



Cochrane
Library

Cochrane Database of Systematic Reviews

Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer (Review)

Lawrie TA, Patel A, Martin-Hirsch PPL, Bryant A, Ratnavelu NDG, Naik R, Ralte A

Lawrie TA, Patel A, Martin-Hirsch PPL, Bryant A, Ratnavelu NDG, Naik R, Ralte A.
Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer.
Cochrane Database of Systematic Reviews 2014, Issue 6. Art. No.: CD010409.
DOI: [10.1002/14651858.CD010409.pub2](https://doi.org/10.1002/14651858.CD010409.pub2).

www.cochranelibrary.com

TABLE OF CONTENTS

ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
SUMMARY OF FINDINGS	4
BACKGROUND	6
Figure 1.	8
OBJECTIVES	9
METHODS	9
Figure 2.	10
RESULTS	12
Figure 3.	13
Figure 4.	15
Figure 5.	16
Figure 6.	19
Figure 7.	20
Figure 8.	21
Figure 9.	22
DISCUSSION	22
AUTHORS' CONCLUSIONS	24
ACKNOWLEDGEMENTS	24
REFERENCES	25
CHARACTERISTICS OF STUDIES	32
DATA	89
Test 1. Per groin: All tests.	91
Test 2. Per woman: All tests.	92
Test 3. Per groin: Blue dye only.	92
Test 4. Per woman: Blue dye only.	92
Test 5. Per groin: Tc-99m only.	93
Test 6. Per woman: Tc-99m only.	93
Test 7. Per groin: Combined test.	93
Test 8. Per woman: Combined test.	94
Test 9. Per groin: Mixed tests.	94
Test 10. Per woman: Mixed tests.	94
Test 11. Per groin: Tumour size < 4 cm.	95
Test 12. Per woman: Tumour size < 4 cm.	95
ADDITIONAL TABLES	95
APPENDICES	97
WHAT'S NEW	103
HISTORY	103
CONTRIBUTIONS OF AUTHORS	103
DECLARATIONS OF INTEREST	103
SOURCES OF SUPPORT	103
DIFFERENCES BETWEEN PROTOCOL AND REVIEW	104
INDEX TERMS	104

[Diagnostic Test Accuracy Review]

Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer

Theresa A Lawrie¹, Amit Patel², Pierre PL Martin-Hirsch³, Andrew Bryant⁴, Nithya DG Ratnavelu⁵, Raj Naik⁶, Angela Ralte⁷

¹Cochrane Gynaecological, Neuro-oncology and Orphan Cancer Group, Royal United Hospital, Bath, UK. ²Gynaecological Oncology, University Hospitals Bristol NHS Foundation Trust, Bristol, UK. ³Gynaecological Oncology Unit, Royal Preston Hospital, Lancashire Teaching Hospital NHS Trust, Preston, UK. ⁴Institute of Health & Society, Newcastle University, Newcastle upon Tyne, UK. ⁵Gynaecological Oncology, Northern Gynaecological Oncology Centre, Gateshead, UK. ⁶Northern Gynaecological Oncology Centre, Gateshead, UK. ⁷Pathology Department, Queen Elizabeth Hospital, Gateshead, UK

Contact: Amit Patel, Gynaecological Oncology, University Hospitals Bristol NHS Foundation Trust, St Michaels Hospital, Southwell Street, Bristol, BS2 8EG, UK. dr_amitpatel@hotmail.com.

Editorial group: Cochrane Gynaecological, Neuro-oncology and Orphan Cancer Group.

Publication status and date: Edited (no change to conclusions), published in Issue 9, 2016.

Citation: Lawrie TA, Patel A, Martin-Hirsch PPL, Bryant A, Ratnavelu NDG, Naik R, Ralte A. Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer. *Cochrane Database of Systematic Reviews* 2014, Issue 6. Art. No.: CD010409. DOI: [10.1002/14651858.CD010409.pub2](https://doi.org/10.1002/14651858.CD010409.pub2).

Copyright © 2016 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Vulval cancer is usually treated by wide local excision with removal of groin lymph nodes (inguinofemoral lymphadenectomy) from one or both sides, depending on the tumour location. However, this procedure is associated with significant morbidity. As lymph node metastasis occurs in about 30% of women with early vulval cancer, accurate prediction of lymph node metastases could reduce the extent of surgery in many women, thereby reducing morbidity. Sentinel node assessment is a diagnostic technique that uses traceable agents to identify the spread of cancer cells to the lymph nodes draining affected tissue. Once the sentinel nodes are identified, they are removed and submitted to histological examination. This technique has been found to be useful in diagnosing the nodal involvement of other types of tumours. Sentinel node assessment in vulval cancer has been evaluated with various tracing agents. It is unclear which tracing agent or combination of agents is most accurate.

Objectives

To assess the diagnostic test accuracy of various techniques using traceable agents for sentinel lymph node assessment to diagnose groin lymph node metastasis in women with FIGO stage IB or higher vulval cancer and to investigate sources of heterogeneity.

Search methods

We searched MEDLINE (1946 to February 2013), EMBASE (1974 to March 2013) and the relevant Cochrane trial registers.

Selection criteria

Studies that evaluated the diagnostic accuracy of traceable agents for sentinel node assessment (involving the identification of a sentinel node plus histological examination) compared with histological examination of removed groin lymph nodes following complete inguinofemoral lymphadenectomy (IFL) in women with vulval cancer, provided there were sufficient data for the construction of two-by-two tables.

Data collection and analysis

Two authors (TAL, AP) independently screened titles and abstracts for relevance, classified studies for inclusion/exclusion and extracted data. We assessed the methodological quality of studies using the QUADAS-2 tool. We used univariate meta-analytical methods to estimate pooled sensitivity estimates.

Main results

We included 34 studies evaluating 1614 women and approximately 2396 groins. The overall methodological quality of included studies was moderate. The studies included in this review used the following traceable techniques to identify sentinel nodes in their participants: blue dye only (three studies), technetium only (eight studies), blue dye plus technetium combined (combined tests; 13 studies) and various inconsistent combinations of these three techniques (mixed tests; 10 studies). For studies of mixed tests, we obtained separate test data where possible.

Most studies used haematoxylin and eosin (H&E) stains for the histological examination. Additionally an immunohistochemical (IHC) stain with and without ultrastaging was employed by 14 and eight studies, respectively. One study used reverse transcriptase polymerase chain reaction analysis (CA9 RT-PCR), whilst three studies did not describe the histological methods used.

The pooled sensitivity estimate for studies using blue dye only was 0.94 (68 women; 95% confidence interval (CI) 0.69 to 0.99), for mixed tests was 0.91 (679 women; 95% CI 0.71 to 0.98), for technetium only was 0.93 (149 women; 95% CI 0.89 to 0.96) and for combined tests was 0.95 (390 women; 95% CI 0.89 to 0.97). Negative predictive values (NPVs) for all index tests were > 95%. Most studies also reported sentinel node detection rates (the ability of the test to identify a sentinel node) of the index test. The mean detection rate for blue dye alone was 82%, compared with 95%, 96% and 98% for mixed tests, technetium only and combined tests, respectively. We estimated the clinical consequences of the various tests for 100 women undergoing the sentinel node procedure, assuming the prevalence of groin metastases to be 30%. For the combined or technetium only tests, one and two women with groin metastases might be 'missed', respectively (95% CI 1 to 3); and for mixed tests, three women with groin metastases might be 'missed' (95% CI 1 to 9). The wide CIs associated with the pooled sensitivity estimates for blue dye and mixed tests increased the potential for these tests to 'miss' women with groin metastases.

Authors' conclusions

There is little difference in diagnostic test accuracy between the technetium and combined tests. The combined test may reduce the number of women with 'missed' groin node metastases compared with technetium only. Blue dye alone may be associated with more 'missed' cases compared with tests using technetium. Sentinel node assessment with technetium-based tests will reduce the need for IFL by 70% in women with early vulval cancer. It is not yet clear how the survival of women with negative sentinel nodes compares to those undergoing standard surgery (IFL). A randomised controlled trial of sentinel node dissection and IFL has methodological and ethical issues, therefore more observational data on the survival of women with early vulval cancer are needed.

PLAIN LANGUAGE SUMMARY

Can tests used to identify the main groin lymph node/s in women with vulval cancer accurately predict whether the cancer has spread to the groin/s?

The issue

Women with vulval cancer that has spread to the groin lymph nodes need additional treatment. The standard treatment usually involves surgical removal of as many groin nodes as possible (known as complete inguinofemoral lymphadenectomy (IFL)). However, only about 30% of women with vulval cancer in whom lymph nodes are not obviously enlarged will have groin involvement; therefore, in about 70% of these women additional surgery is not necessary. As groin surgery often causes later swelling of the legs and other unpleasant side effects, it would be preferable not to undergo the surgery if it is not required; therefore, accurate screening tests to determine who should have surgery are needed.

