AJ, Fragakis G, Terzakis E, Stavrakakis E, Apostolikas N. "Butterfly" operation vs triple incision technique in vulvar cancer: a comparison of morbidity and clinical outcome. *Eur J Gynaecol Oncol* 1996;**17**(1):67–73.

29. Lin JY, DuBeshter B, Angel C, Dvoretsky PM. Morbidity and recurrence with modifications of radical vulvectomy and groin dissection. *Gynecol Oncol* 1992;**47**(1):80–6.

30. Beesley V, Janda M, Eakin E, Obermair A, Battistutta D. Lymphedema after gynecological cancer treatment: prevalence, correlates and supportive care needs. *Cancer* 2007;**109**(12):2607–14.

31. Gould N, Kamelle S, Tillmanns T, Scribner D, Gold M, Walker J, *et al.* Predictors of complications after inguinal lymphadenectomy. *Gynecol Oncol* 2001;**82**(2):329–32.

32. Ryan M, Stainton MC, Slaytor EK. Aetiology and prevalence of lower limb lymphoedema following treatment for gynecological cancer. *Aust NZ J Obstet Gynaecol* 2003;**43**:148–51.

33. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. *Med Dec Making* 1991;**11**:88–94.

34. *CRD's guidance for undertaking reviews in healthcare*. *Systematic reviews*. Centre for Reviews and dissemination; 2009.

35. Whiting P, Rutjes AWS, 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* 2003;**3**(25).

36. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JAC. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 2006;**1**:1–21.

37. van Houwelingen HC, Zwinderman AH, Stijnen T. A bivariate approach to meta-analysis. *Stat Med* 1993;**12**:2273–84.

38. van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Stat Med* 2002;**28**:589–624.

39. Irwig L, Tostesen A, Gatsonis C. Guidelines for meta-analysis evaluating diagnostic tests. *Ann Intern Med* 1994;**120**:667–76.

40. Lijmer JG, Bossuyt PM, Heistercamp SH. Exploring sources of heterogeneity in systematic reviews of diagnostic tests. *Stat Med* 2002;**21**:1525–37.

41. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005;**58**:882–93.

42. Song FJ, Khan KS, Dinnes J, Sutton AJ. Asymetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *Int J Epidemiol* 2002;**31**:88–95.

43. Moher D, Cook DJ, Eastwood S, *et al.* Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Br J Surg* 2000;**87**:1448–54.

44. Jadad AR, Moore RA, *et al.* Assessing the quality of reports of randomised clinical trials: is blinding necessary? *Control Clin Trials* 1996;**17**:1–12.

45. Lathe PM, Braunholtz DA, Hills RK, Khan KS, Lilford R. Measurement of beliefs about effectiveness of laparoscopic uterosacral nerve ablation. *BJOG* 2005;**112**:243–6.

46. Lilford R. Formal measurements of clinical uncertainty: prelude to a trial in perinatal medicine. The Fetal Compromise Group. *BMJ* 1994;**308**:111–2.

47. Honest H, Forbes CA, Duree KH, *et al.* Screening to prevent spontaneous preterm birth: systematic reviews of accuracy and effectiveness literature with economic modelling. *Health Technol Assess* 2010;**13**:1–627.

48. Meads CA, Cnossen JS, Meher S, Juarez-Garcia A, ter Riet G, Duley L, *et al.* Methods of prediction and prevention of pre-eclampsia-systematic reviews of accuracy and effectiveness literature with economic modelling. *Health Technol Assess* 2008;**12**(6).

49. Daniels J, Gray J, Pattison H, Roberts T, *et al.* Rapid testing for group B streptococcus during labour: a test accuracy study with evaluation of acceptability and cost-effectiveness. *Health Technol Assess* 2009;**13**:1–178.

50. Jeffries H. Searching the lived experience of women with cancer of the vulva. *Cancer Nurse* 2009;**32**:E30–6.

Appendix 2 Scoping searches for systematic reviews and Health Technology Assessments

Systematic review

MEDLINE (14 January 2011)

1.	exp Vulval Neoplasms	6177
2.	((vulva or vulval or vulval) adj5 (cancer$ or carcinoma$ or adenocarcinoma$ or carcinogen$ or sarcoma$ or malignan$ or tumo?r$ or neoplas$)).mp. [mp = title, original title, abstract, name of substance word, subject heading word, unique identifier]	6992
3.	#1 or #2	6992
4.	("review" or "review academic" or "review tutorial").pt.	1,547,865
5.	cinahl.tw,sh.	4605
6.	((hand adj2 search$) or (manual$ adj2 search$)).tw,sh.	4358
7.	(electronic database$ or bibliographic database$ or computeri?ed database$ or online database$).tw,sh.	7263
8.	(pooling or pooled or mantel haenszel).tw,sh.	34,931
9.	(retraction of publication or retracted publication).pt.	3069
10.	(peto or dersimonian or der simonian or fixed effect).tw,sh.	1811
11.	(medline or medlars or embase or pubmed).tw,sh.	43,864
12.	#5 or #6 or #7 or #8 or #9 or #10 or #11	84,639
13.	#4 and #12	38,102
14.	meta-analysis.sh.	25,963
15.	(meta-analys$ or meta analys$ or metaanalys$).tw,sh.	45,017
16.	(systematic$ adj5 review$).tw,sh.	26,728
17.	(systematic$ adj5 overview$).tw,sh.	531
18.	(quantitativ$ adj5 overview$).tw,sh.	129
19.	(quantitativ$ adj5 synthesis$).tw,sh.	949
20.	(methodologic$ adj5 review$).tw,sh.	2168
21.	(methodologic$ adj5 overview$).tw,sh.	147
22.	(integrative research review$ or research integration).tw.	67
23.	(quantitativ$ adj5 review$).tw,sh.	2941
24.	#14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23	67,802
25.	#13 or #24	89,293
26.	#3 and #25	43

EMBASE (14 January 2011)

1.	'vulva tumour'/exp	6733
2.	((vulva OR vulval OR vulval) NEAR/5 (cancer* OR carcinoma* OR adenocarcinoma* OR carcinogen* OR sarcoma* OR malignan* OR tumo?r* OR neoplas*)):lnk,ab,ti	4922
3.	#1 OR #2	7756
4.	'review'/exp	1,696,863
5.	medline:lnk,ab,ti OR medlars:lnk,ab,ti OR embase:lnk,ab,ti OR pubmed:lnk,ab,ti	53,845
6.	cinahl:lnk,ab,ti	5675
7.	electronic:ab,ti AND adj:ab,ti AND database*:ab,ti OR (bibliographic NEAR/2 database*):ab,ti	1169
8.	(pooled NEAR/2 analys*):ab,ti OR pooling: ab,ti	11,058
9.	peto:ab,ti OR dersimonian:ab,ti OR (fixed NEAR/2 effect):ab,ti OR mantel:ab,ti AND haenszel:ab,ti	2391
10.	#5 OR #6 OR #7 OR #8 OR #9	66,370
11.	#4 AND #10	34,213
12.	'meta-analysis'/exp	51,889
13.	meta AND analys*:lnk,ab,ti	48,252
14.	(systematic* NEAR/5 review*):lnk,ab,ti	34,901
15.	(systematic* NEAR/5 overview*):lnk,ab,ti	667
16.	(quantitativ* NEAR/5 review*):lnk,ab,ti	1906
17.	(quantitativ* NEAR/5 overview*):lnk,ab,ti	162
18.	(methodologic* NEAR/5 review*):lnk,ab,ti	2743
19.	(methodologic* NEAR/5 overview*):lnk,ab,ti	188
20.	(integrative NEAR/5 (research OR review*)):ab,ti OR (research NEAR/5 integration):ab,ti	2019
21.	(quantitativ* NEAR/5 synthesi*):lnk,ab,ti	1437
22.	#12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21	102,029
23.	#11 OR #22	121,795
24.	#3 AND #23	63
25.	#3 AND #23 AND [embase]/lim	54

Cochrane (14 January 2011)

1.	MeSH descriptor Vulval Neoplasms explode all trees	53
2.	(vulva OR vulval OR vulval) NEAR/5 (cancer* OR carcinoma* OR adenocarcinoma *OR carcinogen* OR sarcoma* OR malignan* OR tumo?r* OR neoplas*)	116
3.	(#1 OR #2)	116

Appendix 3 Data extraction form for diagnostic reviews

Review title

Date (dd/mm/yy)

Reviewer ID

Study ID

Study title	
First author	
Source of publication Journal yy;vol(iss):pp	
Country of publication	
Language	
Publication type	journal □ abstract other (specify): ……

Study eligibility

Population	women suspected to have primary vulval cancer (in stage I or/and II FIGO stage or adequate in TNM classiffication) without clinically suggestive of metastasis patients without primary tumour treatment other
Index test	Tcc Tcc+blue dye blue dye other
Reference standard	histopatology (IFL) clinical follow-up other

If "other" was selected one or more times → discard, if it was not → continue.

Study characteristics

Population	
Trial inclusion criteria	
Trial exclusion criteria	
Number of enrolled patients, N	
Number of patients who completed the study, n (%)	
Age, in years specify the measure:	
Type of initial treatment, n (%)	
Initial staging, n (%)	
Other main baseline parameters (SLN status)	
Tests	
Type of index test used (short description)	
Type of alternative test/comparator (short description)	
Type of reference standard (short description)	
Duration of follow up	

Methods	
Method of enrolment	consecutive arbitrary random not reported
Data Collection	prospective retrospective not reported
Information about drops out	precise information inaccurate information lack of information
Statistical technique used	
Sample size calculation	
Funding source	

Quality assessment	
Representative spectrum	Yes No Unclear
Acceptable reference standard	Yes No Unclear Not applicable
Acceptable delay between tests	Yes No Unclear
Partial verification avoided	Yes No Unclear
Differential verification avoided	Yes No Unclear
Incorporation avoided	Yes No Unclear
Reference standard/ index test results blinded	Yes No Unclear
Relevant clinical information	Yes No Unclear
Uninterpretable results reported	Yes No Unclear
Withdrawals explained	Yes No Unclear

Results

Test			
	Disease present	Disease absent	Total
Test positive			
Test negative			
Total			
Comparator			
	Disease present	Disease absent	Total
Test positive			
Test negative			
Total			

Dichotomous data

Outcome:…… Follow up:……	
Intervention group	
N	n (%)
(95% CI SE p)	
Incomplete outcome data addressed	

N' – number of evaluated patients; n – number of patients with outcome

Reviewer's comments

Appendix 4 Data extraction form for effectiveness reviews

Review title
Date (dd/mm/yy)
Reviewer ID
Study ID

Study title	
First author	
Source of publication Journal yy;vol(iss):pp	
Language	
Publication type	journal □ abstract other (specify): ……

Study eligibility/PICOS Scheme

Population	women with primary vulval cancer in FIGO stage I or/and II, according FIGO classiffication women with primary vulval cancer in T1-2,N0-1,M0 stage, according TNM classiffication other
Intervention	Curative intent: radical vulvectomy modyfied vulvectomy radiation chemotherapy
Comparison	no comparators comparators used (specify)…… comparison within the same group of participants over time
Outcomes	morbidity mortality Quality of Life none of the above
Study design	RCT non-randomised controlled study (specify): …….. other (specify): ……..

If included study is comparative experimental study, then go to the point A ,
If included study is comparative observational study, then go to the point B,
If included study is non- comparative study, then go to the point C

PART A
Comparative Experimental Studies:

1. Study characteristics

Methods/methodological quality	
Study design	RCT NRS
RCT	
Method of randomization	specify and assess the method: …… …. adequate inadequate unclear not reported
Allocation concealment	adequate inadequate unclear not reported Describe……..
Blinding	select blinded subjects: patients investigators/clinicians outcomes assessors no blinding used assess the method: adequate inadequate unclear not reported
Information about drop outs	precise information (number of patients and reasons) inaccurate information lack of information
Rate of loss to follow-up	
Patients lost to follow-up analysed for adverse events	
Was the follow-up adequate to ascertain adverse effects?	Yes No Unclear If "yes", specify……
Statistical technique used	
Was adequate statistical analysis of potential confounders performed?	Yes No Unclear
Intention-to-treat analysis What was the definition of ITT in the study?	implemented not implemented …… …… ……
Sample size calculation	
Was the sensitivity analysis performed?	Yes No Not applicable
How problem with missing data was resolved?	
Were missing data accounted for in the analyses?	Yes No
Post hoc analysis	
Funding source	

NRS	
Control group selection	specify and assess the method: ……... ... adequate inadequate unclear not reported
Allocation concealment	adequate inadequate unclear not reported Describe……..
Blinding	select blinded subjects: patients investigators/clinicians outcomes assessors no blinding used assess the method: adequate inadequate unclear not reported
Information about drop outs	precise information (number of patients and reasons) inaccurate information lack of information
Rate of loss to follow-up	
Patients lost to follow-up analysed for adverse events	
Was the follow-up adequate to ascertain adverse effects?	Yes No Unclear If "yes", specify……
Statistical technique used	
Was adequate statistical analysis of potential confounders performed?	Yes No Unclear
Intention-to-treat analysis What was the definition of ITT in the study?	implemented not implemented …… …… ……
Sample size calculation	
Was the sensitivity analysis performed?	Yes No Not applicable
How problem with missing data was resolved?	
Were missing data accounted for in the analyses?	Yes No
Post hoc analysis	
Funding source	

Population		
Trial inclusion criteria		
Trial exclusion criteria		
	Intervention group	Comparator/control group
Number of enrolled patients		
Number of patients randomised, NR		
Number of patients who completed treatment, n (%)		
Number of patients available for follow up, n (%)		
Age, in years specify the measure:		

Other baseline characteristics (FIGO or TNM stage, tumour size, deep of invasion, tumour cell type, site of disease)		
Were treatment groups comparable at baseline?	Yes No If "no" specify the reasons: …… …… …… ……	
Treatment		
Type of treatment used (technique, no. of sessions)		
Treatment duration		
Duration of follow up		
Outcomes		
Definition and unit of measurement		

Results

Dichotomous data

Outcome:…… Follow up:……			
Intervention group NR / N =		Control group NR / N =	
N’	n (%)	N’	n (%)
Effect estimate RR OR (95% CI SE p)			
Blinding	select blinded subjects: patients investigators/clinicians outcomes assessors no binding used assess the method: adequate inadequate unclear not reported		
Incomplete outcome data addressed			

N’ – number of evaluated patients; n – number of patients with outcome

Time-to-event data

Outcome:…… Follow up:……			
Intervention group NR / N =		Control group NR / N =	
N’	Median	N’	Median
Effect estimate HR (95% CI SE p)			
Blinding	select blinded subjects: patients investigators/clinicians outcomes assessors no binding used assess the method: adequate inadequate unclear not reported		
Incomplete outcome data addressed			

N’ – number of evaluated patients

Continuous data

<table>
<tr><td colspan="8">Outcome:…… Follow up:……</td></tr>
<tr><td colspan="4">Intervention group
NR / N =</td><td colspan="4">Control group
NR / N =</td></tr>
<tr><td>N’</td><td>Mean value at baseline (SD / SE / other)</td><td>Mean endpoint value (SD / SE / other)</td><td>Mean change from baseline (SD / SE / other)</td><td>N’</td><td>Mean value at baseline (SD / SE / other)</td><td>Mean endpoint value (SD / SE / other)</td><td>Mean change from baseline (SD / SE / other)</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Blinding</td><td colspan="7">select blinded subjects:
patients investigators/clinicians outcomes assessors no binding used
assess the method:
adequate inadequate unclear not reported</td></tr>
<tr><td colspan="8">Incomplete outcome data addressed</td></tr>
</table>

N’ – number of evaluated patients

Reviewer's comments

PART B

B) Comparative Observational Studies:

1. Study characteristics

Methods/methodological quality	
Study design	Case – control Cohort
Case – Control	
Is case definition adequate?	independent validation record linkage self reported none
Are the cases representative?	all cases arising from same population or group not known
Selection of controls	same population as cases not known or no
Definition of controls	outcome of interest not present in history no mention of history of outcome
Comparability of cases and controls	Yes No Unclear
Ascertainment of exposure to intervention	secure record structured interview where blind to case/control status interview not blinded to case/control status written self report of medical record only no description
Was the method of ascertainment of exposure for cases and controls the same?	Yes No Unclear
Non-response rate	same for both groups non respondents described rate different and no designation
Cohort	
Is the cohort representative	Yes No Unclear
Selection of non–exposed cohort	same population as exposed cohort not known or no
Ascertainment of exposure	secure record structured interview written self report no description
Demonstration that outcome of interest wasn't present at start of study?	Yes No Unclear
Comparability of cohorts on the basis of the design or analysis	Yes No Unclear
Assessment of outcome	independent or blind assessment record linkage self-report no description
Was follow-up long enough for outcomes to occur?	Yes No Unclear If "yes", specify……
Was follow-up of cohorts adequate?	complete follow-up subjects lost to follow-up unlikely to introduce bias, small number lost (…..%) follow-up rate ….%, and no description of this lost no statement
Are the objectives or the hypothesis of the study stated?	Yes No Unclear
Method of allocation to groups	
For patients who weren't eligible for study, are the reasons why stated?	Yes No
Information about drop outs	precise information (number of patients and reasons)

	inaccurate information lack of information	
Statistical technique used		
Intention-to-treat analysis What was the definition of ITT in the study?	implemented not implemented …… …… ……	
Sample size calculation		
Was loss to follow-up taken into account in the analysis?	Yes No	
Were any confounders mentioned?	Yes, please describe…….. No	
Were confounders accounted for in analyses?	Yes No	
How problem with missing data was resolved?		
Were missing data accounted for in the analyses?	Yes No	
Was the impact of biases assessed?	Yes No Not clearly assessed	
Funding source		
Population		
Trial inclusion criteria		
Trial exclusion criteria		
Is the target population defined?	Yes No	
	Intervention group	Comparator/control group
Number of included patients, N		
Number of patients who completed treatment, n (%)		
Age, in years specify the measure:		
Other baseline characteristics (FIGO or TNM stage, tumour size, deep of invasion, tumour cell type, site of disease)		
Were treatment groups comparable at baseline?	Yes No Not applicable If "no" specify the reasons: …… …… …… ……	
Treatment		
Type of treatment used (technique, no. of sessions)		
Treatment duration		
Duration of follow up		
Outcomes		
Definition and unit of measurement		

Results

Dichotomous data

Outcome:…… Follow up:……			
Intervention group NR / N =		Control group NR / N =	
N’	n (%)	N’	n (%)
Effect estimate RR OR (95% CI SE p)			
Blinding	select blinded subjects: patients investigators/clinicians outcomes assessors no binding used assess the method: adequate inadequate unclear not reported		
Incomplete outcome data addressed			

N’ – number of evaluated patients; n – number of patients with outcome

Time-to-event data

Outcome:…… Follow up:……			
Intervention group NR / N =		Control group NR / N =	
N’	Median	N’	Median
Effect estimate HR (95% CI SE p)			
Blinding	select blinded subjects: patients investigators/clinicians outcomes assessors no binding used assess the method: adequate inadequate unclear not reported		
Incomplete outcome data addressed			

N’ – number of evaluated patients

Continuous data

Outcome:…… Follow up:……							
Intervention group NR / N =				Control group NR / N =			
N'	Mean value at baseline (SD / SE / other)	Mean endpoint value (SD / SE / other)	Mean change from baseline (SD / SE / other)	N'	Mean value at baseline (SD / SE / other)	Mean endpoint value (SD / SE / other)	Mean change from baseline (SD / SE / other)
Blinding	select blinded subjects: patients investigators/clinicians outcomes assessors no binding used assess the method: adequate inadequate unclear not reported						
Incomplete outcome data addressed							

N' – number of evaluated patients

Reviewer's comments

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

PART C

Non-Comparative Studies:

Quality assessment according checklist from "Methods for the development of NICE public health guidance (second edition)"

Type of study……...