Sentinel node assessment involves identifying the main lymph node/s draining the tumour. After the main (sentinel) nodes are identified, they are removed and examined under a microscope to check for cancer cells. Additional surgery depends on the findings of the examination: if cancer cells are found in the nodes, additional surgery is necessary; if the nodes are cancer-free, additional surgery can be avoided.

Why is this review important?

Several studies have been done using dyes or traceable agents to identify sentinel nodes. From these studies, it is not clear whether all of these agents are sufficiently accurate to predict which women have cancerous spread to the groin. This review summarises the evidence and produces overall estimates of the relative accuracies of the available tests.

How was the review conducted?

We included all studies that tested the accuracy of tracer agent/s against the standard method of identifying cancer in the groin nodes (removing all groin nodes (IFL) and examining them under a microscope). Women in these studies had vulval cancer of Federation of

Gynecology and Obstetrics (FIGO) stage IB or higher without obvious signs of cancer in the groin (enlarged or palpable nodes). We only included studies of at least 10 women, and noted any concerns about the quality of studies.

What are the findings?

We included 34 studies (1614 women) that evaluated three techniques: blue dye only, technetium (a radioactive substance) only, or blue dye and technetium combined. Ten studies used all three techniques during the course of the study (one technique per participant). There are two attributes to a test: the ability to identify or detect the sentinel node, and the ability to identify the cancer in the sentinel node. We found that all tests can identify cancer in the groin nodes with good accuracy (more than 90% of nodes with cancer will be accurately identified with any of the tests), although the combined test was the most accurate (95%). The ability of the tests to detect sentinel nodes varied, with the blue dye test only detecting sentinel nodes in 82% of women, compared with 98% for the combined test. If sentinel nodes are not detected, they cannot be examined for cancer cells; therefore, women in whom sentinel nodes are not detected will usually need to undergo IFL.

What does this mean?

The combined and technetium only tests are able to predict accurately which women have cancerous spread to the groin. For a group of 100 women undergoing assessment, the findings mean that approximately one or fewer women having the combined or technetium only tests will undergo an unnecessary IFL, compared with approximately 11 women having the blue dye only test. This is mainly because the blue dye only test is not as good as technetium in identifying sentinel nodes. Fewer women with spread to the groin will be missed with the combined or technetium only tests (1 to 3 out of 30) compared with the blue dye only test (1 to 8 out of 30). It is not clear whether women with negative sentinel nodes (i.e. no spread of cancer to the groin lymph nodes) who do not undergo IFL will live as long as those who undergo IFL. The current best data on survival come from a Dutch study that followed up 259 women with negative sentinel nodes and reported a three-year survival of 97%.

SUMMARY OF FINDINGS

Summary of findings 1. Summary of findings: Traceable agents for sentinel lymph node assessment in vulval cancer

Review question: how does the diagnostic test accuracy of various techniques using traceable agents for sentinel lymph node assessment in vulval cancer compare?

Patients or population: women with FIGO stage IB or higher vulval cancer without palpable/suspicious groin nodes

Settings: tertiary level hospitals

Role: to diagnose groin lymph node metastases

Index tests: blue dye, technetium, combined tests (blue dye and technetium) and mixed tests (blue dye, technetium or combined tests)

Reference standard: histological examination following complete inguinofemoral lymphadenectomy

Studies: prospective (30) and retrospective (4) cohort

Index test	Quantity of evidence	Mean detection rate*	Pooled sensitivity results per woman (95% CI)	Consequences in a cohort of 100 women undergoing SN assessment, assuming the prevalence of groin metastases to be 30%				
				No SNs detected** (undetected)	Women with metastatic nodes diagnosed by index test (TP)	Women with metastatic nodes missed by index test (FN)	Women requiring IFL***	Women not requiring IFL
1. Blue dye	68 women (3 studies)	82%	94% (69% to 99%)	18	23 (17 to 25)	2 (0 to 8)	41	59
2. Technetium	149 women (8 studies)	96%	93% (89% to 96%)	4	27 (26 to 28)	2 (1 to 3)	31	69
3. Combined tests (blue dye + technetium)	390 women (12 studies)	98%	95% (89% to 97%)	2	28 (26 to 29)	1 (1 to 3)	30	70
4. Mixed tests	679 women (7 studies)	95%	91% (71% to 98%)	5	26 (20 to 28)	3 (1 to 9)	32	68

Studies which employed 'mixed tests' used a combination of the index tests 1 to 3 and presented the overall results (i.e. did not present results separately for the different tests). The detection rate is the percentage of patients in which the test located a sentinel node. In patients where no node is detected, the test has no value.

*These mean detection rates are estimates derived from the total number of participants included in the studies for each test (see [Table 2](#)).

**Undetected women require complete inguino-femoral lymphadenectomy (IFL).

***Undetected women + correctly diagnosed women (TPs).

TP = true positives; FN = false negatives

BACKGROUND

Target condition being diagnosed

The target condition being diagnosed is groin lymph node metastases in women with vulval cancer. Vulval cancer is a rare gynaecological cancer with an incidence of 1 to 3 per 100,000 women per year (ONS 2009; Sankaranarayanan 2006; Saraiya 2008). At the time of diagnosis more than half of affected women are aged 70 years or above, and incidence peaks at the age of 75 years and above (ONS 2009; Sankaranarayanan 2006). However, recent epidemiological evidence from The Netherlands suggests that the incidence in younger women is increasing (Schuurman 2013). The majority (75% to 90%) of vulval cancers are squamous cell carcinomas (Saraiya 2008; Stehman 2007), which are staged according to the International Federation of Gynecology and Obstetrics (FIGO) classification (Table 1).

Groin lymph node metastasis is associated with reduced survival in women with vulval cancer and depends on the type, size and location of the vulval lesion (Andreasson 1985; Boyce 1985; Curry 1980; Homesley 1991; Parker 1975; Podratz 1983; Smyczek-Gargya 1997). The risk of groin metastases in women with apparent early-stage vulval cancer (stage IB/II) is approximately 30% (GROINSS-V 2008).

Primary vulval lesions are treated by wide local excision (WLE). Lesions smaller than 2 cm, with a depth of invasion less than 1 mm (FIGO stage IA), do not require removal of lymph nodes (inguinofemoral lymphadenectomy; IFL) from the groin due to the extremely low risk (less than 1%) of metastasis (Hacker 1993). However, in all other cases (FIGO stage IB and higher) removal of all groin lymph nodes (IFL) has been the traditional gold standard of treatment. Vulval tumours away from midline (lateralised) require removal of groin lymph nodes from the same side, whilst midline tumours require removal of lymph nodes from both sides (Hacker 1993; Iversen 1981). Most women with positive groin lymph nodes will require further treatment with radiation after surgery, with its risk of additional morbidity. This treatment approach is highly effective, with a low groin tumour recurrence rate of 1% to 10% (Burger 1995; Hacker 1981; Homesley 1991; Katz 2003). Effective treatment has resulted in a halving of vulval cancer mortality over the last three decades (ONS 2009). However, this treatment approach is associated with significant morbidity related to the wound and lymph drainage in up to 70% of cases (Fotiou 1996; Gaarenstroom 2003; Rouzier 2003; Stehman 1992; Van der Zee 2008). Short-term morbidity includes wound infection, wound disruption, groin lymph collection (lymphocyst) and longer hospital stay. Long-term morbidity includes chronic leg swelling (lymphoedema), chronic and recurrent skin infection (erysipelas) and reduced mobility.

Index test(s)

The lymphatic fluid from the vulval skin is drained by lymphatic channels to the groin lymph nodes. The first lymph node to receive these lymphatic channels on each side is considered to be the sentinel node. Cancer cells from a vulval tumour spread via lymph fluid through lymphatic channels usually to the sentinel node, before spreading to other nodes. A diagnostic test can therefore be employed to detect, excise and examine the sentinel node(s) histologically for cancer cells. This is usually achieved by injecting a traceable agent subcutaneously around the vulval tumour (usually

at four quadrants). This agent spreads via the lymphatic channels to the lymph node, which can be traced using an appropriate tracing method. The first lymph node in each groin region to concentrate the traceable agent is considered the sentinel lymph node. There are two attributes to a test: the ability to identify a sentinel node (detection), and the ability to identify the cancer in the sentinel node (diagnosis).

Various traceable agents and their detection techniques can be employed on their own or in combination. For instance, the most commonly used technique involves a combination of radioactive 99m technetium and patent blue dye. Technetium radiocolloid is injected at four quadrants of the vulval tumour a day before, or on the day of, surgery followed by a scan to detect the sentinel node(s) (lymphoscintigraphy), which are marked on the overlying skin. Patent blue dye is injected around the tumour immediately before surgery. A hand-held gamma camera probe to detect the concentration of the technetium and the visual discolouration of patent blue dye guides the surgeon during the operation to detect the sentinel node.