Methodology description……

Population	
Trial inclusion criteria	
Trial exclusion criteria	
Number of enrolled patients	
Number of patients who completed treatment, n (%)	
Number of patients available for follow up, n (%)	
Age, in years specify the measure:	
Other baseline characteristics (FIGO or TNM stage, tumour size, deep of invasion, tumour cell type, site of disease)	
Treatment	
Type of treatment used (technique, no. of sessions)	
Treatment duration	
Duration of follow up	
Outcomes	
Definition and unit of measurement	

Results

Dichotomous data

Outcome:…… Follow up:……	
Intervention group	
N	n (%)
(95% CI SE p)	
Incomplete outcome data addressed	

N' – number of evaluated patients; n – number of patients with outcome

Time to event data

Outcome:…… Follow up:……	
Intervention group	
N	Median
(95% CI SE p)	
Incomplete outcome data addressed	

N' – number of evaluated patients; n – number of patients with outcome

Continuous data

Outcome:…… Follow up:……			
Intervention group			
N	Mean value at baseline (SD / SE / other)	Mean endpoint value (SD / SE / other)	Mean change from baseline (SD / SE / other)
p			
Incomplete outcome data addressed			

N' – number of evaluated patients; n – number of patients with outcome

Appendix 5 Diagnostic search strategies

Diagnostic part

Searches Ovid MEDLINE (January 2011)

1.	exp Vulvar Neoplasms/	6177
2.	((vulva or vulval or vulvar) adj5 (cancer$ or carcinoma$ or adenocarcinoma$ or carcinogen$ or sarcoma$ or malignan$ or tumo?r$ or neoplas$)).mp. [mp = title, original title, abstract, name of substance word, subject heading word, unique identifier]	6992
3.	1 or 2	6992
4.	exp Sentinel Lymph Node Biopsy/	5846
5.	(sentinel adj2 lymph$ adj2 node$).tw.	4719
6.	(lymphatic adj2 mapping).tw.	907
7.	(SLN adj2 biops$).tw.	797
8.	lymphoscintigraphy.mp.	1824
9.	4 or 5 or 6 or 7 or 8	7903
10.	3 and 9	202

Searches EMBASE (January 2011)

1.	'vulva tumour'/exp	6733
2.	((vulva OR vulval OR vulvar) NEAR/5 (cancer* OR carcinoma* OR adenocarcinoma* OR carcinogen* OR sarcoma* OR malignan* OR tumo?r* OR neoplas*)):lnk,ab,ti	4922
3.	#1 OR #2	7756
4.	'sentinel lymph node biopsy'/exp	5894
5.	sentinel:lnk,ab,ti AND lymph*:lnk,ab,ti AND node*:lnk,ab,ti	7884
6.	(lymphatic NEAR/2 mapping):lnk,ab,ti	1071
7.	(sln NEAR/2 biops*):lnk,ab,ti	995
8.	lymphoscintigraphy:lnk,ab,ti	2283
9.	#4 OR #5 OR #6 OR #7 OR #8	10,783
10.	#3 AND #9 AND [embase]/lim	255

Appendix 6 Effectiveness search strategies

Primary studies effectiveness

EMBASE (14 January 2011)

1.	'vulva tumour'/exp	6733
2.	((vulva OR vulval OR vulvar) NEAR/5 (cancer* OR carcinoma* OR adenocarcinoma* OR carcinogen* OR sarcoma* OR malignan* OR tumo?r* OR neoplas*)):lnk,ab,ti	4922
3.	#1 OR #2	7756
4.	#1 OR #2 AND [embase]/lim	5584
5.	#1 OR #2 AND [embase]/lim AND ([editorial]/lim OR [letter]/lim)	252
6.	#1 OR #2 AND [embase]/lim AND [animals]/lim	94
7.	#1 OR #2 NOT #5 NOT #6 AND [embase]/lim	5238
8.	#1 OR #2 NOT #5 NOT #6 AND [embase]/lim AND ([erratum]/lim OR [note]/lim)	69
9.	#1 OR #2 NOT #5 NOT #6 AND [embase]/lim NOT #8	5169

Ovid MEDLINE (17 January 2011)

1.	exp Vulvar Neoplasms/	6177
2.	((vulva or vulval or vulvar) adj5 (cancer$ or carcinoma$ or adenocarcinoma$ or carcinogen$ or sarcoma$ or malignan$ or tumo?r$ or neoplas$)).mp.	6992
3.	#1 or #2	6992
4.	limit #3 to (editorial or letter)	253
5.	#3 not #4	6739
6.	limit #5 to animals	239
7.	#5 not #6	6500

Appendix 7 Coding manual for case–control studies

A study was awarded with maximum one star (*) for each numbered item within the Section and Exposure categories and a maximum of two stars (**) in the Comparability category.

Section	Question	
Selection	***1. Is the case definition adequate?***	
	Yes, with independent validation*	Requires some independent validation (e.g. > 1 person/record/time/process to extract information, or reference to primary record source such as X-rays or medical/hospital records)*
	Yes, e.g. record linkage or based on self reports	Record linkage (e.g. *International Classification of Diseases* codes in database) or self-report with no reference to primary record
	No description	No description
	2. Representativeness of the cases	
	Consecutive or obviously representative series of cases*	All eligible cases with outcome of interest over a defined period of time, all cases in a defined catchment area, all cases in a defined hospital or clinic, group of hospitals, health maintenance organisation, or an appropriate sample of those cases (e.g. random sample)*
	Potential for selection biases or not stated	Not satisfying requirements in part (a), or not stated
	3. Selection of controls	
	This item assesses whether or not the control series used in the study is derived from the same population as the cases and essentially would have been cases had the outcome been present	
	Community controls*	Community controls (i.e. same community as cases and would be cases if had outcome)
	Hospital controls	Hospital controls, within same community as cases (i.e. not another city) but derived from a hospitalised population
	No description	No description
	4. Definition of controls	
	No history of disease (endpoint)*	If cases are first occurrence of outcome, then it must explicitly state that controls have no history of this outcome. If cases have new (not necessarily first) occurrence of outcome, then controls with previous occurrences of outcome of interest should not be excluded*
	No description of source	No mention of history of outcome

Section	Question	
Comparability	***1. Comparability of cases and controls on the basis of the design or analysis***	
	A. Study controls for . . . (select the most important factor)* B. Study controls for any additional factor.* (This criteria could be modified to indicate specific control for a second important factor)*	A maximum of 2 stars can be allotted in this category Either cases and controls must be matched in the design and/or confounders must be adjusted for in the analysis. Statements of no differences between groups or that differences were not statistically significant are not sufficient for establishing comparability Note: If the odds ratio for the exposure of interest is adjusted for the confounders listed, then the groups will be considered to be comparable on each variable used in the adjustment. There may be multiple ratings for this item for different categories of exposure (e.g. ever vs. never, current vs. previous or never)
Exposure	***1. Ascertainment of exposure***	
	Secure record (e.g. surgical records)*	
	Structured interview where blind to case/control status*	
	Interview not blinded to case/control status	
	Written self report or medical record only	
	No description	
	2. Same method of ascertainment for cases and controls	
	Yes*	
	No	
	3. Non-response rate	
	Same rate for both groups*	
	Non respondents described	
	Rate different and no designation	

Appendix 8 Quality assessment questions for case series

Section	Question	
Introduction	1. Are the objective or the hypothesis of the study stated?	
	YES	If the hypothesis was stated or describe
	NO	If the hypothesis wasn't stated or describe
	UNCLEAR	If the hypothesis wasn't stated or describe but study protocol was adequate for study population
Population	2. Was the study population clinically clearly describe? (age, FIGO stage, TNM classification)	
	YES	If there were obligatory information of FIGO stage or TNM classification and at least one more clinically point was describe
	NO	If there wasn't any data about clinically significant information
	UNCLEAR	If authors gave information only about one point in patients clinically description
	2. Was the study population pathologically clearly describe? (histology, tumour size, location, deep of invasion, node's status)	
	YES	If there were information about histology and one other points
	NO	If there wasn't any data about pathologically state
	UNCLEAR	If only one of points mentioned above was reported
Intervention	3. Was intervention of the vulval tumour clearly describe?	
	YES	If there was information about type of treatment with manual protocol
	NO	If there was information only about type of intervention, e.g. surgery, RT, chemotherapy
	UNCLEAR	If there was information only about type of intervention but without description of treatment type, e.g. radical vulvectomy, hemivulvectomy
	4. Was intervention for groin clearly describe?	
	YES	If there were information about type of treatment with description of technique, type of lymphadenectomy.
	NO	If in he study was only information about groins intervention, e.g. groin dissection
	UNCLEAR	If there were information about groins treatment with information about type of procedures but without description, e.g. IFL

Section	Question	
Follow-up	5. Was follow-up time reported?	
	YES	If there were information at least about mean time of observation for all study groups (intervention) patients
	NO	If there wasn't any information about follow-up time
	UNCLEAR	If information were reported only for part of patients, studies groups or there were problems with evaluation of adequate data, e.g. (…) patients were observed until the end of this study (…)
	6. Were all included patients accounted at the end of follow-up?	
	YES	If all patients and data was estimate at the end
	NO	If not all patients accounted at the end of the study, and there were no information from authors about lost patients
	UNCLEAR	If not all patients accounted at the end of the study but information about lost were given
Outcome	7. Were clinically important outcomes considered? (e.g. survival, deaths)	
	YES	If there were all necessary data
	NO	If there were not any information
	UNCLEAR	If information were given in inadequate way e.g. only *p*-value without count of treatment arm
	8. Were definitions of the outcomes presented in the study?	
	YES	If the definitions were described
	NO	If there were not any descriptions of definitions
	UNCLEAR	If author only mentioned or give generally description, e.g. (…) we used this definition according Moore 2001 (…)
	9. Were all outcomes reported in study consequently given?	
	YES	If all were given
	NO	If all were not given
	UNCLEAR	If only part were estimate
	10. In case it was necessary, was it possible to analyse data for patients that meet our criteria separately?	
	YES	If patients were clinically described at the beginning of the study, and study data was presented for each group separately
	NO	If patients were not clinically described the beginning of the study, and study data wasn't presented for each group separately
	UNCLEAR	If patients were not clinically described at the beginning of the study, but study data was presented for each group separately
	11. Were all clinically important date presented in way possible to estimate?	
	YES	If all were presented
	NO	If none of them were not possible to estimate

Appendix 9 Excluded studies with reasons, diagnostic systematic review

Ansink AC, Sie-Go DM, van der Velden J, Sijmons EA, de Barros LA, Monaghan JM, *et al*. Identification of sentinel lymph nodes in vulval carcinoma patients with the aid of a patent blue V injection: a multicenter study. *Cancer* 1999;**86**:652–6	Lack of data on population – no data about stage
Atienza Merino G. Applicability of the identification and biopsy technique of the sentinel-lymph-node in vulval cancer. *Prog Obstet Ginecol* 2010;**53**:403–11	Wrong type of publication – review
Barton DP, Berman C, Cavanagh D, Roberts WS, Hoffman MS, Fiorica JV, *et al*. Lymphoscintigraphy in vulval cancer: a pilot study. *Gynecol Oncol* 1992;**46**:341–4	Wrong population – stage I–II < 70%
Boran N, Kayikcioglu F, Kir M. Sentinel lymph node procedure in early vulval cancer. *Gynecol Oncol* 2003;**90**:492–3	Lack of data on population – no data about stage
Bowles J, Terada KY, Coel MN, Wong JH. Preoperative lymphoscintigraphy in the evaluation of squamous cell cancer of the vulva. *Clin Nucl Med* 1999;**24**:235–8	Wrong study design – small number of patients, no data about accuracy
Carcopino X, Houvenaeghel G, Buttarelli M, Charaffe-Jauffret E, Gonzague L, Rossi I. Feasibility and morbidity of sentinel lymph node detection in patients with vulval carcinoma. [French]. *Bull Cancer* 2005;**92**:489–97	Wrong population – stage I–II < 70%
Cepni I, Kahraman N, Isiloglu H, Arvas M, Demirkiran F, Uzum F, *et al*. Preoperative assessment of lymph nodes metastases in gynecologic malignancies by pelvic lymphoscintigraphy. *Eur J Lymphol Relat Probl* 1992;**3**:111–18	Small sample size
Crane LM, Pleijhuis RG, Themelis G, Harlaar NJ, Sarantopoulos A, Arts HG, *et al*. Detection of the sentinel lymph node in vulval cancer, using near-infrared fluorescence intraoperative imaging: a technical feasibility study. *Mol Imaging Biol* 2010;**12**:S1164	Small sample size
Crane LM, Themelis G, Buddingh T, Harlaar NJ, Pleijhuis RG, Sarantopoulos A, *et al*. Multispectral real-time fluorescence imaging for intraoperative detection of the sentinel lymph node in gynecologic oncology. *J Visualised Experiments* 2010;**44**:2225	Wrong study design – description of the technique
Crane LMA, Themelis G, Arts HJG, Buddingh KT, Brouwers AH, Ntziachristos V, *et al*. Intraoperative near-infrared fluorescence imaging for sentinel lymph node detection in vulval cancer: First clinical results. *Gynecol Oncol* 2011;**120**:291–5	Wrong study design – small number of patients
De Cesare SL, Fiorica JV, Roberts WS, Reintgen D, Arango H, Hoffman MS, *et al*. A pilot study utilising intraoperative lymphoscintigraphy for identification of the sentinel lymph nodes in vulval cancer. *Gynecol Oncol* 1997;**66**:425–8	Small sample size
de Hullu JA, Piers DA, Hollema H, Aalders JG, Van der Zee AG. Sentinel lymph node detection in locally recurrent carcinoma of the vulva. *BJOG* 2001;**108**:766–8	Wrong study design – case study
Echt ML, Finan MA, Hoffman MS, Kline RC, Roberts WS, Fiorica JV. Detection of sentinel lymph nodes with lymphazurin in cervical, uterine, and vulval malignancies. *South Med J* 1999;**92**:204–8	Wrong population – stage I–II < 70%
Farrell C, Lee ST, Grant MP, Rowe C. Rapid localisation of sentinel lymph nodes in vulval lymphoscintigraphy. *Intern Med J* 2010;**40**:26–7	Wrong type of publication – abstract
Fuste P, Ortega M, Vidal S, Mancebo G, Alameda F, Carreras R. Feasibility of the sentinel lymph node technique in cervical and vulval cancers. [Spanish]. *Medicina Clinica* 2007;**128**:569–71	Small sample size

Garcia JD, Altolaguirre GR, Domingo JM, Ormaetxea GM, Goikoetxea NA, Gonzalvo EA, *et al*. First results in the sentinel lymph node procedure in vulval squamous cell carcinoma. *Prog Obstet Ginecol* 2009;**52**:675–80	Small sample size
Hakam A, Nasir A, Raghuwanshi R, Smith PV, Crawley S, Kaiser HE, *et al*. Value of multilevel sectioning for improved detection of micrometastases in sentinel lymph nodes in invasive squamous cell carcinoma of the vulva. *Anticancer Res* 2004;**24**:1281–6	Wrong population – stage I–II < 70%
Knopp S, Holm R, Trope C, Nesland JM. Occult lymph node metastases in early stage vulval carcinoma patients. *Gynecol Oncol* 2005;**99**:383–7	Wrong intervention – non-sentinel lymph node examination
Kudo G, Toyama H, Hasegawa K, Kuroda M, Hattori H, Ishiguro M, *et al*. Sentinel lymph node navigation surgery in Paget's disease of the vulva. *Clin Nucl Med* 2002;**27**:909–10	Wrong study design – case study
Kurzl R, Friese K. Diseases of the vulva. *Gynakologe* 2009;**42**:245–6	Wrong type of publication – review
Levenback CF, Tian C, Coleman RL, Gold MA, Fowler JM, Judson PL. Sentinel node (SN) biopsy in patients with vulval cancer: A Gynecologic Oncology Group (GOG) study. *J Clin Oncol* 2009;**27**:5505	Wrong type of publication – abstract
Louis-Sylvestre C, Evangelista E, Leonard F, Ittu E, Meignan M, Paniel BJ. Sentinel node localisation should be interpreted with caution in midline vulval cancer. *Gynecol Oncol* 2005;**97**:151–4	Duplicate of Louis-Sylvestre *et al*. 2006[65]
Makowski L, Tjahjadi M, Grubner S, Friedrich Gratz K, Hillemanns P, Hertel H. SPECT-CT improved the detection of inguinal sentinel lymph nodes in vulval cancer. *Arch Gynecol Obstet* 2010;**282**:S199	Wrong type of publication – abstract
Maza S, Taupitz M, Taymoorian K, Winzer KJ, Ruckert J, Paschen C, *et al*. Multimodal fusion imaging ensemble for targeted sentinel lymph node management: Initial results of an innovative promising approach for anatomically difficult lymphatic drainage in different tumour entities. *Eur J Nucl Med Mol Imaging* 2007;**34**:378–83	Wrong study design – case study
Molpus KL, Kelley MC, Johnson JE, Martin WH, Jones HW, III. Sentinel lymph node detection and microstaging in vulval carcinoma. *J Reprod Med* 2001;**46**:863–9	Lack of data on population – no data about stage
Moore RG, Granai CO, Gajewski W, Gordinier M, Steinhoff MM. Pathologic evaluation of inguinal sentinel lymph nodes in vulval cancer patients: a comparison of immunohistochemical staining versus ultrastaging with haematoxylin and eosin staining. *Gynecol Oncol* 2003;**91**:378–82	Lack of data on population – no data about stage
Moore RG, DePasquale SE, Steinhoff MM, Gajewski W, Steller M, Noto R, *et al*. Sentinel node identification and the ability to detect metastatic tumour to inguinal lymph nodes in squamous cell cancer of the vulva. *Gynecol Oncol* 2003;**89**:475–9	Wrong population – stage I–II < 70%
Oonk MH, de Hullu JA, Van der Zee AG. Current controversies in the management of patients with early-stage vulval cancer. *Curr Opin Oncol* 2010;**22**:481–6	Wrong type of publication – review
Oonk MHM, Hollema H, de Hullu JA, Van der Zee AGJ. Prediction of lymph node metastases in vulval cancer: A review. *Int J Gynecol Cancer* 2006;**16**: 963–71	Wrong type of publication – review
Oonk MHM, van de Nieuwenhof HP, de Hullu JA, Van der Zee AGJ. The role of sentinel node biopsy in gynecological cancer: a review. *Curr Opin Oncol* 2009;**21**:425–32	Wrong type of publication – review
Radziszewski J, Bidzinski M, Panek G, Sobiczewski P, Derlatka P, Nasierowska-Guttmejer A, *et al*. Sentinel lymph node in vulval cancer – a pilot study to identify and assess the diagnostic value. *Nowotwory* 2003;**53**:270–4	Lack of data on population
Regauer S. Histopathological work-up and interpretation of sentinel lymph nodes removed for vulval squamous cell carcinoma. *Histopathology* 2009;**55**:174–81	Wrong study design

Robison K, Steinhoff MM, Granai CO, Brard L, Gajewski W, Moore RG. Inguinal sentinel node dissection versus standard inguinal node dissection in patients with vulval cancer: A comparison of the size of metastasis detected in inguinal lymph nodes. *Gynecol Oncol* 2006;**101**:24–7	Wrong type of publication – no data about accuracy
Rodier JF, Janser JC, Routiot T, David E, Ott G, Schneegans O, *et al*. Sentinel node biopsy in vulval malignancies: a preliminary feasibility study. *Oncol Rep* 1999 Nov;**6**:1249–52	Small sample size
Schmidt E, Zambo K, Hartmann T, Dehghani B, Bodis J. Preliminary experiences with sentinel nod detection in cases of vulval malignancy. *Magy Noorv Lapja* 2003;**66**:177–80	Duplicate – published in English in Zambo 2002 (later also excluded)
Sideri M, De Cicco C, Maggioni A, Colombo N, Bocciolone L, Trifirò G, *et al*. Detection of sentinel nodes by lymphoscintigraphy and gamma probe guided surgery in vulval neoplasia. *Tumori* 2000;**86**:359–63	Paper not received
Tavares MG, Sapienza MT, Galeb NA, Jr., Belfort FA, Costa RR, Osorio CA, *et al*. The use of 99mTc-phytate for sentinel node mapping in melanoma, breast cancer and vulval cancer: a study of 100 cases. *Eur J Nucl Med* 2001; **28**:1597–604	Lack of data on population – no data about stage
Terada KY, Coel MN, Ko P, Wong JH. Combined use of intraoperative lymphatic mapping and lymphoscintigraphy in the management of squamous cell cancer of the vulva. *Gynecol Oncol* 1998;**70**:65–9	Small sample size
Terada KY, Shimizu DM, Wong JH. Sentinel node dissection and ultrastaging in squamous cell cancer of the vulva. *Gynecol Oncol* 2000;**76**:40–4	Small sample size
Tjin Asjoe FM, van BE, Ewing P, Burger CW, Ansink AC. Sentinel node procedure in vulval squamous cell carcinoma: a histomorphologic review of 32 cases. The significance of anucleate structures on immunohistochemistry. *Int J Gynecol Cancer* 2008;**18**:1032–6	Wrong study design
Van Der Zee A, Oonk M, Van Hemel B, de Hullu J, Ansink A, Van Der Velde J, *et al*. Sentinel lymph node metastasis in patients with vulval cancer mandates adjuvant groin treatment, independent of size. *Gynecol Oncol* 2009; **112**:S14–15	Wrong type of publication – abstract
Wydra D, Matuszewski R, Romanowicz G, Bandurski T. Evaluation of surgical gamma probes for sentinel node localisation in cervical and vulval cancer. *Nucl Med Rev* 2005;**8**:105–10	Wrong study design – description of the technique
Wydra D, Sawicki S, Emerich J, Romanowicz G. Evaluation of sentinel node detection in vulval cancer. *Nucl Med Rev* 2005;**8**:128–30	Small sample size
Zambo K, Schmidt E, Hartmann T, Kornya L, Dehghani B, Tinneberg HR, *et al*. Preliminary experiences with sentinel lymph node detection in cases of vulval malignancy. *Eur J Nucl Med Mol Imaging* 2002;**29**:1198–200	Small sample size

Appendix 10 Additional data from test accuracy systematic review

TABLE 52 Inclusion and exclusion criteria: test accuracy studies

Study (author, year)	Patients in correct clinical stage (%)	Inclusion criteria	Exclusion criteria
Achimas-Cadariu *et al.*, 2009[50]	95%	SCC; no bulky groin lymph nodes	Other than SCC or non-invasive cancer
Basta *et al.*, 2005[51]	100%	FIGO I–II	NR
Brunner *et al.*, 2008[52]	100%	SCC; T1–T2; without clinically suspicious inguinal nodes	Vulval cancer with an invasion depth < 1 mm, vulval melanoma, adenocarcinoma, basal cell cancer, verrucous carcinoma, prior chemotherapy, pelvic or inguinal RT or prior vulval surgery
Camara *et al.*, 2009[53]	94.1%	Clinical stage I–II	NR
Crosbie *et al.*, 2010[54]	100%	Clinical stage I–II; SCC, > 4 cm, stromal invasion < 1 mm; without clinically suggestive of metastasis	NR
De Cicco *et al.*, 2000[55]	100%	Stage T1–T2; SCC, without clinically suggestive of metastasis	Patients with clinically positive groin nodes, pregnant or lactating patients were excluded from the study
de Hullu *et al.*, 2000[56]	100%	Stage T1–T2; SCC, without clinically suggestive of metastasis	NR
Hampl *et al.*, 2008[57]	94.4%	SCC; T1–T3	Unresectable tumours, suspicious nodes in the groin (detected by ultrasonography) or cytologically or histologically proven lymphatic metastases
Hauspy *et al.*, 2007[58]	95%	Stage T1–T2; SCC, without clinically suggestive of metastasis	NR
Johann *et al.*, 2008[59]	86%	SCC; T1–T2	NR
Klat *et al.*, 2009[60]	100%	Stage T1B–T2; SCC	T3 or T4 tumours and with palpable and enlarged lymph nodes
Levenback *et al.*, 2001[61]	87%	Primary surgical treatment for vulval cancer, regardless of clinical stage or histological features	NR
Lindell *et al.*, 2010[39]	98%	With the intention to learn the procedure; T1–T3 without palpable lymph nodes in the groins, no upper limit for tumour size; (one patient with multifocal tumour)	NR
Louis-Sylvestre *et al.*, 2006[62]	100%		Prior vulval surgery

continued

TABLE 52 Inclusion and exclusion criteria: test accuracy studies (*continued*)

Study (author, year)	Patients in correct clinical stage (%)	Inclusion criteria	Exclusion criteria
		Stage T1–T2; stromal invasion < 1 mm	
Martinez-Palones *et al.*, 2006[63]	92.9%	Stage T1–T2; SCC or melanoma	In situ carcinoma, depth invasion < 1 mm, prior chemotherapy or RT or vulval surgery, and vaginal, rectal or urinary bladder involvement (stages III–IV)
Merisio *et al.*, 2005[64]	100%	T1-T2; histologically confirmed invasive SCC, clinically negative groins, no prior chemotherapy or RT	Clinically positive groin nodes, stage T3–T4, pregnant or lactating
Moore *et al.*, 2008[65]	100%	Clinical stage I–II, SCC; > 4 cm, stromal invasion < 1 mm; without clinically suggestive of metastasis	NR
Nyberg *et al.*, 2007[66]	100%	Vulval cancer	Previous radiation therapy, no surgery at all, incomplete surgery, the primary operation performed at some other hospital or before 2001, or unknown origin of the malignant disease
Pityński *et al.*, 2003[67]	100%	Stage I–II	NR
Radziszewski *et al.*, 2010[68]	100%	Clinical stage I–II; SCC, > 4 cm, stromal invasion < 1 mm, T1–T2, N0, M0	Prior vulval surgery, positive nodes
Rob *et al.*, 2007[69]	100%	Clinical stage I–II; SCC, > 4 cm, stromal invasion < 1 mm; without clinically suggestive of metastasis	NR
Terada *et al.*, 2006[70]	100%	SCC; T1; at least 1 mm of invasion	Locally advanced tumours (T2–4), gross adenopathy in the groin
Vakselj *et al.*, 2007[71]	92%	Vulval cancer	NR
Van den Eynden *et al.*, 2003[72]	100%	Vulval cancer	NR
Van der Zee *et al.*, 2008[73]	100%	SCC; T1–T2; < 4 cm; depth of invasion > 1 mm and clinically non-suspicious inguinofemoral lymph nodes; registered at the University Medical Centre Groningen (Groningen, the Netherlands); amendment – multifocal	NR
Vidal-Sicart *et al.*, 2007[74]	86%	Need to perform an IFL in patients initially proposed for a curative surgical procedure	NR

NR, not reported.

TABLE 53 Treatment description – test accuracy studies

Study (author, year)	Treatment after SLN biopsy		IFL	
	Surgery	RT	Unilateral	Bilateral
Achimas-Cadariu *et al.*, 2009[50]	Modified hemivulvectomy or wide local excision *n* = 19, radical vulvectomy: 27	NR	Lateral T1	NR
Basta *et al.*, 2005[51]	Radical vulvectomy: 39[a]	NR	NR	NR
Brunner *et al.*, 2008[52]	NR	NR	34	10
Camara *et al.*, 2009[53]	Hemivulvectomy: 5, radical vulvectomy: 12	NR	NR	NR
Crosbie *et al.*, 2010[54]	NR	NR	NR	NR
De Cicco *et al.*, 2000[55]	Wide excision, hemivulvectomy, radical vulvectomy	NR	18	19
de Hullu *et al.*, 2000[56]	Radical excision: 59	NR	11	48
Hampl *et al.*, 2008[57]	Hemivulvectomy (35%) or vulvectomy (35%), followed by local tumour resection in 30% of *n* = 127	NR	21	103
Hauspy *et al.*, 2007[58]	Wide excision: 31; radical vulvectomy: 8	2	NR	NR
Johann *et al.*, 2008[59]	Wide excision: 3; hemivulvectomy: 13; radical vulvectomy: 7	NR	1	22
Klat *et al.*, 2009[60]	Radical excision, radical vulvectomy	NR	NR	NR
Levenback *et al.*, 2001[61]	NR	1[b]	28	24
Lindell *et al.*, 2010[39]	Wide excision: 17;[c] hemivulvectomy: 24	NR	24[d]	53
Louis-Sylvestre *et al.*, 2006[62]	Radical vulvectomy or other	NR	NR	NR
Martinez-Palones *et al.*, 2006[63]	Radical vulvectomy: 28	NR	16	12
Merisio *et al.*, 2005[64]	Radical vulvectomy: 20	NR	8	12
Moore *et al.*, 2008[65]	Radical wide excision or radical vulvectomy	RT (patients with SLN+)	16	19
Nyberg *et al.*, 2007[66]	NR	NR	NR	NR
Pityński *et al.*, 2003[67]	Wide excision	NR	NR	NR
Radziszewski *et al.*, 2010[68]	NR	NR	5	51
Rob *et al.*, 2007[69]	Radical excision, radical vulvectomy, other	NR	NR	NR
Terada *et al.*, 2006[70]	Radical excision	Palliative	17	4
Vakselj *et al.*, 2007[71]	NR	9	19	16

continued

TABLE 53 Treatment description – test accuracy studies (*continued*)

Study (author, year)	Treatment after SLN biopsy		IFL	
	Surgery	RT	Unilateral	Bilateral
Van den Eynden *et al.*, 2003[72]	Wide excision, hemivulvectomy, radical vulvectomy	NR	13	14
Van der Zee *et al.*, 2008[73]	Wide local excision: 358;[c] radical vulvectomy: 41; other: 4	Four patients, > 1 intranodal metastasis and/or extra nodal growth was detected, postoperative external beam RT (50 Gy) to the groin/pelvis	183	NR
Vidal-Sicart *et al.*, 2007[74]	Wide excision: 2; radical vulvectomy: 7; other: 17	NR	30	19

Gy, gray (radiation unit); NR, not reported.
a Radical surgery.
b In one patient undergoing SLN biopsy only, the SLN was grossly positive. The surgeon aborted further IFL in favour of RT.
c Wide local excision.
d Two patients with central tumour had unilateral dissection due to complications (e.g. heavy bleeding, severe aspiration).

TABLE 54 Index test and gold standard technical details: test accuracy studies

	Reference standard			Index test (blue dye)		Index test (^{99m}Tc)			
Study ID	IFL/follow-up	Follow-up	H&E/ immunostained	Type	Dose	Type	Dose	Imaging	Lymphoscintigraphy
Achimas-Cadariu *et al.*, 2009[50]	IFL for SLN + and follow-up for SLN-	Annually by screening the in-house records or contacting the office gynaecologists or family doctors	H&E for all and ultrastaging for SLN negative	Isosulfan blue dye	NR	^{99m}Tc colloid albumin	NR	Handheld probe (Gammafinder, G100, World of Medicine, Berlin, Germany)	Performed
Basta *et al.*, 2005[51]	IFL for all	NA	H&E and immunostained for all	Patent Blau (Guerbet GmbH)	2–4 ml	Nanocolloid (Nycomed Amersham Sorin S. r.l., Saluggia, Italy) marked with ^{99m}Tc	2.5 mCi	Handheld gamma probe (Navigator GPS, USSC; Tyco Healthcare, Mansfield, MA, USA)	Performed
Brunner *et al.*, 2008[52]	IFL and follow-up for all	3-month intervals, including inspection, vaginorectal and groin palpation, and in some cases, serum tumour marker evaluation. Mean 35.8 months, SD 122.1 months	H&E for all and immunostained for H&E negative	Isosulfan blue	1 ml	^{99m}Tc microcolloidal-containing albumin (Albures, Pharmaceutical Nycomed Amersham, Braunschweig, Germany)	15 MBq tracer in 0.4 ml saline	Single- or dual-head gamma camera; handheld collimated gamma probe (ScintiProbe MR-100, Pol.Hi.Tech, Carsoli, Italy)	Preoperative lymphoscintigraphy was performed in all
Camara *et al.*, 2009[53]	IFL for all	NA	NR	Blue dye	NR	^{99m}Tc nanocolloid	NR	Handheld gamma probe	Not performed
Crosbie *et al.*, 2010[54]	IFL and follow-up for all	Total follow-up: 5 years, assessment every 3 months for the first 2 years and every 6 months thereafter. Includes symptom enquiry and clinical examination of the vulva and groins	H&E and immunostained for negative on H&E	Patent blue dye	3 ml	^{99m}Tc-labelled human serum albumin nanocolloid (Solco Nuclear, Birfelden, Switzerland)	40 MBq/ 0.2 ml	Gamma camera	Preoperative lymphoscintigraphy was performed in all
	IFL for all	NA	H&E	NA	NA				

continued

TABLE 54 Index test and gold standard technical details: test accuracy studies (*continued*)