Once excised, the sentinel node is sent either for an immediate frozen section examination or for routine paraffin histology (which takes a few days to report). If the sentinel node is found to have cancer cells (positive sentinel node), further surgery to remove all remaining groin nodes (at the same time if reported on frozen section, or at a later date if on paraffin section) will be required. If the sentinel node is reported to be free of cancer cells (negative sentinel node), total removal of groin lymph node and its associated morbidity can be avoided. Ultrastaging techniques such as serial micro-sectioning (at 200 to 250 µm) and immunohistochemistry staining (usually for cytokeratin) are used to detect micro-metastasis (< 2 mm size) in the sentinel node if initial haematoxylin and eosin (H&E) section is negative (Knopp 2005). This has proven to increase the detection rate of lymph node metastasis (GROINSS-V 2008), but its significance in the overall prognosis in vulval cancer remains unclear. It is anticipated that use of ultrastaging will have a significant effect on the diagnostic accuracy of sentinel node analysis.

Sentinel node detection and analysis has been pioneered and has become the standard of care in the surgical management of melanoma and breast cancer (Canavese 2010; Krag 2007; Morton 1990; Morton 2006; Rodier 2007; Thompson 2007; Wang 2011). Success rates of sentinel node detection in vulval cancer with the combined use of 99m technetium and patent blue dye approach have been reported to be between 89% and 100% (de Hullu 2000; Tavares 2001). Failure to detect the sentinel node could be due to the agent failing to reach a sentinel node, too low concentration of agent in the lymph node, or the surgeon not being able to identify the sentinel node. In this situation it is advisable to undergo standard groin lymph node dissection (IFL). In those where sentinel node(s) are identified it is important that the false negative rate of groin lymph node metastasis (i.e. negative sentinel lymph node but presence of positive non-sentinel lymph nodes) is extremely low. A high false negative sentinel node rate will lead to poor outcomes due to avoidance of groin lymph node removal and radiation treatment in cases that would have actually benefited from these therapies.

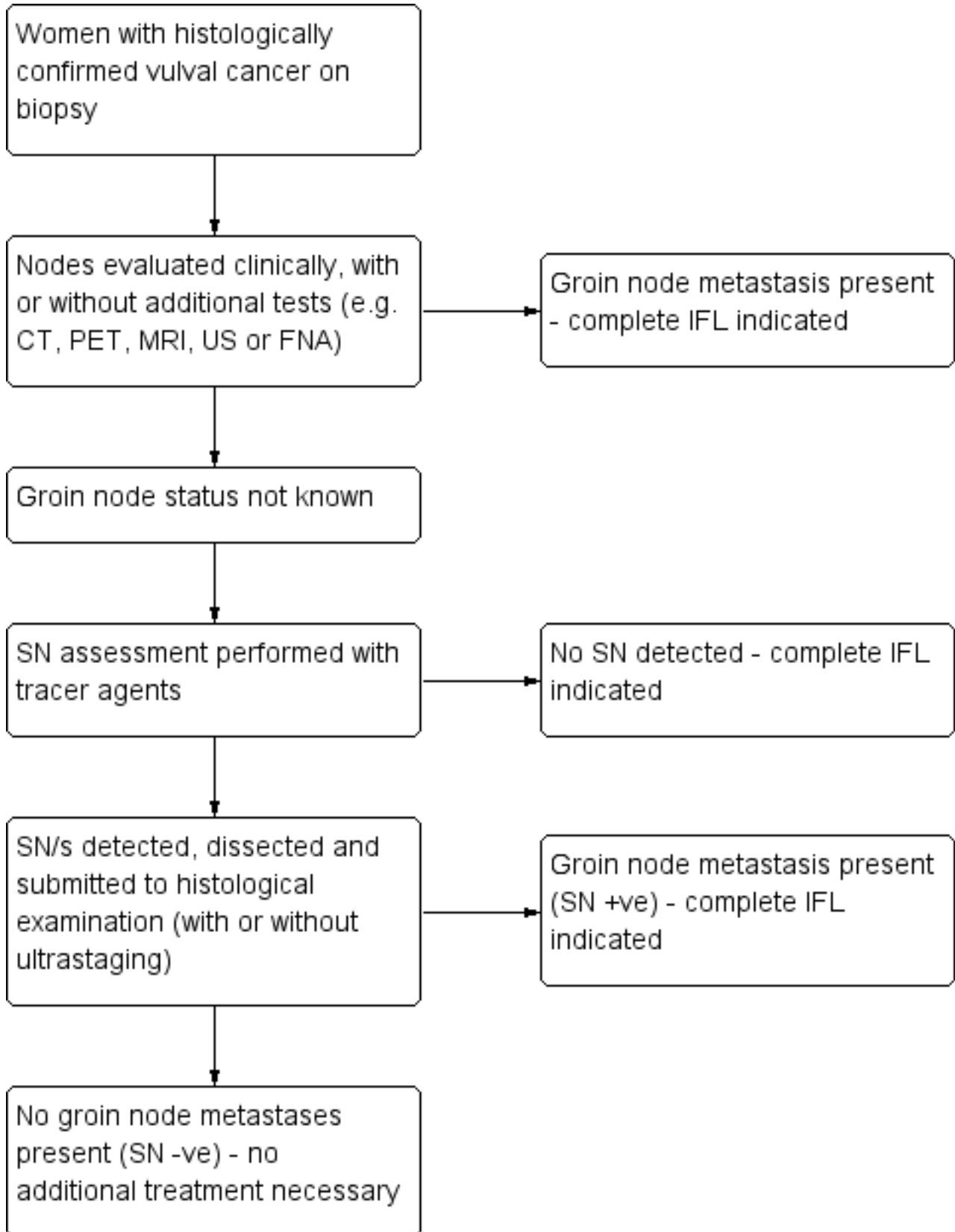
Sentinel node assessment is usually only used in cases where the vulval tumour size is less than 4 cm in maximum diameter with greater than 1 mm depth of invasion, and in cases where groin

lymph node metastasis is not suspected. The maximum tumour dimension of 4 cm, although arbitrarily chosen, is based on a relatively lower risk of lymph node metastasis ($\pm 30\%$; [GROINSS-V 2008](#)) and low failure rate to detect sentinel nodes. False negatives occur more frequently with tumours larger than 4 cm in size ([Levenback 2012](#)).

Clinical pathway

For the clinical pathway of women with early vulval cancer see [Figure 1](#).

Figure 1. Clinical pathway of women with ≥ FIGO stage IB vulval cancer SN sentinel node; CT computed tomography; PET positron emission tomography; MRI magnetic resonance imaging; US ultrasound; FNA fine needle aspiration; IFL inguinofemoral lymphadenectomy



Role of index test(s)

The role of the index test is to predict accurately groin lymph node metastases so that the extent of surgery can be reduced for women without metastases.

Alternative test(s)

Currently, there are no alternative diagnostic tests that predict groin lymph node metastases in vulval cancer with reasonable test accuracy. Various imaging techniques including ultrasound, computed tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography - computed tomography (PET-CT) have been used to evaluate groin lymph node status before definitive surgery. Although they have the advantage of being non-invasive, their ability to confirm (sensitivity) or exclude metastasis (specificity) is limited (Abang Mohammed 2000; Cohn 2002; de Hullu 1999; Hall 2003; Hawnaur 2002; Land 2006; Makela 1993; Moskovic 1999; Sohaib 2002), and therefore they are not routinely used in clinical practice.

Rationale

Surgical excision of tumour and lymphatic staging remains a cornerstone of the management in vulval cancer. For very early-stage disease (FIGO IA), WLE without lymphatic staging is an accepted method of treatment due to the low risk (less than 1%) of lymph node metastasis (Hacker 1993). For FIGO stage IB disease or above, a WLE of vulval tumour along with the removal of groin lymph nodes from one or both sides (depending on the tumour location) is the traditional treatment of choice (Hacker 1993; Iversen 1981). This treatment, however, is associated with significant morbidity related to wound and lymph drainage in up to 70% of cases (Gaarenstroom 2003; Rouzier 2003; Stehman 1992; Van der Zee 2008). The overall rate of lymph node metastasis in vulval cancer is reported to be 25% to 50% (Creasman 1997; Simonsen 1984; Sutton 1991). The node-negative cases are unlikely to benefit from removal of groin lymph nodes and many will suffer from unnecessary associated surgical morbidity. Most women with positive groin lymph nodes will require further treatment with radiation, with its risk of additional associated morbidity.

The concept of sentinel node detection and analysis has been successfully applied to guide the management of melanoma and breast cancer (Canavese 2010; Krag 2007; Morton 1990; Morton 2006; Rodier 2007; Thompson 2007). The surgical morbidity of axillary lymph node dissection has been reduced without adverse effect on breast cancer outcomes (Canavese 2010; Krag 2007; Rodier 2007; Wang 2011). A similar benefit is possible with sentinel node assessment in vulval cancer as well. The sentinel lymph node is the first lymph node in the groin region to which the vulval cancer cells would spread via the lymphatic channels. The histological analysis of the sentinel groin node is considered to be representative of all other remaining non-sentinel groin lymph nodes draining to the same anatomical side. The use of sentinel node assessment will therefore triage only those women with positive sentinel node for further groin node dissection, avoiding surgical morbidity in the remaining sentinel node-negative women. If sentinel node detection and analysis has very high sensitivity with an extremely low false negative rate in predicting groin lymph node metastasis, its use in routine clinical practice can be envisaged. This review aims to analyse the diagnostic accuracy of sentinel node assessment in vulval cancer.