	Reference standard			Index test (blue dye)		Index test (^{99m}Tc)			
Study ID	IFL/follow-up	Follow-up	H&E/ immunostained	Type	Dose	Type	Dose	Imaging	Lymphoscintigraphy
De Cicco *et al.*, 2000[55]						^{99m}Tc-colloids Nanocoll (Nycomed Amersham Sorin S. R.L., Saluggia, Italy)	0.1 ml of 5 MBq; total dose 10–20 MBq	Gamma-detecting probe (MR 100, Po. li.tech. L'Aquila, Italy)	Preoperative lymphoscintigraphy was performed in all
de Hullu *et al.*, 2000[56]	IFL for all	NA	H&E for all and immunostained for negative on H&E	Blue V dye (Laboratoire Guerbet, Aulney-Sous-Bois, France)	2 ml	^{99m}Tc-labelled nanocolloid (Solco Nuclear, Birfelden, Switzerland)	0.2–0.6 ml of 60 MBq	Handheld probe (Neoprobe; Neoprobe Corporation, Dublin, Ireland)	Preoperative lymphoscintigraphy was performed in all
Hampl *et al.*, 2008[57]	IFL for all	NA	H&E for all and immunostained (unclear if for all)	Patentblue V 2.5% in aqua dest with 0.9 NaCl and 0.05% disodium phosphate (Laboratoire Guerbet, Aulney-Sous-Bois, France)	0.5–1.0 ml	^{99m}Tc-labelled nanocolloid with a particle size of b80 nm (Nycomed Amersham Sorin S. r.l., Saluggia, Italy)	0.2 ml 60–120 MBq	Every 30 minutes, handheld gamma counter	Performed
Hauspy *et al.*, 2007[58]	IFL for all	NA	H&E and immunostained for all	Lymphazurin blue dye	4 ml	Technetium–sulfur colloid	0.1–0.2 mCi	Handheld gamma probe (Navigator GPS; Tyco Healthcare, Mansfield, MA, USA)	Not routinely performed in all patients
Johann *et al.*, 2008[59]	IFL for all	NA	Step sectioning, frozen section	Blue dye	NR	^{99m}Tc-labelled nanocolloids, particle size ≤ 80 nm (Nanocoll™, GE Healthcare, Amersham Health, Braunschweig, Germany)	Four aliquots of 15 MBq	Dual-head gamma camera up to 2 hours post injection (SLN, standard gamma-probe)	Performed
	IFL for all	NA			4 ml		50 MBq		

	Reference standard			Index test (blue dye)		Index test (^{99m}Tc)			
Study ID	IFL/follow-up	Follow-up	H&E/ immunostained	Type	Dose	Type	Dose	Imaging	Lymphoscintigraphy
Klat *et al.*, 2009[60]			H&E and immunostained for all	Bleu patenté V (2.5%, Guebert, Paris, France)		Radionuclide-labelled nanocolloid (Senti-Scint, FJC National Research Institute for Radiobiology and Radiohygiene, Budapest, Hungary)		Handheld gamma counter (NEO 2000; Neoprobe Corporation, Dublin, Ohio OH, USA; Europrobe, Euromedical, Le Chesnay, France)	Preoperative lymphoscintigraphy was performed in all
Levenback *et al.*, 2001[61]	IFL for all	NA	H&E for all and immunostained for negative on H&E	Isosulfan blue dye	1–4 ml	NA	NA	NA	Not performed
Lindell *et al.*, 2010[39]	IFL for all	NA	H&E for all and immunostained for negative on H&E	Blue dye (metyltionincloride, 10 mg/ml, ATL, Stockholm, Sweden)	0.25 ml	Human serum albumin colloid labelled with ^{99m}Tc (Nanocoll, Nycomed Amersham Sorin Srl, Saluggia,Italy)	40 MBq	One head of a Triad XLT gamma camera equipped with a parallel-hole collimator (Trionix Inc., Twinsburg, OH, USA); handheld gamma-probe (Europrobe1, Eurorad, France)	Performed
Louis-Sylvestre *et al.*, 2006[62]	IFL for all	NA	H&E for all and immunostained for negative on H&E	Blue dye	2 ml	^{99m}Tc	30 MBq	Manual probe de détection (Europrobe, Euromédical Instruments, The Chesnay France)	Preoperative lymphoscintigraphy was performed in all
Martinez-Palones *et al.*, 2006[63]	IFL for all	NA	H&E all and immunostained for negative on H&E	Isosulfan blue dye (Lymphazurin 1%, United States Surgical Co., Norwalk, CT, USA)	2–4 ml	^{99m}Tc-labelled nanocolloid (Albu-res®, Pharmaceutical Nycomed Amersham, Bruanschweig, Germany)	NR	Gamma camera	Dynamic lymphoscientigraphy was performed in all

continued

TABLE 54 Index test and gold standard technical details: test accuracy studies (*continued*)

	Reference standard			Index test (blue dye)		Index test (^{99m}Tc)			
Study ID	IFL/follow-up	Follow-up	H&E/ immunostained	Type	Dose	Type	Dose	Imaging	Lymphoscintigraphy
Merisio *et al.*, 2005[64]	IFL for all	NA	H&E for all and immunostained for negative on H&E	NA	NA	Radioactively labelled ^{99m}Tc pertechnetate, nanocolloid particles (< 80 nμ in diameter)	Average total dose 11 MBq (range 10–20 MBq) diluted in a 20 : 1-l volume	Adac Thirus and Picker Prisma 2000 (Aurora, OH, USA) XP gamma cameras; ScintiProbe MR-100 surgical probe (IFL)	Preoperative lymphoscintigraphy was performed in all
Moore *et al.*, 2008[65]	IFL for SLN positive and follow-up for SLN negative	At 3-month intervals for the first 2 years and then every 6 months thereafter	H&E for all and immunostained for negative on H&E	Methylene blue	NR	^{99m}Tc–sulphur colloid	2 mCi	Handheld gamma camera	Preoperative lymphoscintigraphy or examination with a handheld gamma counter regardless of the tumour location
Nyberg *et al.*, 2007[66]	IFL for all	NA	NR	Patent blue dye	NR	^{99m}Tc colloid	NR	Handheld gamma probe intraoperatively (Navigator GPS, Tyco HealthCare, Norwalk, CT, USA).	Not performed at all
Pityński *et al.*, 2003[67]	IFL for all	NA	NR	Patent Blue	2 ml	^{99m}Tc nanocolloid	1 ml; 50 MBq	Dual-head gamma camera (Siemens, Munich, Germany); gamma radiation detector (Navigator GPS; Tyco Healthcare, Mansfield, MA, USA)	Not routinely performed

	Reference standard			Index test (blue dye)		Index test (^{99m}Tc)			
Study ID	IFL/follow-up	Follow-up	H&E/ immunostained	Type	Dose	Type	Dose	Imaging	Lymphoscintigraphy
Radziszewski *et al.*, 2010[68]	IFL for all	NA	H&E and Immunostained for all	2.5% of patent blue V	NR	^{99m}Tc colloid albumin	1.2 mCi	Double-headed high-resolution VariCam gamma probe (Elscint, Haifa, Israel) with a low-energy collimator	Preoperative lymphoscintigraphy was performed in all
Rob *et al.*, 2007[69]	IFL for all	NA	H&E and immunostained for all	Blue dye (Patentblau V BYC, Gueden, Germany, or bleu patenté V 2.5%, Guerbet, France)	2 ml	NA	NA	NA	NA
Terada *et al.*, 2006[70]	IFL for SLN + and follow-up for all	Median follow-up 4.6 years (range 2–8 years)	H&E for all and immunostained for negative on H&E	Isosulfan blue dye	0.5 ml	^{99m}Tc-labelled sulphur colloid	0.5 mCi	Handheld gamma probe (Neoprobe Corporation, Dublin, Ohio, OH, USA)	Performed
Vakselj *et al.*, 2007[71]	IFL for SLN + and follow-up for SLN negative	Total follow-up: 4 years	NR	Methylene blue dye	NR	^{99m}Tc-labelled nanocolloid	NR	Gamma camera	Static and dynamic lymphoscintigraphies were performed in all
Van den Eynden *et al.*, 2003[72]	IFL for all (phase 1) IFL for SLN + and follow-up for SLN negative (phase 2)	Phase 2: short follow-up, maximum 17 months	H&E for all and immunostained for negative on H&E	Patent blue violet	0.5–2.0 mL	^{99m}Tc nanocolloid	0.2 ml; 40 MBq	Handheld gamma probe	Performed

continued

TABLE 54 Index test and gold standard technical details: test accuracy studies (*continued*)

	Reference standard			Index test (blue dye)		Index test (^{99m}Tc)			
Study ID	IFL/follow-up	Follow-up	H&E/ immunostained	Type	Dose	Type	Dose	Imaging	Lymphoscintigraphy
Van der Zee *et al.*, 2008[73]	IFL for SLN + and follow-up for SLN negative	For 2 years at intervals of every 2 months. Stopping rules: occurrence of groin recurrences	H&E for all and immunostained for negative on H&E	Patent blue-V (Laboratoire Guerbet, Aulney-Sous-Bois, France)	2.0 ml	^{99m}Tc-labelled nanocolloid (Solco Nuclear, Birsfelden, Switzerland)	0.5 mL of 100 MBq	Single-head gamma camera; handheld gamma-ray detection probe (Neoprobe, Dublin, Ohio, OH, USA)	Performed
Vidal-Sicart *et al.*, 2007[74]	IFL and follow-up for all	Over a mean period of 24 months	H&E for all and immunostained for negative on H&E	Isosulfan blue (Lymphazurin, BenVenue Labs, Bedford, OH, USA), methylene blue (Laboratory Dr Carreras, Barcelona, Spain)	1 ml	^{99m}Tc nanocolloid (Lymphoscint Amersham, Saluggia, Italy)	0.75–1 mCi	Handheld gamma probe (Navigator, USSC, Norwalk, CT, USA)	Performed

MBq, megabecquerel; mCi, millicurie; NA, not applicable; NR, not reported.

TABLE 55 Adverse events reported according to surgical procedures: test accuracy studies

Outcome	Study	SLN biopsy + IFL	SLN biopsy only	SLN biopsy + IFL or follow-up
Short term	Moore *et al.,* 2008[65] *n* = 36	NA	NA	Wound breakdown, *n* = 0 (0%) Wound cellulitis, *n* = 0 (0%) Cellulitis, *n* = 0 (0%) Postoperative groin lymphocele, *n* = 2 (5.5%) Cellulitis arising in the labia majora, *n* = 1 (2.8%)
	Terada *et al.,* 2006[70] *n* = 21	Wound cellulitis, *n* = 2 (9.5%) Seroma, *n* = 1 (4.8%)	NA	NA
	Van den Eynden *et al.,* 2003[72] *n* = 17	NA	NA	Cellulitis, *n* = 1 (5.9%) Lymphocele, *n* = 2 (11.8%)
	Van der Zee *et al.,* 2008[73] *n* = 264	Wound breakdown, *n* = 16 (34%)[a] Wound cellulitis, *n* = 10 (21.3%)[a]	Wound breakdown, *n* = 31 (11.7%) Wound cellulitis, *n* = 12 (4.5)	NA
Long term	Crosbie *et al.,* 2010[54] *n* = 32	NA	NA	Wound infection, *n* = 10 (31%) Wound dehiscence, *n* = 8 (25%) Lymphocyst, *n* = 7 (22%) Chronic lymphoedema, *n* = 5 (16%)
	Van der Zee *et al.,* 2008[73] *n* = 264	Lymphoedema, *n* = 30 (25.2%)[b] Recurrent erysipelas, *n* = 19 (16.2%)[c]	Lymphoedema, *n* = 5 (1.9%) Recurrent erysipelas, *n* = 1 (0.4%)	NA

NA, not applicable.

Appendix 11 Excluded studies with reasons, effectiveness systematic review

Akashi K, Kudo R, Sato T, Tanaka S. Administration of bleomycin in cancer of the female sex organs and the follow-up results. *Sanfujinka No Jissai – Pract Gynecol Obstet* 2011;**11**:1970	Paper not received
Akl A, Akl M, Boike G, Hebert III J, Graham J. Preliminary results of chemoradiation as a primary treatment for vulval carcinoma. *Int J Radiat Oncol Biol Phys* 2000;**48**:415–20	Wrong study design: small study population, $n = 8$
Andersen BL, Hacker NF. Psychosexual adjustment after vulval surgery. *Obstet Gynecol* 1983;**62**:457–62	Wrong study design and data presentation: lack of data, only subjective description
Andreasson B, Bock JE, Visfeldt J. Prognostic role of histology in squamous cell carcinoma in the vulval region. *Gynecol Oncol* 1982;**14**:373–81	Wrong study population: 54% patients met the included criteria, but most data were presented for all study group
Andreasson B, Bock JE, Weberg E. Invasive cancer in the vulval region. *Acta Obstet Gynecol Scand* 1982;**61**:113–19	Wrong study population: 13% of patients had others malignancies
Andreasson B, Moth I, Jensen SB, Bock JE. Sexual function and somatopsychic reactions in vulvectomy-operated women and their partners. *Acta Obstet Gynecolog Scand* 1986;**65**:7–10	Wrong study design: no information about FIGO stage
Andreasson B, Nyboe J. Value of prognostic parameters in squamous cell carcinoma of the vulva. *Gynecol Oncol* 1985;**22**:341–51	Inadequate data presentation: no information about effectiveness of treatment
Ansink AC, Van Tinteren H, Aartsen EJ, Heintz APM. Outcome, complications and follow-up in surgically treated squamous cell carcinoma of the vulva 1956–1982. *Eur J Obstet Gynecol Reprod Biol* 1991;**42**:137–43	Paper waiting to be received
Arvas M, Kose F, Gezer A, Demirkiran F, Tulunay G, Kosebay D. Radical versus conservative surgery for vular carcinoma. *Int J Gynecol Obstet* 2005;**88**:127–33	More than 25% FIGO stage III and IV
Atlante G, Lombardi A, Mariani L, Vincenzoni C. Carcinoma of the vulva (1981–1985): Analysis of a radio-surgical approach. *Eur J Gynaecol Oncol* 1989;**10**:341–8	Paper waiting to be received
Ayhan A, Tuncer ZS, Akarin R, Yucel I, Develioglu O, Mercan R *et al*. Complications of radical vulvectomy and inguinal lymphadenectomy for the treatment of carcinoma of the vulva. *J Surg Oncol* 1992;**51**:243–5	Paper waiting to be received
Bafna UD, Devi K, Naik A, Hazra S, Sushma N, Babu N. Carcinoma of the vulva: a retrospective review of 37 cases at a regional cancer centre in South India. *J Obstet Gynaecol* 2004;**24**:403–7	Wrong study population: too small a study group ($n = 5$), all patients presented as case reports
Bakalianou K, Salakos N, Iavazzo C, Paltoglou G, Papadias K, Gregoriou O, *et al*. Paget's disease of the vulva. A ten-year experience. *Eur J Gynaecol Oncol* 2008;**29**:368–70	Paper waiting to be received
Balat O, Edwards CL, Verschraegen C, Delclos L. The long term results of radiotherapy with or without surgery in management of advanced vulvar cancer: report of 76 patients. *Eur J Gynaecol Oncol* 2000;**21**:426–9	More than 25% FIGO stage III and IV

Balat O, Edwards C, Delclos L. Complications following combined surgery (radical vulvectomy versus wide local excision) and radiotherapy for the treatment of carcinoma of the vulva: Report of 73 patients. *Eur J Gynaecol Oncol* 2000;**21**:501–3	Paper waiting to be received
Baltzer J. Precancerous lesions and early stages of vulval cancer. *Arch Gynecol Obstet* 1989;**245**:498–503	Paper waiting to be received
Barton DP. The prevention and management of treatment related morbidity in vulval cancer. *Best Pract Res Clin Obstet Gynaecol* 2003;**17**:683–701	Wrong study design: review
Beissert M, Steinbach C, Neef G. Treatment of cancer of the vulva. *Zentralbl Gynakol* 1974;**96**:1268–73	Wrong study design: lack of information about patients' clinical stage
Bnedet JL, Turko JL, Fairey RN, Boyes DA. Squamous carcinoma of the vulva: results of treatment 1938 to 1976. *Am J Obstet Gynecol* 1979;**134**:201–7	Sample recruited before 1980
Bergen S, DiSaia PJ, Liao SY, Berman ML. Conservative management of extramammary Paget's disease of the vulva. *Gynecol Oncol* 1989;**33**:151–6	Inadequate data presentation: data format not allowing to obtain meaningful information
Berget A, Larsen JF, Pedersen PH. CO_2 laser in Gynaecology. II. Practical experience. *Ugeskr Laeg* 1982;**144**:3350–3	Paper waiting to be received
Beriwal S, Heron DE, Kim H, King G, Shogan J, Bahri S, *et al*. Intensity-modulated radiotherapy for the treatment of vulval carcinoma: A comparative dosimetric study with early clinical outcome. *Int J Radiat Oncol Biol Phys* 2006;**64**:1395–400	Wrong study design: small study population, $n < 10$
Berman ML, Soper JT, Creasman WT, Olt GT, SiSaia PJ. Conservative surgical management of superficially invasive stage I vulval carcinoma. *Gynecol Oncol* 1989;**35**:352–7	Wrong study population: 16% of patients had concomitant cancer
Bienkiewicz A, Gottwald L, Akoel K M, Lech W, Welfel J, Suzin J. Clinical analysis of 105 cases of vulval cancer. *Ginekol Pol* 2002;**73**:913–18	Inadequate data presentation: lack of information about each FIGO stage group treatment and adequate results presentation/mixed population; the problem and data format was inadequate
Black D, Tornos C, Soslow RA, Awtrey CS, Barakat R R, Chi DS. The outcomes of patients with positive margins after excision for intraepithelial Paget's disease of the vulva. *Gynecol Oncol* 2007;**104**:547–50	Wrong study design: lack of information about FIGO stage/women with vulval Paget's disease (80% primary), no FIGO staging; no comparison with other intervention (all patients underwent surgical excision of primary VI Paget's disease)
Blecharz P, Urbanski K, Karolewski K, Klimek M, Pudelek J, Bieda T, *et al*. Prognostic factors in patients with vulvar cancer in the material of the Crocow Division of Centre of Oncology. *Ginecol Oncol* 2007;**5**:22–8	More than 25% FIGO stage III and IV
Bognel C, Prade M, Charpentier P, Michel G. Paget's disease of the vulva. Seven clinicopathological cases. *Gynecologie* 1980;**31**:527–33	Wrong study design: case report based on seven cases
Boidi Trotti A, Tardy A, Burke P, Tomassone W. Results of radiating treatment with fast electrons in 24 cases of cancer of the vulva. *Minerva Ginecol* 1980;**32**:1011–12	Wrong study design: case report based on three cases
Bokhman J V, Maximov S J, Ebert AD. Efficiency of radical treatment in vulval carcinoma. Analysis of 148 cases. *Zentralbl Gynakol* 1997;**119**:166–72	Paper waiting to be received
Bosquet JG, Kinney WK, Russell AH, Gaffey TA, Magrina JF, Podratz KC. Risk of occult inguinofemoral lymph node metastasis from squamous carcinoma of the vulva. *Int J Radiat Oncol Biol Phys* 2003;**57**:419–24	Wrong study design: study is a retrospective follow-up