OBJECTIVES

To assess the diagnostic test accuracy of various techniques using traceable agents for sentinel lymph node assessment to diagnose groin lymph node metastasis in women with FIGO stage IB or higher vulval cancer and to investigate sources of heterogeneity.

METHODS

Criteria for considering studies for this review

Types of studies

We included all prospective and retrospective studies that compared and reported diagnostic test accuracy statistics of sentinel node assessment (detection and histological examination) with the reference standard of histological examination of inguinofemoral lymph node dissection (IFL). We included studies that reported the number of sentinel node procedures (each side counted separately; so-called 'per groin' data) and the number of women who underwent a sentinel node procedure (whereby a bilateral sentinel node procedure was reported as one case; so-called 'per woman' data), provided that we could construct two-by-two tables of these data. We excluded studies reporting fewer than 10 sentinel node procedures, as well as studies for which construction of a two-by-two table for either 'per groin' or 'per woman' data was not possible. For studies that included women with clinically suspicious, palpable or metastases-positive groin nodes, we attempted to exclude these women from the extracted data. Where this was not possible, we excluded studies in which these cases exceeded 10% of the total numbers of women or groins assessed.

Participants

Women diagnosed with International Federation of Gynecology and Obstetrics (FIGO) stage IB or higher vulval cancer without clinically suspicious nodes. We considered all ages, histological types of tumour, and all techniques and settings of sentinel node detection, dissection and histological examination in this review.

Index tests

Tracer agents used to identify sentinel nodes for histological assessment.

We expected that trial reports should specify accurately the technique used and include the following:

- Description of agent
- Technique, amount, location and timing of injection of agent
- Method used to trace and detect sentinel node
- Definition of what was regarded as a sentinel node
- Description of the histological method used to assess the sentinel node

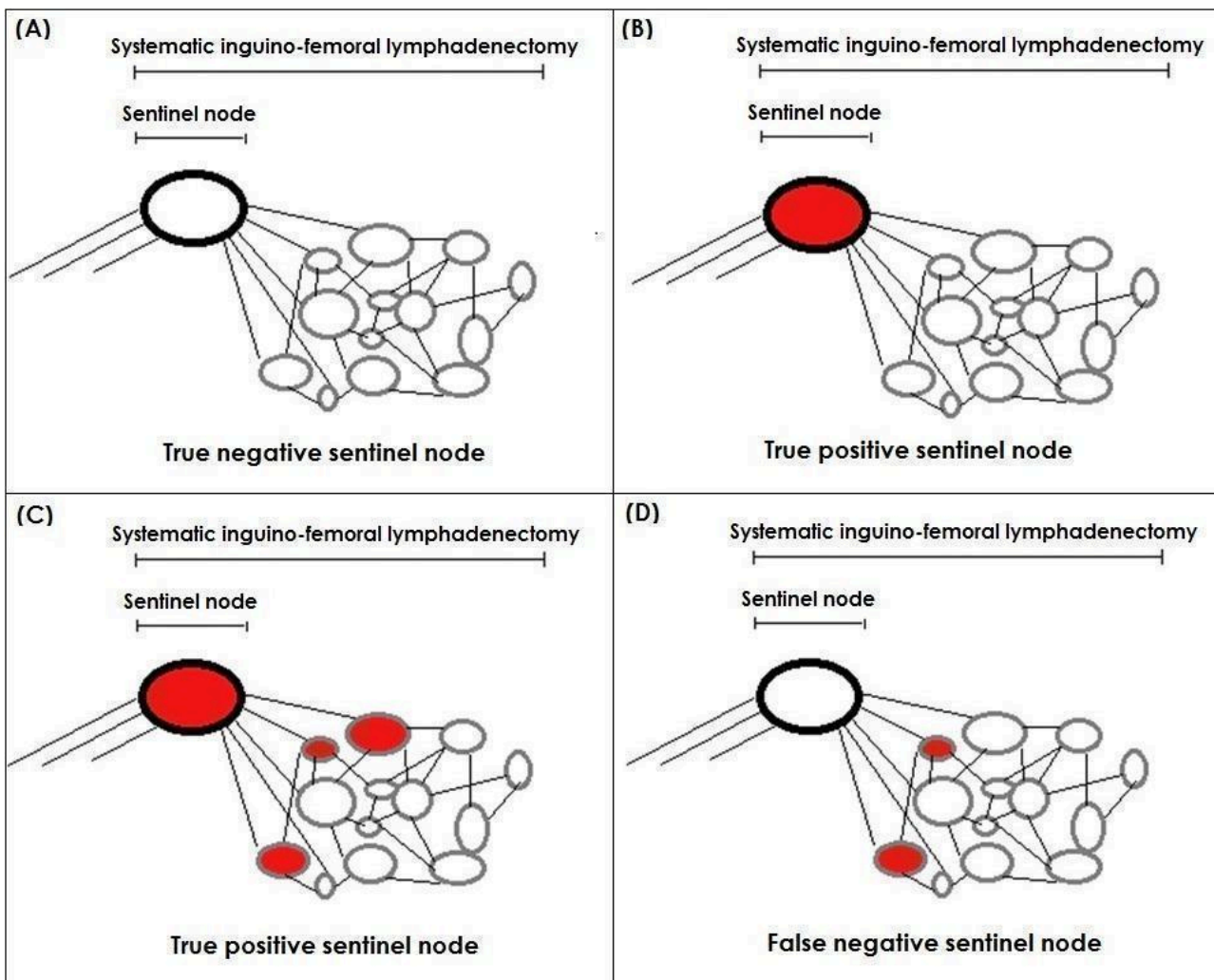
A sentinel node should have been defined as the first lymph node that showed adequate concentration of tracing agent (e.g. greatest radioactive signal in groin basin detected on hand-held gamma probe in case of radioactive tracer agent, or a node that appeared visually blue intra-operatively) (de Hullu 1998). The sentinel node should then have been removed and subjected to standard histological examination or by frozen section with at least

haematoxylin and eosin (H&E) staining. If the sentinel node was found to be malignant, it was defined as a positive sentinel node. If a sentinel node did not show any malignancy, it was defined as a negative sentinel node. If a sentinel node could not be identified, it was defined as 'failure to detect sentinel node' (and not index test negative). Details of reason for failure to detect sentinel nodes should also have been reported where possible, along with the outcome of lymph node status on reference standard.

A complete groin lymph node dissection (IFL) would include the sentinel node in the specimen, thus creating a situation where

the index test result becomes part of the reference standard (incorporation). Realistically, therefore, false positive tests would not exist in this situation (see Figure 2). In an event where the sentinel node was identified, assessed and deemed histologically positive but the remaining groin nodes were negative, the index test would still be regarded as a true positive. We did not anticipate that false positive tests would be reported in the included studies, but if we encountered them, unless further information was available from the author to create a protocol-compliant two-by-two table, we planned to exclude them from the review.

Figure 2. Possible outcomes of sentinel lymph node assessment followed by total groin lymph node removal (inguinofemoral lymphadenectomy). (A) Negative sentinel and rest of the groin nodes (True negative), (B) Positive sentinel node but negative rest of the groin nodes (True positive), (C) Positive sentinel and groin nodes (True positive) and (D) Negative sentinel but positive groin nodes (False negative)



We anticipated that many studies would report a combination of various sentinel node assessment techniques (mixed tests). When analysing the diagnostic test accuracy of a single technique, we only included these studies in the analyses if a separate two-by-two table for the technique in question could be constituted. Similarly, not all women received combined techniques (i.e. blue dye and technetium) in these studies. When analysing the diagnostic test accuracy of the combined technique, these studies were only

included in the analyses if all cases received the combined technique in question, or a separate two-by-two table for cases who received the combined technique in question could be constituted. Where possible, we attempted to obtain separate diagnostic test accuracy data from investigators of studies in which different index test data had been combined.

Target conditions

Groin (inguinofemoral) lymph node metastases in FIGO stage IB or higher vulval cancer.

Reference standards

Histological examination of systematic groin lymph node dissection (IFL) was the reference standard. The reference standard was to be subjected to the standard histological assessment with at least H&E staining. If any of the removed nodes (including the sentinel node) showed cancer metastasis histologically, the reference standard was considered positive. Studies were to report the reference standard result by each side of groin node removal or by women/cases.