Bosquet JG, Magrina JF, Gaffey TA, Hernandez JL, Webb MJ, Cliby WA, *et al*. Long-term survival and disease recurrence in patients with primary squamous cell carcinoma of the vulva. *Gynecol Oncol* 2005;**97**:828–33	Wrong study design: authors do not provide information about FIGO stage of disease, there are no opportunities to extract interesting data
Boutselis JG. Radical vulvectomy for invasive squamous cell carcinoma of the vulva. *Obstet Gynecol* 1972;**39**:827–36	Inadequate data presentation: results presented are cumulative for all patients and all interventions
Boyce J, Fruchter RG, Kasambilides E, Nicastri AD, Sedlis A, Remy JC. Prognostic factors in carcinoma of the vulva. *Gynecol Oncol* 1985;**20**:364–77	More than 25% FIGO stage III and IV
Bozzetti F, Lupi G, Di Re F. Evaluation of lymph node involvement in squamous cell carcinoma of the vulva and therapeutic implications. *Tumori* 1974;**60**:269–77	Paper waiting to be received
Brinton LA, Nasca PC, Mallin K, Baptiste MS, Wilbanks GD, Richart RM. Case–control study of cancer of the vulva. *Obstet Gynecol* 1990;**75**:859–66	Wrong study design: epidemiological investigation of rare reproductive tumours
Bryson SCP, Dembo AJ, Colgan TJ, Thomas GM, Deboer G, Lickrish GM. Invasive squamous-cell carcinoma of the vulva – defining low and high-risk groups for recurrence. *Int J Gynecol Cancer* 1991;**1**:25–31	Paper waiting to be received
Burger MP, Hollema H, Emanuels AG, Krans M, Pras E, Bouma J. The importance of the groin node status for the survival of T1 and T2 vulval carcinoma patients. *Gynecol Oncol* 1995;**57**:327–34	Wrong study design: the main aim was correlation between groin node status and risk factors
Burke TW, Morris M, Roh MS, Levenback C, Gershenson DM. Perineal reconstruction using single gracilis myocutaneous flaps. *Gynecol Oncol* 1995;**57**:221–5	Wrong study population: patients after reconstruction
Busch M. Long-term results of radiotherapy alone for carcinoma of the vulva. *Adv Ther* 1999;**16**: 89–100	Paper waiting to be received
Butler JSB, Milliken DA, Dina R, Eccles SA, Maghami SG, Jameson C, *et al*. Isolated groin recurrence in vulval squamous cell cancer (VSCC). The importance of node count. *Eur J Gynaecol Oncol* 2010;**31**:510–13	Wrong study design: analyses of risk factors
Calista D. Topical 1% cidofovir for the treatment of vulval intraepidermal neoplasia (VIN1) developed on lichen sclerosus. *Int J Dermatol* 2009;**48**:535–6	Wrong study design: case report, VIN1 treated with topical 1% cidofovir
Callies R, Kock U, Zeller GX. Results of treatment of vulval malignoma – retrospective analysis of 119 cases. *Zentralbl Gynakol* 1984;**106**:440–55	Inadequate data presentation: population description
Campagnutta E, Scarabelli C, Juzzolino C. The treatment of carcinoma of the vulva using cis-platinum locoregional endoarterial chemotherapy. *Minerva Ginecol* 1984;**36**:407–9	Wrong study design: case report
Carcopino X, Houvenaeghel G, Buttarelli M, Charaffe-Jauffret E, Gonzague L, Rossi I. Feasibility and morbidity of sentinel lymph node detection in patients with vulval carcinoma. *Bull Cancer* 2005;**92**:489–97	Wrong intervention: diagnostic study
Cavanagh D, Fiorica JV, Hoffman MS, Roberts WS, Bryson SCP, LaPolla JP, *et al*. Invasive carcinoma of the vulva. Changing trends in surgical management. *Am J Obstet Gynecol* 1990;**163**:1007–15	Wrong study population: only 67% patients met the inclusion criteria
Cavanagh D, Shepherd JH. The place of pelvic exenteration in the primary management of advanced carcinoma of the vulva. *Gynecol Oncol* 1982;**13**:318–22	Wrong study design: data presented in cumulative way for all patients (also with advanced disease)
Chan JK, Sugiyama V, Pham H, Gu M, Rutgers J, Osann K, *et al*. Margin distance and other clinic-pathologic prognostic factors in vulvar carcinoma: a multivariate analysis. *Gynecol Oncol* 2007;**104**:636–41	More than 25% FIGO stage III and IV

Chardot C, Dartois D, Keil L. Long term results in vulva carcinoma and discussion about the extent of surgical exeresis. 92 Observations. *Ann Med Nancy Est* 1978;**17**:1323–30	Paper waiting to be received
Cheng X, Zang RY, Wu XH, Li ZT, Cai SM, Zhang ZY. Recurrence patterns and prognostic factors in Chinese patients with squamous cell carcinoma of the vulva treated with primary surgery. *Int J Gynecol Cancer* 2009;**19**:158–62	Wrong study population: only 35% patients met inclusion criteria
Choo YC. Invasive squamous carcinoma of the vulva in young patients. *Gynecol Oncol* 1982;**13**:158–64	Sample recruited before 1980
Chu J, Tamimi HK, Ek M, Figge DC. Stage I vulval cancer: criteria for microinvasion. *Obstet Gynecol* 1982;**59**:716–19	Wrong study population: patients with other tumour
Crane LMA, Themelis G, Arts HJG, Buddingh KT, Brouwers AH, Ntziachristos V, *et al*. Intraoperative near-infrared fluorescence imaging for sentinel lymph node detection in vulval cancer: first clinical results. *Gynecol Oncol* 2010;**120**:291–5	Wrong intervention: diagnostic study
Crane LM, Pleijhuis RG, Themelis G, Harlaar NJ, Sarantopoulos A, Arts HG, *et al*. Detection of the sentinel lymph node in vulval cancer, using near-infrared fluorescence intraoperative imaging: a technical feasibility study. *Mol Imaging Biol* 2010;**12**:S1164	Wrong study design: diagnostic study
Creasman WT. New 123 gynaecologic cancer staging. *Obstetr Gynecol* 1990;**75**:287–8.	Wrong study design: report from the FIGO Committee Meeting
Curtin JP, Rubin SC, Jones WB, Hoskins WJ, Lewis J. Paget's disease of the vulva. *Gynecol Oncol* 1990;**39**:374–7	Inadequate data presentation: all results presented per patient
Danby CS, Margesson LJ. Approach to the diagnosis and treatment of vulval pain. *Dermatol Ther* 2010;**23**:485–504	Paper waiting to be received
de Hullu JA, Van Der Zee AG. Surgical treatment of early-stage vulva carcinoma and the complications of the operation. *Ned Tijdschr Geneeskd* 2005;**149**:336–42	Paper waiting to be received
Dean RE, Taylor ES, Weisbrod DM, Martin JW. The treatment of premalignant and malignant lesions of the vulva. *Am J Obstet Gynecol* 1974;**119**:59–68	Wrong study population: only 67% of patients met inclusion criteria
Dinh VT, Ton TP: Surgical treatment of cancer of the vulva. A study of 21 cases. *J Gynecol Obstet Biol Reprod* 1976;**5**:113–22	Paper waiting to be received
DiSaia PJ, Creasman WT, Rich WM. An alternate approach to early cancer of the vulva. *Am J Obstet Gynecol* 1979;**133**:825–32	Sample recruited before 1980
Donaldson ES, Powell DE, Hanson MB, Van N, Jr. Prognostic parameters in invasive vulval cancer. *Gynecol Oncol* 1981;**11**:184–90	Paper waiting to be received
Durdevic S, Hadzic B, Petrovic D. Radical vulvectomy with inguino-femoral lymphadenectomy in the surgical treatment of vulval carcinoma. *Medicinski Pregled* 2000;**53**:607–12	Paper waiting to be received
Dvoretsky PM, Bonfiglio TA, Helmkamp BF. The pathology of superficially invasive, thin vulval squamous cell carcinoma. *Int J Gynecol Pathol* 1984;**3**:331–42	Wrong study design: analyses of metastases' pathology
Ebert A, Baege A, Maximov S and Bokhman JV. 148 cases of vulval cancer – results of radical surgical treatment. *Eur J Cancer* 1997;**33**:942	Paper waiting to be received
Egwuatu VE, Ejeckam GC. An analysis of tumours of the female genital tract in Enugu, Nigeria (1973–1979): a hospital based tumour registry review. *Bull Cancer* 1980;**67**:535–9	Inadequate data: no information about FIGO or TNM stage, patients without other gynaecological tumours

Eifel PJ. Radiotherapy versus radical surgery for 123 gynaecologic neoplasms: carcinomas of the cervix and vulva. *Front Radiat Ther Oncol* 1993;**27**:130–42	Paper waiting to be received
Eke AC, Alabi-Isama LI, Akabuike JC. Management options for vulval carcinoma in a low resource setting. *World J Surg Oncol* 2010;**8**:94	Wrong study population: all stages included in the investigation
Elit L, Hancock G, Carey M, Dal Bello D, Allen HH. Comparing the morbidity of single versus separate incision surgical approaches to vulval cancer. *J Gynecol Tech* 1999;**5**:147–50	Paper waiting to be received
Eva LJ, Ganesan R, Chan KK, Honest H, Malik S, Luesley DM. Vulval squamous cell carcinoma occurring on a background of differentiated vulval intraepithelial neoplasia is more likely to recur: A review of 154 cases. *J Reprod Med Obstet Gynecol* 2008;**53**:397–401	Inadequate data presentation
Eva LJ, Ganesan R, Chan KK, Honest H, Luesley DM. Differentiated-type vulval intraepithelial neoplasia has a high–risk association with vulval squamous cell carcinoma. *Int J Gynecol Cancer* 2009;**19**:741–4	Wrong study design: inadequate study question
Fairey RN, MacKay PA, Benedet JL. Radiation treatment of carcinoma of the vulva, 1950–1980. *Am J Obstet Gynecol* 1985;**151**:591–7	Paper waiting to be received
Fanning J, Lambert HCL, Hale TM, Morris PC, Schuerch C. Paget's disease of the vulva: Prevalence of associated vulval adenocarcinoma, invasive Paget's disease, and recurrence after surgical excision. *Am J Obstet Gynecol* 1999;**180**:24–7	Wrong study population: only 58% of patients met inclusion criteria
Fariaseisner R, Cirisano FD, Grouse D, Leuchter RS, Karlan BY, Lagasse LD, *et al*. Conservative and individualized surgery for early squamous carcinoma of the vulva – the treatment of choice for stage-I and II (T1–2N0–1M0) disease. *Gynecol Oncol* 1994;**53**:55–8	Paper waiting to be received
Faul CM, Mirmow D, Huang Q, Gerszten K, Day R, Jones MW. Adjuvant radiation for vulvar carcinoma: improved local control. *Int J Radiat Oncol Biol Phys* 1997;**38**:381–9	More than 25% FIGO stage III and IV
Feakins RM, Lowe DG. Basal cell carcinoma of the vulva: A clinicopathologic study of 45 cases. *Int J Gynecol Pathol* 1997;**16**:319–24	Paper waiting to be received
Fioretti P, Gadducci A, Prato G, Tavella N, Fanucchi A, Facchini V. The influence of some prognostic factors on the clinical outcome of patients with squamous cell carcinoma of the vulva. *Eur J Gynaecol Oncol* 1992;**8**:97–104	More than 25% FIGO stage III and IV
Fiorica JV, Roberts WS, LaPolla JP, Hoffman MS, Barton DPJ, Cavanagh D. Femoral vessel coverage with dura mater after inguinofemoral lymphadenectomy. *Gynecol Oncol* 1991:**42**:217–21	Wrong study design: information only about primary risk factors of disease
Fonseca-Moutinho JA, Coelho MC, Dilva DP. Vulvar squamous cell carcinoma. Prognostic factors for local recurrence after primary en bloc radical vulvectomy and bilateral groin dissection. *J Reprod Med* 2000;**45**:672–8	More than 25% FIGO stage III and IV
Frankendal B, Larsson LG, Westling P. Carcinoma of the vulva. Results of an individualized treatment schedule. *Acta Radiol Ser Ther Phys Biol* 1973;**12**:165–74	Wrong study population: five patients had other tumours
Frankman O, Kabulski Z. Malignancy grading and prognosis from a biopsy only in cases of electrocoagulated squamous cell carcinoma of the vulva, stages I and II. *Int J Gynecol Obstet* 1983;**21**:119–24	Paper waiting to be received

Friedrich J, Wilkinson EJ, Fu YS. Carcinoma in situ of the vulva: a continuing challenge. *Am J Obstet Gynecol* 1980;**136**:830–43	Paper waiting to be received
Frischbier HJ, Thomsen K, Schmermund HJ, Oberheuser F, Hohne G, Lohbeck HU. Radiotherapy of carcinoma of the vulva – treatment results of electron therapy in 446 patients 1956 to 1978. *Geburtsh Frauenheilk* 1985;**45**:1–5	Sample recruited before 1980
Gaarenstroom KN, Kenter GG, Trimbos JB, Agous I, Amant F, Peters AA, et *al*. Postoperative complications after vulvectomy and inguinofemoral lymphadenectomy using separate groin incisions. *Int J Gynecol Oncol* 2003;**13**:22–7	Wrong study population: only 69.7% of patients met inclusion criteria
Gadducci A, Prato B, Fanucchi A, Bonuccelli A, Cristofani R, Facchini V. Disease-free survival in patients with squamous cell carcinoma of the vulva treated with radical vulvectomy and bilateral inguinal-femoral lymphadenectomy: Analysis of prognostic variables. *Cancer J* 1993;**6**:269–73	Wrong study population: only 73% of patients met inclusion criteria
Garcia I, Tejerizo L, Garcia S, Hernandez H, Velasco M, Lanchares P. Prognosis factors in cancer of the vulva. *Eur J Gynaecol Oncol* 1993;**14**:386–91	Paper waiting to be received
Geisler JP, Manahan KJ, Buller RE. Neoadjuvant chemotherapy in vulval cancer: Avoiding primary exenteration. *Gynecol Oncol* 2006;**100**:53–7	Inadequate data presentation: no information about patients' FIGO stage, all patients presented as case report
Gerbaulet A, Sendra F, Gallez D, Pejovi-Lenfant MH, Haie-Meder C, Michel G, *et al*. The role of radiotherapy in vulval carcinoma. Experience in the Institute Gustave-Roussy and review of the literature. *Oncologia* 1991;**14**:567–73	Paper waiting to be received
Gerszten K, Selvaraj RN, Kelley J, Faul C. Preoperative chemoradiation for locally advanced carcinoma of the vulva. *Gynecol Oncol* 2005;**99**:640–4	Wrong study population: only 33% of patients met inclusion criteria
Ghebre R, Petzel SV, Glubka B, Lindgren B. Quality of life, body image and sexual health for women with vulval cancer. *Gynecol Oncol* 2009;**112**:S165	Paper waiting to be received
Goetze B, Ebert A. Treatment results in cancer of the vulva – analysis of the 113 cases. *Eur J Cancer* 1995;**31**:S249	Wrong study design: poster
Gomez ED, Trincado JM. Effects of chemotherapy on cancer of the vulva. *Prog Obstet Ginecol* 1987;**30**:729–34	Paper waiting to be received
Gonzalez Bosquet J, Magrina J F, Magtibay PM, Gaffey TA, Cha SS, Jones MB, *et al*. Patterns of inguinal groin metastases in squamous cell carcinoma of the vulva. *Gynecol Oncol* 2007;**105**:742–6	Wrong study design: analyses of risk factors
Gordinier ME, Malpica A, Burke TW, Bodurka DC, Wolf JK, Jhingran A, *et al*. Groin recurrence in patients with vulval cancer with negative nodes on superficial inguinal lymphadenectomy. *Gynecol Oncol* 2003;**90**: 625–8	Wrong study design: small population, $n < 10$
Gould N, Kamelle S, Tillmanns T, Scribner D, Gold M, Walker J, *et al*. Predictors of complications after inguinal lymphadenectomy. *Gynecol Oncol* 2001;**82**:329–32	Wrong study population: included also patients with advance stage of disease
Green MS, Naumann R W, Elliot M, Hall JB, Higgins RV, Grigsby JH. Sexual dysfunction following vulvectomy. *Gynecol Oncol* 2000;**77**:73–7	Wrong study population: 51% patients of patients had vulval cancer and dysplasia
Grimshaw RN, Murdoch JB, Monaghan JM. Radical vulvectomy and bilateral inguinal-femoral lymphadenectomy through separate incisions – experience with 100 cases. *Int J Gynecol Cancer* 1993;**3**:18–23	Paper waiting to be received