Systematic groin lymph node removal includes removal of inguinal and femoral lymph nodes. Traditionally this includes removal of lymph nodes above and parallel to the inguinal ligament up to the pubic tubercle medially and lymph nodes from the femoral triangle (parallel to femoral vessels and sapheno-femoral junction including cribriform fascia) up to and including the deep fascia of muscle forming the base of the femoral triangle. Dissection deeper to the deep fascia or into the adductor canal is usually not required. However, there remains some uncertainty regarding the ideal extent and adequacy of surgical dissection ([Hudson 2004](#)).

Search methods for identification of studies

Electronic searches

The electronic searches were performed by the Diagnostic Test Accuracy Working Group Trial Search Co-ordinator, Anne Eisinga. This included searches of the following electronic databases:

- MEDLINE (OvidSP) (1946 to February 2013, week four);
- EMBASE (OvidSP) (1974 to March 2013, week 10).

The search strategies are outlined in the appendix ([Appendix 1](#)). As these searches would have identified any possible reviews on the topic we did not search other databases, e.g. DARE, as stated in the protocol. We did not apply language restrictions to the electronic searches and, where necessary, had non-English articles of relevant studies translated.

Searching other resources

We reviewed the reference list of all relevant studies retrieved from electronic searches and used the 'related articles' feature of PubMed to identify additional potentially relevant studies. We did not handsearch the conference proceedings of the International Gynaecological Cancer Society (IGCS), the European Society of Gynaecological Oncology (ESGO), the Society of Gynaecologic Oncologists (SGO) and the American Society of Clinical Oncologist (ASCO) from 2000 to the present as planned in the protocol, as abstracts for this period from these societies were identified by the electronic searches. Where the electronic searches identified conference abstracts in the absence of a full report, we attempted to contact the investigators for more information.

Data collection and analysis

Selection of studies

We used the reference manager software Endnote® to remove duplicates from all titles and abstracts retrieved from the literature

search ([Endnote 2012](#)). Amit Patel (AP) and Theresa Lawrie (TAL) independently examined all eligible references. We excluded studies that clearly did not meet the inclusion criteria and obtained full-text articles of those that appeared potentially relevant. AP and TAL independently assessed the full-text articles for their eligibility and in the event of disagreement involved other authors. We documented clearly the reasons for exclusion of potentially relevant studies.

Data extraction and management

We extracted the following variables from each included study to a specifically designed Excel® spreadsheet:

- Author, year of publication and journal (including language)
- Country
- Setting
- Inclusion criteria
- Exclusion criteria
- Study design and flow of patient pathway
- Population
- Sample size
- Details of diagnosis of vulval cancer (diagnostic biopsy or radical excision)
- Pathological parameters
 - Size of tumour
 - Histological type
 - Unifocal or multi-focal
 - Lympho-vascular invasion
 - Previous history of vulval surgery
- Details of any suspected groin node involvement prior to sentinel node assessment
- Additional tests performed to assess groin lymph node status prior to the sentinel node assessment
- Experience of the surgeons
- Index test
 - Method(s)
 - Details of tracer agent used, amount, dilution
 - Method of application
 - Timing of application in relation to sentinel node excision
 - Method used to detect sentinel node
 - Method used for histological assessment of sentinel node
 - Results
 - Detection rate of sentinel node (total intended versus total detected)
 - False negative sentinel node test (categorised by negative sentinel node)
 - Rate of adverse events associated with index test
- Reference standard
 - Unilateral or bilateral
 - Is positive sentinel node (but negative remaining groin nodes) regarded as positive reference standard?
 - Average lymph node yield (quality marker for reference standard)
 - Extent of surgical dissection of groin lymph nodes
 - Method used for histological assessment
 - Rate of adverse effects associated with reference standard

- QUADAS-2 items (see [Assessment of methodological quality](#) below)
- Data for two-by-two table

We piloted the data extraction spreadsheet including QUADAS-2 items using two included studies. We matched the data between the two authors (TAL and AP) and resolved differences by revisiting the original articles.

Assessment of methodological quality

AP and TAL performed the assessment of methodological quality. In the event of disagreement, other co-authors were involved. We assessed treatment pathways in detail for each included study. We also assessed the description of index and reference standard tests for each included study to determine if these were described in sufficient detail to enable the reader to reproduce the technique. We assessed study methodological quality using the QUADAS-2 tool ([Whiting 2011](#)) as described in [Appendix 3](#), and reported the results in detail in a tabular and graphical form. We also summarised results in the text.

Statistical analysis and data synthesis

To determine diagnostic test accuracy (DTA), we performed separate analyses with 'per groin' and 'per women' data. We created two-by-two tables in Review Manager software ([RevMan 2012](#)) and calculated sensitivity (diagnostic accuracy statistic for proportion of women with the disease who were correctly identified from the test) for each included study. Specificity was always 100% as, when the reference standard was negative, sentinel node was always negative. We presented diagnostic accuracy statistics for each study in a paired forest plot (where specificity was always 100%). As a result, sensitivity statistics were lined up on y-axis (sensitivity) crossing x-axis at 0 (1-specificity) on receiver operating characteristic (ROC) space. As there was no variability in specificity, we used the univariate model to pool these sensitivity data by removing the logit specificity and correlation parameters from the standard bivariate model ([Reitsma 2005](#)), thus simplifying the model to a univariate random-effects logistic regression model. The analysis was carried out using the `xmlogit` command and macro procedures in Stata IC version 12.0 ([Stata 2012](#)).

We estimated mean detection rates for each technique by combining the number of participants detected in each included study (numerator), divided by the total number of participants in which the technique was attempted (denominator), and multiplied by 100. We used these crude rates to illustrate the clinical consequences of the DTA results.

Investigations of heterogeneity

We anticipated that multiple factors could influence DTA statistics for sentinel node detection and analyses in vulval cancer. These factors may lead to heterogeneity in the analyses. In the univariate analysis of sensitivity we did not quantify heterogeneity using the I^2 statistic or make inference about heterogeneity using the variance parameter for logit sens. Instead of quantifying heterogeneity, we investigated heterogeneity where possible. We explored the effect of heterogeneity by investigating forest plots limited to relevant study level subgroup co-variables. We attempted to explore the potential effects of the following covariates:

- Index test used in sentinel node detection, e.g. blue dye and/or technetium
- Techniques used in sentinel node histological assessment, e.g. ultrastaging, immunohistochemical (IHC) staining
- Size of tumour (less than 4 cm)
- Experience of surgeon/s

We were unable to explore heterogeneity related to other variables (including histology, site and focality of tumour, previous vulval surgery and the use of imaging techniques prior to enrolment) due to insufficient relevant data.

Sensitivity analyses

During the review process we discovered that several studies had included some women with clinically suspicious or palpable groin nodes. As this was not predicted at the protocol stage, and raised concerns about applicability, we decided to include these studies if the number of women or groins affected was 10% or less of the total sample, or if investigators supplied sufficient information for us to exclude these data from the study results. We noted our concerns regarding the applicability of the samples assessed in these studies and performed sensitivity analysis by excluding these studies to assess their effect on the review results. We also performed sensitivity analyses related to other methodological quality items (QUADAS-2) including the type of study (retrospective versus prospective).

Assessment of reporting bias

Where possible, we explored patient withdrawals and drop-outs from individual studies. We included data from studies presented at conferences but not published in full, and attempted to obtain further details of these studies, to minimise publication bias. Where studies had been published more than once, at various stages of enrolment, we checked that the data in the earlier and later reports corresponded and, if not, we attempted to obtain clarification from the investigators.

RESULTS

Results of the search

The combined de-duplicated 2013 MEDLINE and EMBASE searches yielded 2020 records. Two review authors (TAL, AP) independently screened these titles and abstracts, selecting 103 records for classification. After obtaining the full texts, we excluded 51 records (pertaining to 47 studies/reports) mainly for the following reasons:

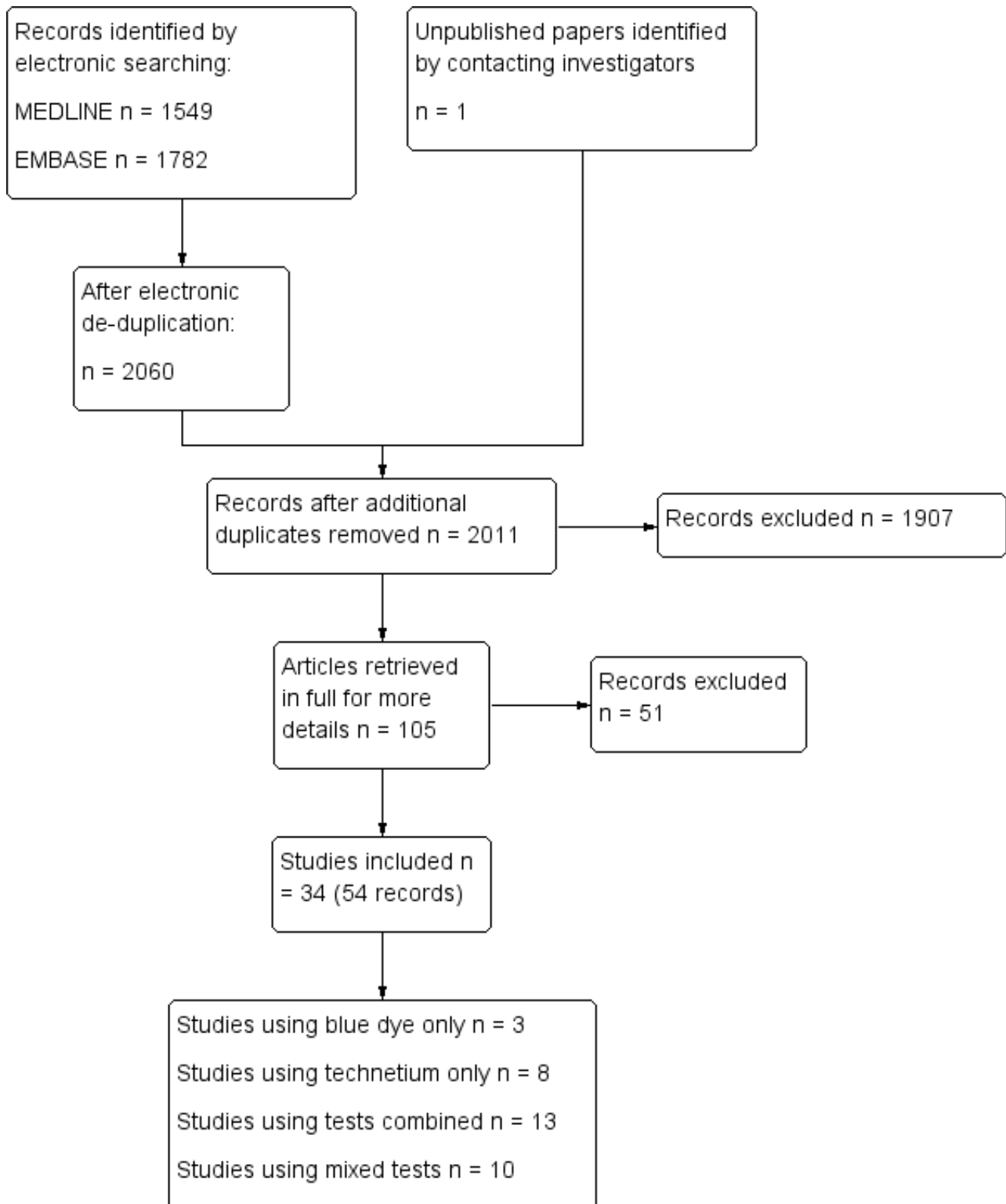
- they were reviews, editorials, case reports or letters to the editor (eight);
- we were unable to construct two-by-two tables from the available data (12);
- the reference standard had not been consistently applied (12);
- they were not studies assessing sentinel node test accuracy (11);
- more than 10% of participants had clinically suspicious nodes and we were unable to separate these data from the other participants (three); or
- the sample size was less than 10 (one).

For further details see [Characteristics of excluded studies](#).

We included 34 studies comprising 54 citations (see [Figure 3](#)). For the purposes of this review, we emailed the investigators of 20 studies for further information and/or data. We obtained unpublished information and/or data for six of these studies ([Levenback 2012](#); [Morotti 2011](#); [Nyberg 2007](#); [Rob 2007](#); [Sawicki](#)

[2010](#); [Trifiro 2010](#)), including an unpublished manuscript ([Morotti 2011](#)). The latter study would otherwise have been excluded had we not received these unpublished data, as we were unable to construct two-by-two tables from the published conference abstract alone.

Figure 3. Study flow diagram.



The included studies evaluated the following index tests:

- Blue dye only (three studies; [Ansink 1999](#); [Echt 1999](#); [Levenback 2001](#)).
- Technetium only (eight studies; [Boran 2003](#); [DeCesare 1997](#); [Goni 2011](#); [Klar 2011](#); [Merisio 2005](#); [Sideri 2000](#); [Trifiro 2010](#); [Zekan 2012](#)).
- Technetium in combination with blue dye (combined tests; 13 studies; [Basta 2005](#); [Camara 2009](#); [Crosbie 2010](#); [de Hullu 2000](#); [Johann 2008](#); [Klat 2009](#); [Louis-Sylvestre 2006](#); [Martinez-Palones 2006](#); [Moore 2003a](#); [Morotti 2011](#); [Radziszewski 2010](#); [Vidal-Sicart 2007](#); [Zambo 2002](#)).
- A combination of the above tests (mixed tests; 10 studies; [Akrivos 2011](#); [HAMPL 2008](#); [Hauspy 2007](#); [Levenback 2012](#); [Li 2009](#); [Lindell 2010](#); [Nyberg 2007](#); [Rob 2007](#); [Sawicki 2010](#); [Sliutz 2002](#)).

Methodological quality of included studies

Of the 34 studies, we considered four included studies to be at a high risk of bias for flow and timing, and three studies raised high concerns regarding applicability ([Figure 4](#)). However, in general, we considered the quality of included studies to be moderate, with the risk of bias mostly low or unclear ([Figure 5](#)).

Figure 4. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study

	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Akrivos 2011	+	+	?	?	+	+	+
Ansink 1999	?	?	?	?	?	+	+
Basta 2005	?	?	?	?	?	+	+
Boran 2003	?	-	?	?	?	+	+
Camara 2009	?	?	+	?	?	+	+
Crosbie 2010	?	?	?	?	+	+	+
DeCesare 1997	+	?	?	?	?	+	+
de Hullu 2000	+	?	+	?	?	+	+
Echt 1999	?	?	?	?	?	+	+
Goni 2011	?	?	?	?	+	+	+
HAMPL 2008	?	?	?	?	+	+	+
Hauspy 2007	+	?	?	?	?	+	+
Johann 2008	+	?	?	?	?	+	+
Klar 2011	+	+	?	+	+	+	+
Klat 2009	?	?	?	?	+	+	+
Levenback 2001	?	?	?	-	-	+	+
Levenback 2012	?	?	?	?	+	+	+
Li 2009	?	?	?	?	?	?	+
Lindell 2010	+	?	?	?	+	+	+
Louis-Sylvestre 2006	+	?	?	?	-	+	+
Martinez-Palones 2006	+	?	?	?	+	+	+
Merisio 2005	?	?	?	?	?	+	+
Moore 2003a	?	?	?	?	+	+	+
Morotti 2011	+	+	?	+	+	+	+
Nyberg 2007	?	?	?	?	-	+	+
Radziszewski 2010	+	-	?	?	+	+	+

Figure 4. (Continued)

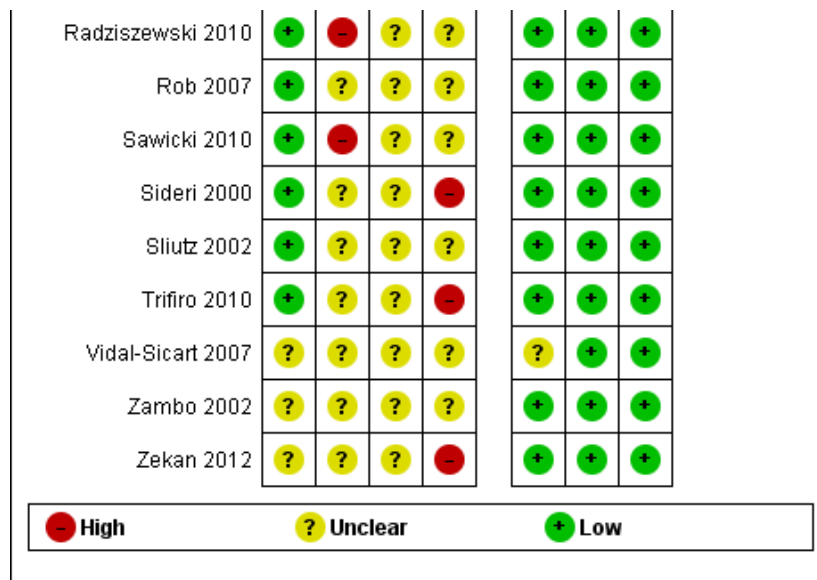
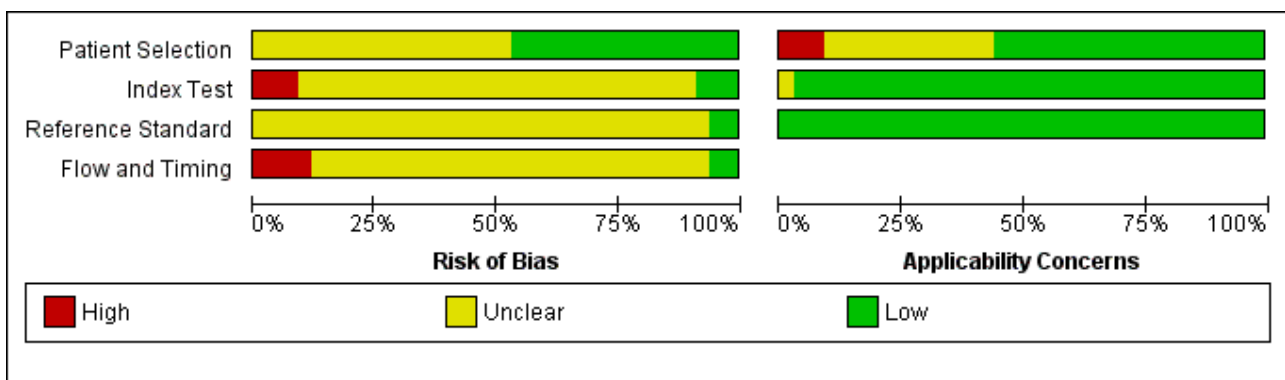


Figure 5. Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies



Types of studies

We included 30 prospective and four retrospective studies. We did not include case-control studies. The sampling method was consecutive in 20 studies, and not clearly described in the other 14 studies. All were conducted in university hospitals and tertiary care settings.

Patient selection

We considered the participants of most studies to be representative of patients in clinical practice. Most participants:

- had squamous cell cancer (SCC) of the vulva, except for [Trifiro 2010](#) (melanomas only) and [Levenback 2001](#) (67% SCC, 33% other histology);
- were between the ages of 29 and 95 years, with reported mean and median ages ranging from 58 to 75 years (eight studies did not report age);
- did not have clinically suspicious nodes.

Four studies included some women with suspicious nodes in their study samples ([DeCesare 1997](#); [Levenback 2001](#); [Louis-Sylvestre 2006](#), [Vidal-Sicart 2007](#)). We excluded data for these women from our data extraction for [Vidal-Sicart 2007](#) and [Louis-Sylvestre 2006](#). For the other two studies, the women with clinically suspicious nodes comprised 10% or less of the participants. We considered these studies to be at an unclear risk of selection bias.

Vulval lesions were midline in 582 women, lateralised in 308 women and the location was not described for 724 women. Tumour size was either not reported (22 studies) or inconsistently reported (12 studies): some studies reported the number of tumours greater than and less than 2 cm, some reported a 4 cm cut-off and some reported continuous data (mean size). Depth of tumour, lymphovascular space invasion (LVSI) and grade were rarely reported.

Few described whether any withdrawals or exclusions had occurred during or after the selection process. It was mainly this lack of clarity

that increased the proportion of studies in which the risk of bias relating to patient sample selection was 'unclear'.

Index test methods

The index test methods of the included studies were highly applicable to this review, with a low risk of potential bias (Figure 5). They comprised the following techniques: blue dye only (three studies), technetium only (eight studies), a combination of blue dye and technetium (13 studies) or mixed tests (any or all of the previous three techniques used within each study; 10 studies). Most studies reported that index test contents were injected peritumourally, in two to four sites around the tumour, or at 3, 6, 9 and 12 o'clock. Blue dye was injected pre-operatively, after general anaesthesia for all studies. When technetium was used, the timing was subject to some variation: 13 studies injected technetium on the day before the operation (Akrivos 2011; Crosbie 2010; de Hullu 2000; Goni 2011; Johann 2008; Louis-Sylvestre 2006; Martinez-Palones 2006; Morotti 2011; Radziszewski 2010; Sideri 2000; Trifiro 2010; Vidal-Sicart 2007; Zambo 2002), two studies injected it between 14 and 18 hours pre-operatively (Basta 2005; Merisio 2005), eight studies injected it two to four hours pre-operatively (Hampl 2008; Hauspy 2007; Lindell 2010; Klat 2009; Moore 2003a; Rob 2007; Sliutz 2002; Zekan 2012), and three studies injected it intra-operatively or within two hours of surgery (DeCesare 1997; Klat 2009; Sawicki 2010). In three studies, the timing was unclear (Camara 2009; Levenback 2012; Li 2009), and in one study injections were given either on the day of surgery or on the day before surgery (Nyberg 2007).

Where studies employed mixed tests, administering the various index tests alone (e.g. blue dye or technetium tests alone) and in combination (e.g. technetium and blue dye), separate data were frequently not reported. Where possible, we emailed investigators to request separate data and obtained these data for Rob 2007 and Levenback 2012. In Hauspy 2007, for a subset of women, the choice of one index method (blue dye) was dependent on the success of the other method (Tc-99m) and therefore separate data would not have been meaningful.

Most studies reported using H&E stains to diagnose groin metastases. Fourteen studies additionally employed ultrastaging and immunohistochemical (IHC) stains to improve detection (Akrivos 2011; Crosbie 2010; de Hullu 2000; Klar 2011; Klat 2009; Levenback 2001; Levenback 2012; Lindell 2010; Hampl 2008; Hauspy 2007; Merisio 2005; Morotti 2011; Rob 2007; Vidal-Sicart 2007), and eight studies reported using IHC stains, but not ultrastaging (Basta 2005; Boran 2003; Goni 2011; Louis-Sylvestre 2006; Martinez-Palones 2006; Radziszewski 2010; Sliutz 2002; Trifiro 2010). One study also employed reverse transcriptase polymerase chain reaction (RT-PCR) analysis (CA9 RT-PCR) to enhance detection (Radziszewski 2010), presenting results with and without the RT-PCR analysis. Due to the experimental nature of this test, we did not use the RT-PCR results in our analyses. Histological methods were not described in three studies (Johann 2008; Nyberg 2007; Sawicki 2010).

For blue dye, most studies considered blue lymph nodes and draining lymphatics that turned blue after the index test injection to indicate the presence of a sentinel node. For technetium, sentinel nodes were usually detected intra-operatively using a hand-held gamma probe. Nodes were reported to be 'hot' if the measured radioactivity was five or 10 times greater than the background

activity (e.g. Hauspy 2007; Levenback 2012; Sawicki 2010) or greater than 5% (e.g. Akrivos 2011) or 10% (e.g. Rob 2007) of the activity of the injection site. Several studies reported continuing dissection if more 'hot' nodes were identified (e.g. defined as activity of >5% or >10% of the activity of the injection site or the 'hottest' sentinel node) (e.g. Boran 2003; Klar 2011; Martinez-Palones 2006; Zekan 2012). Sixteen studies reported the 'mean sentinel node yield' per groin (ranged from 1 to 2.7 sentinel nodes); three studies reported the median sentinel node yield per groin (ranged from 1 to 5 sentinel nodes); and 15 studies did not report either the mean or median sentinel node yield per groin.

For 'surgeon/s experience', either this variable was not reported in the studies, or the element of a 'learning curve' was described; i.e. the surgeon/s gained the necessary experience (10 sentinel node procedures) over the course of the study. Only two studies reported that participating surgeons had performed a minimum of 10 procedures (Klar 2011; Morotti 2011).

Reference standards

All studies reported using histological examination of inguinofemoral lymphadenectomy (IFL) as the reference standard. Few studies described the extent of surgical dissection, therefore it was not possible to determine whether heterogeneity existed in this regard. Most studies reported performing bilateral IFL for midline lesions and unilateral IFL for lateralised lesions. Eleven studies defined midline lesions, either as a lesion within 1 cm of the midline (Ansink 1999; de Hullu 2000; Hampl 2008; Hauspy 2007; Klar 2011; Lindell 2010; Louis-Sylvestre 2006), or within 2 cm of the midline (Levenback 2012; Merisio 2005; Sideri 2000; Zekan 2012), with the other 23 studies not reporting their definitions of 'midline'. Only one study reported blinding the assessors of the reference standard to the results of the index test histology (de Hullu 2000). Thirteen studies (38%) reported that reference standard and index test specimens were sent separately to the laboratory for examination; however, it was unclear to us whether this was supposed to reflect some degree of assessor blinding (see [Characteristics of included studies](#)). Therefore, we considered most studies to be at an 'unclear risk' of bias for this item.

Flow and timing

Most studies reported that enrolled women underwent the index test within 24 hours of surgery, sentinel node removal and IFL procedures were performed during the same operation, and specimens were sent to the laboratory immediately thereafter for examination. However, the risk of bias for patient flow in the majority of the studies was 'unclear' overall. This occurred mainly due to the lack of clarity in most studies regarding the signalling question related to 'additional imaging tests'. In only four studies was it described that women had undergone additional imaging tests (ultrasound or CT) to exclude groin lymph node metastases (Hampl 2008; Morotti 2011; Radziszewski 2010; Rob 2007). When this question was excluded from the 'Risk of bias' assessment, the overall risk of bias for patient flow was low.

Other methodological issues

There were some unit of analysis issues, with nine studies reporting test accuracy data 'per women' only (Basta 2005; Camara 2009; Johann 2008; Nyberg 2007; Rob 2007; Sawicki 2010; Sliutz 2002; Trifiro 2010; Zekan 2012). We emailed contact authors of these

studies and obtained 'per groin' data for [Nyberg 2007](#); [Rob 2007](#); [Sawicki 2010](#) and [Trifiro 2010](#).