Gutierrez AR, Rodriguez OA, Rodriguez CS, Chicote MJV, Carreras P S, Fresnadillo J L R, *et al*. Vulval carcinoma: study of 35 patients treated in our service. *Neoplasia* 1997;**14**:211–14	Paper waiting to be received
Hacker NF, Berek JS, Lagasse LD. Individualization of treatment for stage I squamous cell vulval carcinoma. *Obstet Gynecol* 1984;**63**:155–62	Wrong study population: patients with other types of cancer
Hacker NF, Leuchter RS, Berek JS, Castaldo TW, Lagasse LD. Radical vulvectomy and bilateral inguinal lymphadenectomy through separate groin incisions. *Obstet Gynecol* 1981;**58**:574–9	Wrong study population: patients with other types of cancer
Hacker NF, van der Velden J. Conservative management of early vulval cancer. *Cancer* 1993;**71**:1673–7	Wrong study design: review
Hacker NF. Current treatment of small vulval cancers 3495. *Oncology* 1990;**4**:21–5	Paper waiting to be received
Hampl M, Hantschmann P, Michels W, Hillemanns P. Validation of the accuracy of the sentinel lymph node procedure in patients with vulval cancer: results of a multicenter study in Germany. *Gynecol Oncol* 2008;**111**:282–8	Wrong study population: only 44.8% of patients met inclusion criteria
Han SC, Kim DH, Higgins SA, Carcangiu ML, Kacinski BM. Chemoradiation as primary or adjuvant treatment for locally advanced carcinoma of the vulva. *Int J Radiat Oncol Biol Phys* 2000;**47**:1235–44	Wrong study population: patients with local advanced vulval tumour
Hanprasertpong J, Chichareon S, Wootipoom V, Buhachat R, Tocharoenvanich S, Geater A. Clinico-pathological profile of vulval cancer in Southern Thailand: analysis of 66 cases. *J Med Assoc Thailand* 2005;**88**:575–81	Paper waiting to be received
Harberthur F, Almendral AC, Ritter B. Therapy of vulval carcinoma. *Eur J Gynaecol Oncol* 1993;**14**:218–27	Paper waiting to be received
Hatta N, Yamada M, Hirano T, Fujimoto A, Morita R. Extramammary Paget's disease: treatment, prognostic factors and outcome in 76 patients. *Br J Dermatol* 2008; **158**:313–18	Wrong study population: 28% of patients were men
Hauspy J, Beiner M, Harley I, Ehrlich L, Rasty G, Covens A. Sentinel lymph node in vulval cancer. *Cancer* 2007;**110**:1015–23	Wrong intervention: only diagnostic procedures
Heaps JM, Fu YS, Montz FJ, Hacker NF, Berek JS. Surgical-pathologic variables predictive of local recurrence in squamous cell carcinoma of the vulva. *Gynecol Oncol* 1990;**38**:309–14	Wrong study design: prognostic factors associated with the increased risk of cancer recurrence
Hefler LA, Grimm C, Six L, Seebacher V, Polterauer S, Joura E, *et al*. Inguinal sentinel lymph node dissection vs. Complete inguinal lymph node dissection in patients with vulval cancer. *Anticancer Res* 2008;**28**:515–17	Paper waiting to be received
Heidenreich W, Majewski A. Treatment and results in carcinoma of the vulva – a survey of 234 patients treated from 1951 to 1976. *Geburtsh Frauenheilk* 1986;**46**:136–9	Inadequate data presentation: results presented in cumulative way for patients with primary and advanced disease
Helgason NM, Hass AC, Latourette HB. Radiation therapy in carcinoma of the vulva. A review of 53 patients. *Cancer* 1972;**30**:997–1000	Paper waiting to be received
Hidano A, Nakajima S. Earliest features of the strawberry mark in the newborn. *Br J Dermatol* 1972;**87**:138–44	Wrong study design: wrong population, 125 intervention and study subjects Paper waiting to be received

Hoffman JS, Kumar NB, Morley GW. Microinvasive squamous carcinoma of the vulva: search for a definition. *ObstetGynecol* 1983;**61**:615–18	
Hoffman MS, Roberts WS, Finan MA, Fiorica JV, Bryson SCP, Ruffolo EH, *et al*. A comparative study of radical vulvectomy and modified radical vulvectomy for the treatment of invasive squamous cell carcinoma of the vulva. *Gynecol Oncol* 1992;**45**:192–7	Wrong study population: only 64% of patients met inclusion criteria
Homesley HD, Bundy BN, Sedlis A, Adcock L. Radiation therapy versus pelvic node resection for carcinoma of the vulva with positive groin nodes. *Obstet Gynecol* 1986;**68**:733–40	Paper waiting to be received
Homesley HD, Bundy BN, Sedlis A, Yordan E, Berek JS, Jahshan A, *et al*. Assessment of current International Federation of gynecology and Obstetrics staging of vulvar carcinoma relative to prognostic factors for survival (A Gynecologic Oncology Group study). *Am J Obstet Gynecol* 1991;**164**:997–1004	More than 25% FIGO stage III and IV
Hopkins MP, Reid GC, Vettrano I, Morley GW. Squamous cell carcinoma of the vulva: prognostic factors influencing survival. *Gynecol Oncol* 1991;**43**:113–17	More than 25% FIGO stage III and IV
Hopkins MP, Reid GC, Morley GW. Radical vulvectomy, the decision for the incision. *Cancer* 1993;**72**:799–803	Same study as above
Hruby G, MacLeod C, Firth I. Radiation treatment in recurrent squamous cell cancer of the vulva. *Int J Radiat Oncol Biol Phys* 2000;**46**:1193–7	Wrong study population: patients with recurrent vulval cancer
Hunter DJS. Carcinoma of the vulva: a review of 361 patients. *Gynecol Oncol* 1975;**3**:117–23	Wrong study design: lack of information about FIGO or TNM stage
Husseinzadeh N, Wesseler T, Schneider D, Schellhas H, Nahhas W. Prognostic factors and the significance of cytologic grading in invasive squamous cell carcinoma of the vulva: a clinicopathologic study. *Gynecol Oncol* 1990;**36**:192–9	Wrong study design: risk factors evaluation
Iglesias AG, Lopez LCT, Sanchez MHG, Hernandez JH, Martin MJV, Perez JLL. Prognosis factors in cancer of the vulva. *Eur J Gynaecol Oncol* 1993;**14**:386–91	Paper waiting to be received
Iversen T, Aalders JG, Christensen A, Kolstad P. Squamous cell carcinoma of the vulva: a review of 424 patients, 1956–1974. *Gynecol Oncol* 1980;**9**:271–9	Wrong study population: only 5% of patients met inclusion criteria
Iversen T, Abelier V, Aalders J. Individualised treatment of stage I carcinoma of the vulva. *Obstetr Gynecol* 1981;**57**:85–9	Sample recruited before 1980
Iversen T, Abeler V, Kolstad P. Squamous cell carcinoma in situ of the vulva. A clinical and histopathological study. *Gynecol Oncol* 1981;**11**:224–9	Wrong study population: 80% of patients had carcinoma in situ
Jacek JS and Janusz E. Characteristic features of recurrences of squamous cell carcinoma of the vulva. *Ginekol Pol* 2010;**81**:12–19	Wrong study design: prognostic factors evaluation, not diagnostic accuracy
Janda M, Obermair A, Cella D, Crandon AJ, Trimmel M. Vulval cancer patients' quality of life: A qualitative assessment. *Int J Gynecol Cancer* 2004;**14**:875–81	Inadequate data presentation: all results were given for all patients enrolled in the study
Janda M, Obermair A, Cella D, Perrin LC, Nicklin JL, Ward BG, *etal*. The functional assessment of cancer-vulval: reliability and validity. *Gynecol Oncol* 2005;**97**:568–75	Wrong study population: 77.6% patients with SCC and 85% with adequate vulval tumour

Jeppesen JT, Sell A, Skjoldborg H. Treatment of cancer of the vulva. *Acta Obstet Gynecol Scand* 1972;**51**:101–7	Wrong study population: < 50% of patients met inclusion criteria
Johann S, Klaeser B, Krause T, Mueller MD. Comparison of outcome and recurrence-free survival after sentinel lymph node biopsy and lymphadenectomy in vulval cancer. *Gynecol Oncol* 2008;**110**:324–8	Wrong intervention: diagnostic treatment
Jolicoeur M, Nguyen TV, David S, Devieux A, Goffin F, Gauthier P, *et al*. *Conservative treatment for vulval cancer: chemoradiation and high dose rate brachytherapy*. European Society for Therapeutic Radiology and Oncology Joint Brachytherapy Meeting. Barcelona, Spain; May 2004	Wrong study design: poster
Judson PL, Jonson AL, Paley PJ, Bliss RL, Murray KP, Downs LS Jr, *et al*. A prospective, randomised study analyzing sartorius transposition following inguinal–femoral lymphadenectomy. *Gynecol Oncol* 2004;**95**:226–30	Wrong study population: only 52.4% of patients met inclusion criteria
Kacerovska D, Nemcova J, Petrik R, Michal M, Kazakov DV. Lymphoepithelioma-like carcinoma of the Bartholin gland. *Am J Dermatopathol* 2008;**30**:586–9	Wrong study design: case report
Kaltenbach FJ, Keil G. Autoradiographic and histologic observations on vulva carcinoma under local bleomycin treatment. *Strahlentherapie* 1979;**75**:185–90	Paper waiting to be received
Kaya S, Grillo M, Gent HJ. Results of the various treatment methods in vulval cancer. *Arch Gynecol Obstet* 1991;**250**:127–9	Wrong study population: included patients after secondary treatment
Keys H. Gynecologic Oncology Group randomised trials of combined technique therapy for vulval cancer. *Cancer* 1993;**71**:1691–6	Wrong study population: 55% of patients met inclusion criteria
Khobjai A, Srisomboon J, Charoenkwan K, Phongnarisorn C, Suprasert P, Siriaree S, *et al*. Radical surgery for T1 and T2 squamous cell carcinoma of the vulva through separate incisions. *J Med Assoc Thai* 2005;**88**(Suppl. 2):75–81	Paper waiting to be received
Kirby TO, Rocconi RP, Numnum TM, Kendrick JE, Wright J, Fowler W, *et al*. Outcomes of Stage I/II vulval cancer patients after negative superficial inguinal lymphadenectomy. *Gynecol Oncol* 2005;**98**:309–12	Wrong study design: letter to authors
Knopp S, Holm R, Trope C, Nesland JM. Occult lymph node metastases in early stage vulval carcinoma patients. *Gynecol Oncol* 2005;**99**:383–7	Wrong study design: lack of information about proportion between patient in stage I, II and III
Kodama S, Kaneko T, Saito M, Yoshiya N, Honma S, Tanaka KA. A clinicopathologic study of 30 patients with Paget's disease of the vulva. *Gynecol Oncol* 1995;**56**:63–70	Wrong study design: case report
Kohler U, Schone M, Pawlowitsch T. Results of an individualized surgical treatment of carcinoma of the vulva from 1973 to 1993. *Zentralbl Gynakol* 1997; **119**(Suppl. 1):8–16	Inadequate data presentation: cumulative for all patients (in all stages)
Konefka T, Olszewski J, Makarewicz H, Emerich J. Analysis of intra- and postoperative complications and postoperative course in patients surgically treated for vulval cancer. *Przeglad Lekarski* 1999;**56**:100–3	Inadequate data presentation and population: only 60% of patients met inclusion criteria
Kouvaris JR, Kouloulias VE, Kondi-Pahpiti A, Kokakis JD, Vlahos LJ. Impact of inguinal dissection on prognosis of early-stage squamous cell carcinoma of the vulva – a retrospective analysis 1768. *Onkologie* 2003;**26**:564–7	Paper waiting to be received

Kouvaris JR, Kouloulias VE, Plataniotis GA, Balafouta EJ, Vlahos LJ. Dermatitis during radiation for vulval carcinoma: Prevention and treatment with granulocyte–macrophage colony-stimulating factor impregnated gauze. *Wound Repair Regen* 2001;**9**:187–93	Wrong study population: 22.9% of patients met inclusion criteria
Kouvaris J, Kouloulias V, Loghis C, Sykiotis C, Balafouta M, Vlahos L. Prognostic factors for survival in invasive squamous cell vulval carcinoma: A univariate analysis. *Gynecol Obstet Invest* 2001;**51**:262–5	Inadequate data presentation: information only about statistical analyses
Kraemer B, Guengoer E, Solomayer EF, Wallwiener D, Hornung R. Stage I carcinoma of the Bartholin's gland managed with the detection of inguinal and pelvic sentinel lymph node. *Gynecol Oncol* 2009;**114**:373–4	Wrong study design: case report
Kreienberg R, Beck T, Bartzke G, Henne M and Friedberg F. Results of surgical treatment of carcinoma of the vulva. *Geburtsh Frauenheilk* 1990;**50**:375–82	Wrong study design and data presentation
Kubicki J, Samborska B, Lembrych S. Clinical analysis of 58 cases of vulvectomy. *Ginekol Pol* 1987;**58**:816–19	Wrong study population: only 17.9% of patients had vulval cancer
Kucera H, Weghaupt K. Radical vulvectomy using warm knife and irradiation of the inguinal lymph nodes for invasive squamous cell carcinoma of the vulva. *Geburtsh Frauenheilk* 1986;**46**:595–600	Wrong study population: only 58.2% of patients met the inclusion criteria
Kucera H. Treatment of carcinoma of the vulva at the 1st University Clinic of Gynecology in Vienna (386 cases). *Strahlentherapie* 1980;**156**:598–600	Wrong study design: literature review
Kunos C, Simpkins F, Gibbons H, Tian C, Homesley H. Radiation therapy compared with pelvic node resection for node-positive vulval cancer: a randomised controlled trial. *Obstet Gynecol* 2009;**114**:537–46	Wrong study population: only patients with groin nodes tumour
Kuppers V, Bender HG. Principles of surgical treatment of vulval cancer and precancerous lesions. *Gynakologe* 1993;**26**:293–7	Wrong study design: case report
Kurzl R, Messerer D. Prognostic factors in squamous cell carcinoma of the vulva: a multivariate analysis. *Gynecol Oncol* 1989;**32**:143–50	Wrong study design and unclear information: only information about node status was given without any others clinical descriptions
Lahousen M, Pickel H. About treatment of vulval carcinoma. *Zentralbl Gynakol* 1988;**110**:1001–5	Wrong study population: 66% of patients met inclusion criteria
Lahousen M. Invasive vulvar cancer: treatment and results. *Arch Gynecol Obstet* 1989;**245**:517–23	Same study as above
Landrum LM, Lanneau GS, Skaggs VJ, Gould N, Walker JL, McMeekin DS, *et al*. Gynecologic Oncology Group risk groups for vulval carcinoma: Improvement in survival in the modern era. *Gynecol Oncol* 2007;**106**:521–5	Wrong study population and data presentation
Lanneau GS, Argenta PA, Lanneau MS, Riffenburgh RH, Gold MA, McMeekin DS, *et al*. Vulval cancer in young women: demographic features and outcome evaluation editorial comment. *Obstet Gynecolo Surv* 2009;**64**:661–2	Wrong study design: analyses of risk factors
Lanneau GS, Argenta PA, Lanneau MS, Riffenburgh RH, Gold MA, McMeekin DS, *et al*. Vulval cancer in young women: demographic features and outcome evaluation. *Am J Obstet Gynecol* 2009;**200**:645.e1–5	More than 25% patients FIGO stage III and IV
Lasser A, Cornog JL, Morris JM. Adenoid squamous cell carcinoma of the vulva. *Cancer* 1974;**33**:224–7	Wrong study design: no information about FIGO stage, some of the patients had other tumour

Le T, Elsugi R, Hopkins L, Faught W, Fung-Kee-Fung M. The definition of optimal inguinal femoral nodal dissection in the management of vulva squamous cell carcinoma. *Ann Surg Oncol* 2007;**14**:2128–32	Paper waiting to be received
Leminen A, Forss M, Paavonen J. Wound complications in patients with carcinoma of the vulva: Comparison between radical and modified vulvectomies. *Eur J Obstet Gynecol Reprod Biol* 2000;**93**:193–7	Inadequate data presentation: all results were presented for all cumulative patients
Leuchter RS, Hacker NF, Voet RL, Berek JS, Townsend DE, Lagasse LD. Primary carcinoma of the Bartholin gland: a report of 14 cases and review of the literature. *Obstet Gynecol* 1982;**60**:361–8	Wrong study design: case report
Levato F, Bianchi A, Lenzi B, Pansini F, Randazzo F, Ferretti S, Grandi E, Mollica G. Surgical treatment of invasive carcinoma of the vulva. A reappraisal. *Eur J Gynaecol Oncol* 1992;**13**:99–104	Paper waiting to be received
Levenback CF, Tian C, Coleman RL, Gold MA, Fowler JM, Judson PL. Sentinel node (SN) biopsy in patients with vulval cancer: a Gynecologic Oncology Group (GOG) study. *J Clin Oncol* 2009;**27**:5505	Paper waiting to be received
Levin AO, Carpenter KM, Fowler JM, Brothers BM, Andersen BL, Maxwell GL. Sexual morbidity associated with poorer psychological adjustment among 127 gynaecological cancer survivors. *Int J Gynecol Cancer* 2010;**20**:461–70	Wrong study population: only 11.6% of patients were after vulval carcinoma treatment
Lifshitz S, Savage JE, Yates SJ, Buchsbaum HJ. Primary epidermoid carcinoma of the vulva. *Surg Gynecol Obstet* 1982;**155**:59–61	Wrong study population: lack of any clinical data
Lin JY, DuBeshter B, Angel C, Dvoretsky PM. Morbidity and recurrence with modifications of radical vulvectomy and groin dissection. *Gynecol Oncol* 1992;**47**:80–6	More than 25% patients FIGO stage III and IV
Lobraico RV, Waldow SM, Harris DM, Shuber S. Photodynamic therapy for cancer of the lower female genital tract. *Colposcopy Gynecol Laser Surg* 1986;**2**:185–99	Paper waiting to be received
Luo B. Treatment and analysis of 54 cases of vulval carcinoma. *Chung-Hua Fu Chan Ko Tsa Chih [Chinese J Obstet Gynecol]* 1990;**25**:156–8	Paper waiting to be received
Magrina JF, Gonzalez-Bosquet J, Weaver AL, Gaffey TA, Leslie KO, Webb MJ, *et al*. Squamous cell carcinoma of the vulva stage IA: long-term results. *Gynecol Oncol* 2000;**76t**:24–7	Wrong study population: 20% patients after primary treatment
Magrina JF, Gonzalez-Bosquet J, Weaver AL, Gaffey TA, Webb MJ, Podratz KC, *et al*. Primary squamous cell cancer of the vulva: radical versus modified radical vulvar surgery. *Gynecol Oncol* 1998;**71**:116–21	More than 25% patients FIGO stage III and IV
Magrina JF, Webb MJ, Gaffey TA, Symmonds RE. Stage I squamous cell cancer of the vulva. *Am J Obstet Gynecol* 1979;**134**:453–9	Wrong study population: patients with other malignancies
Mak RH, Halasz LM, Tanaka CK, Ancukiewicz M, Schultz DJ, Russell AH, *et al*. Outcomes after radiation therapy with concurrent weekly platinum-based chemotherapy or every-3–4-week 5-fluorouracil-containing regimens for squamous cell carcinoma of the vulva. *Gynecol Oncol* 2011;**120**:101–7	Wrong study population: 73% of patients with advanced disease
Makinem J, Salmi T, Gronroos M. Individually modified treatment of invasive squamous cell vulvar cancer: 10 year experience. *Ann Chirurg Gynaecol* 1987;**76** (Suppl. 202): 68–71	More than 25% patients FIGO stage III and IV

Malfetano JH, Piver MS, Tsukada Y, Reese P. Univariate and multivariate analyses of 5-year survival, recurrence, and inguinal node metastases in stage I and II vulval carcinoma. *J Surg Oncol* 1985;**30**:124–31	Paper waiting to be received
Malmstrom H, Janson H, Simonsen E, Stenson S, Stendahl U. Prognostic factors in invasive squamous cell carcinoma of the vulva treated with surgery and irradiation. *Acta Oncol* 1990;**29**:915–19	Wrong study design: no data which could be extracted
Manci N, Marchetti C, Esposito F, De Falco C, Bellati F, Giorgini M, *et al*. Inguinofemoral lymphadenectomy: Randomised trial comparing inguinal skin access above or below the inguinal ligament. *Ann Surg Oncol* 2009; **16**:721–8	Paper waiting to be received
Manci N, Marchetti C, Esposito F, *et al*. Inguinofemoral lymphadenectomy: randomised trial comparing inguinal skin access above or below the inguinal ligament. *Ann Surg Oncol*. 2009;**16**:721–8	Wrong study population: included patients with advanced disease
Maricic Z, Kolaric K, Krusic J. Our experience in the treatment of carcinoma of the vulva. *Strahlentherapie* 1976;**151**:495–503	Wrong study design and population: only 42% of patients met the inclusion criteria
Martinez-Palones JM, Perez-Benavente MA, Gil-Moreno A, Diaz-Feijoo B, Roca I, Garcia-Jimenez A, *et al*. Comparison of recurrence after vulvectomy and lymphadenectomy with and without sentinel node biopsy in early stage vulval cancer. *Gynecol Oncol* 2006;**103**: 865–70	Wrong intervention: only diagnostic treatment
Marzetti L, Framarino d, Tagliaferri T, Khosravi L, Paolillo MF. Carcinoma of the vulva: our experience. *Gior Ital Oncol* 1989;**9**:35–8	Paper waiting to be received
Matkowski R, Dryl J, Kornafel J. Analysis of treatment results of vulval cancer. *Ginekol Pols* 2004;**75**:720–8	Inadequate data presentation: results were presented in a cumulative way for all FIGO stages
Meriggi G. Primary carcinoma of the vulva. Clinical study. *Quad Clin Ostet Ginecol* 1972;**27**:253–60	Paper waiting to be received
Micheletti L, Borgno G, Barbero M, Preti M, Cavanna L, Nicolaci P, *et al*. Deep femoral lymphadenectomy with preservation of the fascia lata: Preliminary report on 42 invasive vulval carcinomas. *J Reprod Med Obstet Gynecol* 1990;**35**:1130–3	Paper waiting to be received
Miecznikowski A, Starzewski J. Surgical treatment of vulval cancer. *Eur J Gynaecol Oncol* 1993;**14**:392–7	Paper waiting to be received
Miyazawa K, Nori D, Hilaris BS, Lewis J. Role of radiation therapy in the treatment of advanced vulval carcinoma. *J Reprod Med Obstet Gynecol* 1983;**28**:539–41	Paper waiting to be received
Mohr A, Rieken S, Hof H, Bischof M, Combs SE, Debus J, *et al*. Nodal state determines outcome in vulval cancer. *Int J Radiat Oncol Biol Phys* 2010;**78**:S418	Paper waiting to be received
Molinie V, Paniel BJ, Lessana-Leibowitch M, Moyal-Barracco M, Pelisse M, Escande JP. Paget's disease of the vulva. *Ann Dermatol Venereol* 1993;**120**:522–7	Wrong study design and data presentation
Montana GS, Thomas GM, Moore DH, Saxer A, Mangan CE, Lentz SS, *et al*. Preoperative chemo-radiation for carcinoma of the vulva with N2/N3 nodes: a gynaecologic oncology group study. *Int J Radiat Oncol Biol Phys* 2000;**48**:1007–13	Wrong study population: patients with advanced disease
Montemaggi P, Luzi S, Morganti AG, Smaniotto D. Combined ERT and IRT in carcinoma of the vulva. *Rays Int J Radiol Sci* 1991;**16**:48–52	Paper waiting to be received

Moore DH. Chemotherapy and radiation therapy in the treatment of squamous cell carcinoma of the vulva: are two therapies better than one? *Gynecol Oncol* 2009;**113**:379–83	Wrong study design: review
Moore RG, Robison K, Brown AK, DiSilvestro P, Steinhoff M, Noto R, *et al*. Isolated sentinel lymph node dissection with conservative management in patients with squamous cell carcinoma of the vulva: a prospective trial. *Gynecol Oncol* 2008;**109**:65–70	Wrong study design
Morris WIC. Experiences with vulval cancer. *J R Coll Surg Edinburgh* 1973;**18**:327–40	Paper waiting to be received
Moth I, Andreasson B, Jensen SB, Bock JE. Sexual function and somatopsychic reactions after vulvectomy. A preliminary report. *Danish Medical Bulletin* 1983;**30**(Suppl. 2):27–30	Wrong study design: no information about FIGO or TNM stage
Mulayim N, Silver DF, Schwartz PE, Higgins S. Chemoradiation with 5-fluorouracil and mitomycin C in the treatment of vulval squamous cell carcinoma. *Gynecol Oncol* 2004;**93**:659–66	Wrong study population: patients with advanced disease
Muller RP, Fischedick AR, Schnepper E. Clinical symptoms and high-voltage therapy (electron therapy) of the vulval carcinoma. *Strahlentherapie* 1982;**158**:594–7	Wrong study population: only 66% of patients met inclusion criteria
Nahhas WA and Brown M. Gynecologic surgery in the aged. *J Reprod Med Obstet Gynecol* 1990;**35**:550–4	Paper waiting to be received
Narendra H, Ray S, Rao L, Geetha V. Malignant extrarenal rhabdoid tumour of the vulva in an adult. *J Cancer Res Ther* 2010;**6**:82–5	Wrong study design: case report
Newcomb PA, Weiss NS, Daling JR. Incidence of vulval carcinoma in relation to menstrual, reproductive, and medical factors. *J Natl Cancer Inst* 1984;**73**:391–6	Paper waiting to be received
Nicoletto MO, Parenti A, Bianco PD, Lombardi G, Pedrini L, Pizzi S *et al*. Vulvar cancer: prognostic factors. *Anitcancer Res* 2010;**30**:2311–8	More than 25% patients FIGO stage III and IV
Nyberg RH, Iivonen M, Parkkinen J, Kuoppala T, Maenpaa JU. Sentinel node and vulval cancer: A series of 47 patients. *Acta Obstet Gynecol Scand* 2007;**86**:615–19	Wrong intervention: diagnostic treatment
Ofodile FA, Oluwasanmi JO. Post-circumcision epidermoid inclusion cysts of the clitoris. *Plast Reconstructive Surg* 1979;**63**:485–6	Paper waiting to be received
Onnis A, Marchetti M, Valente S: Surgical management of invasive vulval carcinoma. A new operative technique 'non mutilant radical vulvectomy'. *Eur J Gynaecol Oncol* 1980;**1**:45–51	Paper waiting to be received
Oonk MHM, van de Nieuwenhof HP, de Hullu JA, Van der Zee AGJ. The role of sentinel node biopsy in gynecological cancer: A review. *Curr Opin Oncol* 2009;**21**:425–32	Wrong study design: diagnostic review
Oonk MHM, van Os MA, de Bock GH, de Hullu JA, Ansink AC, Van der Zee AGJ. A comparison of quality of life between vulval cancer patients after sentinel lymph node procedure only and inguinofemoral lymphadenectomy. *Gynecol Oncol* 2009;**113**:301–5	Wrong study design: diagnostic study
Oonk MH, van Hemel BM, Hollema H, de Hullu JA, Ansink AC, Vergote I, *et al*. Size of sentinel-node metastasis and chances of non-sentinel-node involvement and survival in early stage vulval cancer: Results from GROINSS-V, a multicentre observational study. *Lancet Oncol* 2010; **11**:646–52	Wrong study design: diagnostic study

Origoni M, Dindelli M, Ferrari D, Frigerio L, Rossi M, Ferrari A. Surgical staging of invasive squamous cell carcinoma of the vulva. Analysis of treatment and survival. *Int Surg* 1996;**81**:67–70	Paper waiting to be received
Paley PJ, Johnson PR, Adcock LL, Cosin JA, Chen MD, Fowler JM *et al*. The effect of sartorius transposition on wound morbidity following inguinal–femoral lymphadenectomy. *Gynecol Oncol* 1997;**64**:237–41	Wrong study design: no information about patients' clinical stage
Palli C, Lluch J, Valero S. Sexuality, communication and emotions: a situational study in women affected by gynecologic cancer. *Psicooncologia* 2010;**7**:153–73	Paper waiting to be received
Papalas J, Selim M, Lewin M. Granular cell tumour of the VUVLA: A clinicopathologic study of 16 lesions. *J Cutaneous Pathol* 2009;**36**:112	Wrong study population: patients with granular and Schwann cell tumours
Parker RT, Duncan I, Rampone J, Creasman W. Operative management of early invasive epidermoid carcinoma of the vulva. *Am J Obstet Gynecol* 1975;**123**:349–55	Wrong study population and study design: 22% of patients had recurrent disease
Parmley TH, Woodruff JD, Julian CG. Invasive vulval Paget's disease. *Obstet Gynecol* 1975;**46**:341–6	Wrong study design and population: only seven patients had invasive vulval Paget's disease, other cases had intraepithelial changes
Parthasarathy A, Cheung MK, Osann K, Husain A, Teng NN, Berek JS, *et al*. The benefit of adjuvant radiation therapy in single-node-positive squamous cell vulval carcinoma. *Gynecol Oncol* 2006;**103**:1095–9	Wrong study population: all patients had FIGO stage III disease
Patsner B, Mann J. Radical vulvectomy and 'sneak' superficial inguinal lymphadenectomy with a single elliptic incision. *Am J Obstet Gynecol* 1988;**158**:464–9	More than 25% patients FIGO stage III and IV
Pellegrino A, Fruscio R, Maneo A, Corso S, Battistello M, Chiappa V, *et al*. Harmonic scalpel versus conventional electrosurgery in the treatment of vulval cancer. *Int J Gynecol Obstet* 2008;**103**:185–8	Inadequate data presentation: all results were presented for all cumulative patients
Pengsaa P, Pothinam S, Udomthavornsuk B. Vulval carcinoma at Srinagarind Hospital, Khon Kaen, Thailand. *Eur J Gynaecol Oncol* 1993;**14**:56–62	Paper waiting to be received
Perez CA, Grigsby PW, Chao C, Galakatos A, Garipagaoglu M, Mutch D, *et al*. Irradiation in carcinoma of the vulva: factors affecting outcome. *Int J Radiat Oncol Biol Phys* 1998;**42**:335–44	More than 25% patients FIGO stage III and IV
Pescetto G. Criteria for the type of treatment of the vulval carcinoma. *Eur J Gynaecol Oncol* 1980;**1**:32–6	Paper waiting to be received
Petereit DG, Mehta MP, Buchler DA and Kinsella TJ. A retrospective review of nodal treatment for vulval cancer. *Am J Clin Oncol Cancer Clin Trials* 1993;**16**:38–42	Paper waiting to be received
Petereit DG, Mehta MP, Buchler DA, Kinsella TJ. Inguinofemoral radiation of N0,N1 vulval cancer may be equivalent to lymphadenectomy if proper radiation technique is used. *Int J Radiat Oncol Biol Phys* 1993; **27**:963–7	Paper waiting to be received
Petru E. Changes of the FIGO staging of gynecologic malignancies 2009. *Geburtsh Frauenheilk* 2010; **70**:269–72	Wrong study design: literature review
Petzel S, Gebre R, Austin S. Women's health-related quality of life and sexual health following diagnosis and surgical treatment for vulval cancer and vulva dysplasia. *Psycho-Oncology* 2007;**16**:S1–110	Paper waiting to be received

Pierson RL, Figge PK, Buchsbaum HJ. Surgery for gynecologic malignancy in the aged. *Obstet Gynecol* 1975;**46**:523–7	Wrong study population: only 66% of patients met inclusion criteria
Pirtoli L, Rottoli ML. Results of radiation therapy for vulval carcinoma. *Acta Radiol Ser Oncol Radiat Ther Phys Biol* 1982;**21**:45–8	Wrong study population and study design: poor information about patients' clinical stage, only few of them were described
Piura B, Glezerman M. Cancer of the vulva in South Israel. A clinico-pathologic study of 28 cases. *Cervix Lower Female Genital Tract* 1989;**7**:21–8	Paper waiting to be received
Pliskow S. Vulval Paget's disease. Clinicopathological review of 14 cases. *J Florida Med Assoc* 1990;**77**:667–71	Paper waiting to be received
Podratz KC, Symmonds RE, Taylor WF, Williams TJ. Carcinoma of the vulva: analysis of treatment and survival. *Obstet Gynecol* 1983;**61**:63–74	Inadequate data presentation: results presented cumulative way for II, III and IV FIGO stage
Puig-Tintore LM, Ordi J, Vidal-Sicart S, Lejarcegui J A, Torne A, Pahisa J, *et al*. Further data on the usefulness of sentinel lymph node identification and ultrastaging in vulval squamous cell carcinoma. *Gynecol Oncol* 2003;**88**:29–34	Wrong study design: diagnostic study
Rakar S. Surgical treatment of vulval cancer. *Eur J Gynaecol Oncol* 1992;**13**:319–21	Paper waiting to be received
Raspagliesi F, Hanozet F, Ditto A, Solima E, Zanaboni F, Vecchione F, *et al*. Clinical and pathological prognostic factors in squamous cell carcinoma of the vulva. *Gynecol Oncol* 2006;**102**:333–7	Wrong study design: no information about patients' FIGO stage disease
Ratner ES, Foran KA, Schwartz PE, Minkin MJ. Sexuality and intimacy after gynecological cancer. *Maturitas* 2010; **66**:23–6	Paper waiting to be received
Redman M, Alheit HD, Winkler C, Herrmann T. Results in radiotherapy of vulvar carcinoma. *Radiobiol Radiother* 1986;**27**:189–95	More than 25% patients FIGO stage III and IV
Reid GC, DeLancey JOL, Hopkins MP, Roberts JA, Morley GW. Urinary incontinence following radical vulvectomy. *Obstet Gynecol* 1990;**75**:852–8	Inadequate data presentation: all results were given for all study group, unclear group description
Rhodes C A, Cummins C, Shafi MI. The management of squamous cell vulval cancer: a population based retrospective study of 411 cases. *Br J Obstet Gynaecol* 1998;**105**:200–5	Wrong study population: 8% of patients had multifocal tumours
Roberts WS, Kavanagh JJ, Greenberg H, Bryson SC, LaPolla JP, Townsend PA, *et al*. Concomitant radiation therapy and chemotherapy in the treatment of advanced squamous carcinoma of the lower female genital tract. *Gynecol Oncol* 1989;**34**:183–6	Wrong study population: most patients had advanced disease
Robison K, Steinhoff MM, Granai CO, Brard L, Gajewski W, Moore RG. Inguinal sentinel node dissection versus standard inguinal node dissection in patients with vulval cancer: A comparison of the size of metastasis detected in inguinal lymph nodes. *Gynecol Oncol* 2006;**101**:24–7	Wrong study design: diagnostic study
Rodolakis A, Diakomanolis E, Voulgaris Z, Akrivos T, Vlachos G, Michalas S. Squamous vulvar cancer: a clinically based individualisation of treatment. *Gynecol Oncol* 2000; **78**:346–51	More than 25% patients FIGO stage III and IV
Rogers L, Howard B, Van Wijk L, Wei W, Dehaeck K, Soeters R, *et al*. Chemoradiation in advanced vulval carcinoma. *Int J Gynecol Obstet* 2009;**107**:S323	Inadequate data presentation: no possibilities to extract any data

Rosen C, Malmstrom H. Invasive cancer of the vulva. *Gynecol Oncol* 1997;**65**:213–17	Wrong study population: only 59% of patients met inclusion criteria
Rouzier R, Haddad B, Dubernard G, Dubois P, Paniel BJ. Inguinofemoral dissection for carcinoma of the vulva: Effect of modifications of extent and technique on morbidity and survival. *J Am Coll Surg* 2003;**196**:442–50	Wrong study design: no information for effectiveness
Rowley KC, Gallion HH, Donaldson ES, Van Nagell JR, Higgins RV, Powell DE, *et al*. Prognostic factors in early vulval cancer. *Gynecol Oncol* 1988;**31**:43–9	Wrong study design: no information about patients' FIGO stage of disease
Scheistroen M, Trope C. Combined bleomycin and irradiation in preoperative treatment of advanced squamous cell carcinoma of the vulva. *Acta Oncol* 1993;**32**:657–61	Wrong study population: only women with advanced disease
Schmeisser G, Krafft W, Schirmer A. Efficacy of supplementary treatment with bleomycin in carcinoma of the vulva. *Zentralbl Gynakol* 1979;**101**:350–3	Wrong study design: population was described in an unclear way
Schmidt W, Schmid H, Villena-Heinsen C, Kuhn W, Jochum-Merger N, Von Fournier D. Treatment results in malignant tumours of the vulva 1970–1990. *Geburtsh Frauenheilk* 1992;**52**:749–57	Wrong study population: < 70% of patients met inclusion criteria
Schnurch HG. Vulval cancer. diagnosis and treatment. *Urologe A* 2004;**43**:849–59	Wrong study design: review
Sedlis A, Homesley H, Bundy BN, Marshall R, Yordan E, Hacker N, *et al*. Positive groin lymph nodes in superficial squamous cell vulval cancer. *Am J Obstet Gynecol* 1987;**156**:1159–64	Wrong study population: 57.4% of patients met the inclusion criteria
Sengupta BS. Vulvar carcinoma in premenopausal Jamaican women. *Int J Gynaecol Obstet* 1980;**17**:526–30	More than 25% patients FIGO stage III and IV
Senn B, Mueller MD, Cignacco EL, Eicher M. Period prevalence and risk factors for postoperative short-term wound complications in vulval cancer: a cross-sectional study 55. *Int J Gynecol Cancer* 2010;**20**:646–54	Wrong study population: only 36.1% of patients met the inclusion criteria
Sesti F, Santeusanio G, Anemona L, Schiaroli S, Farne C, Piccione E. Verrucous carcinoma of the vulva. *J Obstet Gynaecol* 1991;**11**:79–80	Wrong study design: case report
Shepherd JH, Crowther ME. Complications of gynaecological cancer surgery: a review. *J R Soc Med* 1986;**79**:289–93	Paper waiting to be received
Sherman KJ, Daling JR, Chu J, Weiss NS, Ashley RL, Corey L. Genital warts, other sexually transmitted diseases, and vulval cancer. *Epidemiology* 1991;**2**:257–62	Paper waiting to be received
Shimm DS, Fuller AF, Orlow EL. Prognostic variables in the treatment of squamous cell carcinoma of the vulva. *Gynecol Oncol* 1986;**24**:343–58	Wrong study population: 66% patients met inclusion criteria, 86% had primary disease
Simonsen E, Johnsson JE and Trope C. Radical vulvectomy with warm-knife and open-wound techniques in vulval malignancies. *Gynecol Oncol* 1984;**17**:22–31	Wrong study population: only 60% of patients met inclusion criteria
Simonsen E, Johnsson JE, Trope C. Stage I squamous cell carcinoma of the vulva. *Acta Radiol Oncol* 1984;**23**:443–8	Wrong study population: also included patients with other malignancies
Simonsen E, Nordberg UB, Johnsson JE. Radiation therapy and surgery in the treatment of regional lymph nodes in squamous cell carcinoma of the vulva. *Acta Radiol Oncol* 1984;**23**:433–42	Wrong study population: only 65% of patients met inclusion criteria

Slevin NJ, Pointon RCS. Radical radiotherapy for carcinoma of the vulva. *Br J Radiol* 1989;**62**:145–7	Wrong study population: e.g. patients after primary treatment
Smith AM and Rauld HF. Radical vulvectomy and lymphadenectomy in the management of vulval cancer: The experience of a district general hospital. *J Obstet Gynaecol* 1985;**6**:57–61	Paper waiting to be received
SmyczekGargya B, Volz B, Geppert M and Dietl J. A multivariate analysis of clinical and morphological prognostic factors in squamous cell carcinoma of the vulva. *Gynecol Obstet Invest* 1997;**43**:261–7	Wrong study population: only 53% of patients met the inclusion criteria
Song YN, Yang JX, Shen K, Huang HF, Pan LY. Clinical features of recurrence of vulval squamous cell carcinoma: Analysis of 18 cases. *Nat Med J China* 2008;**88**:1347–9	Wrong study design and data presentation
Stegner HE. Ultrastructure of preneoplastic lesions of the vulva. *J Reprod Med* 1986;**31**:815–20	Paper waiting to be received
Stehman FB, Bundy B N, Ball H, Clarke-Pearson DL, Lagasse LD. Sites of failure and times to failure in carcinoma of the vulva treated conservatively: A Gynecologic Oncology Group study. *Am J Obstet Gynecol* 1996;**174**:1128–33	Inadequate data presentation: mixed for all stages
Stellman RE, Goodwin JM, Robinson J, Dansak D, Hilgers RD. Psychological effects of vulvectomy. *Psychosomatics* 1984;**25**:779–83	Wrong study design: no information about FIGO stage
Stroup AM, Harlan LC, Trimble EL. Demographic, clinical, and treatment trends among women diagnosed with vulval cancer in the United States. *Gynecol Oncol* 2008; **108**:577–83	Inadequate data presentation: unclear information about treatment and clinical patients' stage
Sugawa T, Hashimoto M, Suzuki M. Clinical aspects and treatment of vulval cancer in Japan. *Acta Obstet Gynaecol Jpn* 1980;**32**:177–86	Paper waiting to be received
Suzuki M, Watanabe M, Sato A. Effect of bleomycin on 131 gynaecologic 1311 carcinoma. *Gann Monogr Cancer Res* 1976;**19**:221–30	Paper waiting to be received
Tamburini M, Filiberti A, Ventafridda V, De Palo G. Quality of life and psychological state after radical vulvectomy. *J Psychosom Obstet Gynecol* 1986;**5**:263–9	Paper waiting to be received
Tatra G, Caucig H. Therapeutic results in primary carcinoma of the vulva. *Wien Klin Wochenschr* 1973;**85**:429–32	Inadequate data presentation: no clear information about treatments and survival results, mixed data for patients with FIGO I, II and III
Tebes S, Cardosi R, Hoffman M. Paget's disease of the vulva. *Am J Obstet Gynecol* 2002;**187**:281–4	Wrong study design: lack of clinical information about patients disease stage
ten Bokkel Huinik WW. Chemotherapy and complications in gynaecologic cancer. *Curr Options Gynecol* 1991;**3**:930–2	Wrong study design: review
Terada KY, Shimizu DM, Jiang CS, Wong JH. Outcomes for patients with T1 squamous cell cancer of the vulva undergoing sentinel node biopsy. *Gynecol Oncol* 2006;**102**:200–3	Paper waiting to be received
Thomas G, Dembo A, DePetrillo A, Pringle J, Ackerman I, Bryson P, *et al*. Concurrent radiation and chemotherapy in vulval carcinoma. *Gynecol Oncol* 1989;**34**:263–7	Wrong study population: 60% patients with recurrent disease
Tombolini V, Raffetto N, Santarelli M, Valerianai M, Necozione S, Masedu F, *et al*. Carcinoma of the vulva: clinical results of exclusive and adjuvant radiotherapy. *Anticancer Res* 2005;**25**:3089–94	More than 25% patients FIGO stage III and IV

Trelford JD, Deer DA and Ordorica E. Ten-year prospective study in a management change of vulval carcinoma. *Am J Obstet Gynecol* 1984;**150**:288–96	Wrong study design: diagnostic review
Van de Wiel HBM, Weijmar Schultz WCM, Hallensleben A, Thurkow FG, Bouma J, Verhoeven AC. Sexual functioning of women treated for cancer of the vulva. *Sex Marit Ther* 1990;**5**:73–82	Paper waiting to be received
van der Velden J, Ansink AC. Re: 'Outcomes of stage I/II vulval cancer patients after negative superficial inguinal lymphadenectomy'. *Gynecol Oncol* 2006;**100**:219–20	Wrong study design: letter to authors
Van der Zee AGJ, Oonk MH, de Hullu JA, Ansink AC, Vergote I, Verheijen RH, *et al*. Sentinel node dissection is safe in the treatment of early-stage vulval cancer. *J Clin Oncol* 2008;**26**:884–9	Wrong study design: diagnostic study
Vavra N, Kucera H, Weghaupt K. Evaluation of different risk factors in the follow-up of stage I vulval carcinoma and their influence on the results of therapy. *Wien Klin Wochenschr* 1990;**102**:289–94	Inadequate data presentation: all results presented in cumulative way for all stages
Vidal-Sicart S, Puig-Tintore LM, Lejarcegui JA, Paredes P, Ortega ML, Munoz A, *et al*. Validation and application of the sentinel lymph node concept in malignant vulval tumours. *Eur J Nucl Med Mol Imaging* 2007;**34**:384–91	Wrong intervention/study design: diagnostic only
Volk M, Schmidt Matthiesen H. Treatment of vulval cancer. *Arch Gynakol* 1977;**223**:145–62	Wrong study design: unclear information about treatment and patients' clinical stage
Wagner W, Prott FJ, Weissmann J, Niewohner-Desbordes U, Ostkamp K, Alfrink M. Vulval carcinoma: a retrospective analysis of 80 patients. *Arch Gynecol Obstet* 1999; **262**:99–104	Inadequate data presentation: data format did not allow reader to obtain meaningful information
Weghaupt K. Electro-resection and -coagulation in the therapy of vulva carcinoma. *Fortschr Med* 1978;**96**:1629–34	Wrong study design: review
Weghaupt K. Therapy of vulval carcinoma and its results. *Arch Gynakol* 1974;**216**:151–66	Wrong study design: review
Weijmar S, van d, Bouma J, Janssens J, Littlewood J. Psychosexual functioning after the treatment of cancer of the vulva. A longitudinal study. *Cancer* 1990;**66**:402–7	Paper waiting to be received
Weijmar Schultz WCM, Wijma K, Van de Wiel HBM. Sexual rehabilitation of radical vulvectomy patients. A pilot study. *J Psychosom Obstet Gynecol* 1986;**5**:119–26	Wrong study design: small study group ($n = 10$), no information about FIGO sage
Wharton JT, Gallager S, Rutledge FN. Microinvasive carcinoma of the vulva. *Am J Obstet Gynecol* 1974;**118**:159–62	Wrong study design/unclear information: lack of information about patients' FIGO stage of disease
Weijmar Shultz WC, van der Wiel HB, Bouma J, Janssens J, Littlewood J. Psychosexual functioning after the treatment of cancer of the vulva. *Cancer* 1990;**66**:402–7	Sample too small
Wolber L, Mahner S, Eulenburg C, Choschzick M, Hager M, Kock L, *et al*. Relevance of the number of metastatic lymph-nodes for disease recurrence in vulval cancer. *Arch Gynecol Obstet* 2010;**282**:S194	Paper waiting to be received
Wydra D, Emerich J, Ciach K, Sawicki S, Marciniak A. The role of pelvic exenteration for treatment of pelvic malignancy – a nine-year experience. *Eur J Gynaecol Oncol* 2005;**26**:418–22	Wrong study design: wrong study question
Wydra D, Matuszewski R, Romanowicz G, Bandurski T. Evaluation of surgical gamma probes for sentinel node localisation in cervical and vulval cancer. *Nucl Med Rev* 2005;**8**:105–10	Paper waiting to be received

Yamashiro T, Hasumi K, Masubuchi K. Clinical study of 84 primary vulval carcinomas. *Jpn J Cancer Chemother* 1989;**16**:1677–82	Paper waiting to be received
Yamashiro T, Teshima H, Yokosuka K, Shimizu Y, Hirai Y, Chen JT, *et al*. Clinical study of the vulval Paget's disease: Report of 8 cases. *J Jpn Soc Cancer Ther* 1989;**24**:2563–8	Paper waiting to be received
Yoder BJ, Rufforny I, Massoll NA, Wilkinson EJ. Stage IA vulval squamous cell carcinoma: An analysis of tumour invasive characteristics and risk. *Am J Surg Pathol* 2008;**32**:765–72	Wrong study design: no information about treatment and effectiveness
Yoonessi M, Goodell T, Satchidanand S. Microinvasive squamous carcinoma of the vulva. *J Surg Oncol* 1983;**24**:315–21	Wrong study design: unclear information about patients' stage, intervention
Zerner J. Basal cell carcinoma of the vulva. A report of six cases and review of the literature. *J Maine Med Assoc* 1974;**65**:127–9	Paper waiting to be received
Zhang SH, Sood AK, Sorosky JI, Anderson B, Buller RE. Preservation of the saphenous vein during inguinal lymphadenectomy decreases morbidity in patients with carcinoma of the vulva. *Cancer* 2000;**89**:1520–5	Wrong study design: inadequate study question

Appendix 12 Additional data from effectiveness systematic review

TABLE 56 Adverse events reported in the case series (fuller details)

Study	Intervention	Time of AE	AEs: *n* (%)	
Helm *et al.*, 1992[37]	Single incision	Early surgical complication/ RT	Wound breakdown/rupture: 6 (19)[a]	
			Groin breakdown: 11 (34)[b]	
		Reaction	Cellulitis: 5 (16)	
			Seroma: 6 (19)	
			Major complication: 2 (6)[c]	
	Triple incision	Early surgical complication/ RT	Wound breakdown/rupture: 2 (6)[a]	
			Groin breakdown: 6 (19)[b]	
		Reaction	Cellulitis: 7 (22)	
			Seroma: 10 (31)	
			Major complication: 1 (3)[c]	
Katz *et al.*, 2003[90]	Without IFL	Late surgical complication/ RT	Major complication: 2 (1)[d]	Lymphocyst or lymphoedema: 6 (5)[d]
		Reaction		Skeletal complications: 2 (2)[e]
	With IFL and RT			Lymphocyst or lymphoedema: 4 (7)[e]
				Skeletal complications: 3 (5)[f]
	RT		Skeletal complications: 8 (16)[f]	
Hallak *et al.*, 2007[98]	All interventions	Total number of patients who experienced an AE	64 (21.8)	
	After IFL	Early surgical complications	Wound breakdown/rupture: 5 (19); wound infection: 7 (26); bleeding: 2 (7); thrombosis: 0 (0); none: 15 (56)	
		Late surgical complications	Lymphoedema: 1 (7); none: 20 (91); erysipelas: 0 (0); vaginal stenosis: 0 (0); other: 1 (5)	
	Without IFL	Early surgical complications	Wound breakdown/rupture: 11 (6); wound infection: 10 (5); bleeding: 4 (2), thrombosis: 2 (1); none: 141 (77)	
		Late surgical complications	Lymphoedema: 0 (0); none: 161 (96); erysipelas: 1 (1); vaginal stenosis: 2 (1); other: 3 (2)	

continued

TABLE 56 Adverse events reported in the case series (fuller details) (*continued*)

Study	Intervention	Time of AE	AEs: *n* (%)
Vavra *et al.*, 1990[102]	All interventions	Total	14 (13.8)
		Early surgical complications	Wound breakdown/rupture: 2 (2)[g]; postoperative urinary incontinence: 4 (4)[g]; major complication: 2 (2)[g]
		Late surgical complications	Lymphoedema: 1 (1)[g]; vaginal stenosis: 4 (4)[g]

a In Helm *et al.*,[37] wound complications included: major breakdown in three patients (19%) in the single incision group and two patients (6%) in the triple incision group. Minor breakdowns occurred in three patients (9%) and one patient (3%), respectively.
b In Helm et al.,[37] major groin breakdown was experienced in six groin patients (19%) in the single incision group and one patient (3%) in the triple incision group. Minor breakdown occurred in five patients (16%) and five patients (16%), respectively.
c In Helm *et al.*,[37] major complications in the single incision group were massive haematemesis associated with portal hypertension and another was a ruptured femoral artery (both patients died). The major complication in the triple incision group was a non-fatal stroke (experienced by one patient).
d In Katz *et al.*,[90] two patients died due to postoperative cardiopulmonary complications, categorised as early complications.
e In Katz *et al.*,[90] those patients who experienced lymphocyst and lymphoedema also experienced neurogenic pain.
f In Katz *et al.*,[90] skeletal complications included hip fracture (in one patient) and fractures in osteoporotic bones (in one patient) in the IFL group; complications included fractures in osteoporosis bones (in three patients) in the IFL + RT group; complications included hip fracture or hip replacements (in six patients), asymptomatic pelvic insufficiency fracture (in one patient) and osteopenia and pain in multiple intra and extrapelvic site, including the hips (in one patient) in the RT group.
g In Vavra *et al.*,[102] two patients experienced pulmonary embolism, one of whom died on day 20.

TABLE 57 Complication after RT reported in case series

Study	Timing of AEs	AEs: *n* (%)
Hallak *et al.*, 2007[98]	Early complications	Acute erythema: 118 (51)
		Wet dermatitis: 72 (31)
		Cutaneous infection: 6 (3)
		Other (diarrhoea, urinary tract infection): 8 (4)
	Late complications	Cutaneous pigmentation: 87 (38)
		Fibrosis (inguinal region): 63 (27)
		Telangiectasia: 17 (7)
		Lymphoedema: 16 (7)[a]
		Thrombosis: 2 (1)
		Other: 9 (4)

a In Hallak *et al.*,[98] in two cases IFL was also performed and in three cases a more extended irradiation (including vulva and the pelvic region) was performed due to advanced disease.

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library