[De Cicco 2000](#) and [Sideri 2000](#) were two reports of the same consecutive case series of women in Italy: women in the former report were recruited between May 1996 and September 1998 ([De Cicco 2000](#)); women in the later report were recruited from May 1998 to July 1999 ([Sideri 2000](#)). Taking into consideration that the latter paper was an extension of the [De Cicco 2000](#) data set, there remained inconsistencies between the reports with regard to the number of groins, number and types of procedures, and number and types of lesions (lateralised or midline). Contact authors were emailed for clarification, however they were unable to locate these data due to the time lapse since the study. We used the data in the later publication for this review.

[Zekan 2012](#) was also reported twice, as an article and a conference abstract. There are four fewer women reported in the published article of 2012 than the (earlier) conference abstract. We were unable to obtain clarification of this discrepancy and noted our concerns regarding the possibility of withdrawals in the 'Risk of bias' assessment for flow and timing.

In addition to using ultrastaging and IHC stains, [Radziszewski 2010](#) also used a RT-PCR test. As this was the only study to use this histology method, we did not include the accuracy results of this test in our analysis. By using the PCR test, double the number of true positives were detected; these were mainly micrometastases, for which the clinical significance is unknown.

Findings

1. Detection rates (the ability of a test to identify a sentinel node)

Sentinel node detection rates across included studies varied according to the index test method used and the unit of analysis reported (i.e. 'per groin' or 'per woman' data) ([Table 2](#)).

- For blue dye only, detection rates ranged from 55% to 75% for 'per groin' data (mean 65%; five studies, 228 groins) and 69% to 88% for 'per woman' data (mean 82%; four studies, 131 women).
- For technetium only, detection rates ranged from 68% to 100% for 'per groin' data (mean 84%; six studies, 189 groins) and 75% to 100% for 'per woman' data (mean 96%; eight studies, 159 women).
- For blue dye/technetium combined, detection rates ranged from 75% to 100% for 'per groin' data (mean 93%; nine studies, 607 groins) and 88% to 100% for 'per woman' data (mean 98%; 11 studies, 371 women).
- For mixed tests, detection rates ranged from 70% to 99% for 'per groin' data (mean 81%; seven studies, 1349 groins) and 92% to 100% for 'per woman' data (mean 95%; nine studies, 827 women).

Less than half of the studies reported the status of undetected nodes and it was not meaningful to analyse these limited data.

2. Test accuracy (the ability of a test to identify cancer)

We included test accuracy data from 1614 women with FIGO stage 1B or higher vulval cancer in our sensitivity meta-analyses.

2.1. Test accuracy according to index test method

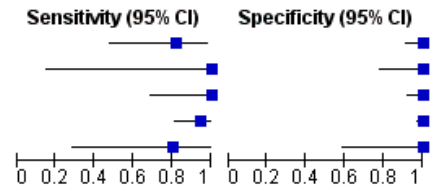
Per groin data

For each index test, the pooled estimates for sensitivity 'per groin' were as follows ([Figure 6](#)):

Figure 6. Forest plot of tests: 3 Per groin: Blue dye only, 5 Per groin: Tc-99m only, 7 Per groin: Combined test, 9 Per groin: Mixed tests.

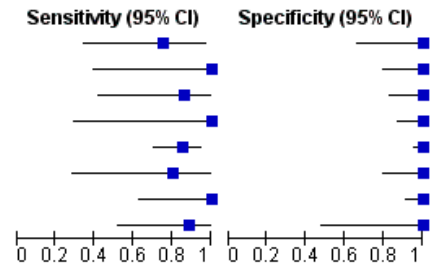
Per groin: Blue dye only

Study	TP	FP	FN	TN	Test	Sensitivity (95% CI)	Specificity (95% CI)
Ansink 1999	9	0	2	41	B	0.82 [0.48, 0.98]	1.00 [0.91, 1.00]
Echt 1999	2	0	0	15	B	1.00 [0.16, 1.00]	1.00 [0.78, 1.00]
Levenback 2001	10	0	0	47	B	1.00 [0.69, 1.00]	1.00 [0.92, 1.00]
Levenback 2012	33	0	2	117	M	0.94 [0.81, 0.99]	1.00 [0.97, 1.00]
Rob 2007	4	0	1	7	M	0.80 [0.28, 0.99]	1.00 [0.59, 1.00]



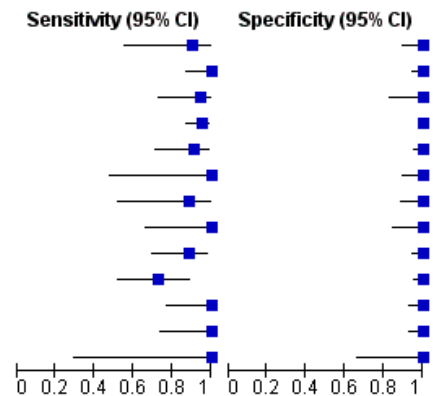
Per groin: Tc-99m only

Study	TP	FP	FN	TN	Test	Sensitivity (95% CI)	Specificity (95% CI)
Boran 2003	6	0	2	9	T	0.75 [0.35, 0.97]	1.00 [0.66, 1.00]
DeCesare 1997	4	0	0	16	T	1.00 [0.40, 1.00]	1.00 [0.79, 1.00]
Goni 2011	6	0	1	20	T	0.86 [0.42, 1.00]	1.00 [0.83, 1.00]
Klar 2011	3	0	0	26	T	1.00 [0.29, 1.00]	1.00 [0.87, 1.00]
Levenback 2012	34	0	6	78	M	0.85 [0.70, 0.94]	1.00 [0.95, 1.00]
Merisio 2005	4	0	1	16	T	0.80 [0.28, 0.99]	1.00 [0.79, 1.00]
Sideri 2000	8	0	0	42	T	1.00 [0.63, 1.00]	1.00 [0.92, 1.00]
Trifiro 2010	8	0	1	5	T	0.89 [0.52, 1.00]	1.00 [0.48, 1.00]



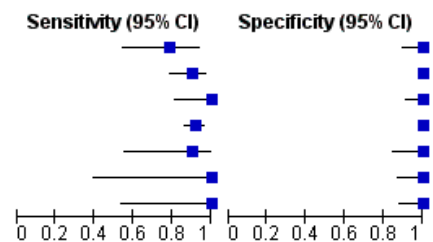
Per groin: Combined test

Study	TP	FP	FN	TN	Test	Sensitivity (95% CI)	Specificity (95% CI)
Crosbie 2010	9	0	1	35	BT	0.90 [0.55, 1.00]	1.00 [0.90, 1.00]
de Hullu 2000	27	0	0	68	BT	1.00 [0.87, 1.00]	1.00 [0.95, 1.00]
Klat 2009	17	0	1	20	BT	0.94 [0.73, 1.00]	1.00 [0.83, 1.00]
Levenback 2012	73	0	4	246	M	0.95 [0.87, 0.99]	1.00 [0.99, 1.00]
Lindell 2010	20	0	2	72	BT	0.91 [0.71, 0.99]	1.00 [0.95, 1.00]
Louis-Sylvestre 2006	5	0	0	34	BT	1.00 [0.48, 1.00]	1.00 [0.90, 1.00]
Martinez-Palones 2006	8	0	1	30	BT	0.89 [0.52, 1.00]	1.00 [0.88, 1.00]
Moore 2003a	9	0	0	22	BT	1.00 [0.66, 1.00]	1.00 [0.85, 1.00]
Morotti 2011	23	0	3	66	BT	0.88 [0.70, 0.98]	1.00 [0.95, 1.00]
Radziszewski 2010	19	0	7	83	BT	0.73 [0.52, 0.88]	1.00 [0.96, 1.00]
Rob 2007	14	0	0	48	M	1.00 [0.77, 1.00]	1.00 [0.93, 1.00]
Vidal-Sicart 2007	12	0	0	48	BT	1.00 [0.74, 1.00]	1.00 [0.93, 1.00]
Zambo 2002	3	0	0	9	BT	1.00 [0.29, 1.00]	1.00 [0.66, 1.00]



Per groin: Mixed tests

Study	TP	FP	FN	TN	Test	Sensitivity (95% CI)	Specificity (95% CI)
Akrivos 2011	15	0	4	33	M	0.79 [0.54, 0.94]	1.00 [0.89, 1.00]
Hampel 2008	46	0	5	179	M	0.90 [0.79, 0.97]	1.00 [0.98, 1.00]
Hauspy 2007	18	0	0	40	M	1.00 [0.81, 1.00]	1.00 [0.91, 1.00]
Levenback 2012	140	0	12	441	M	0.92 [0.87, 0.96]	1.00 [0.99, 1.00]
Li 2009	9	0	1	22	M	0.90 [0.55, 1.00]	1.00 [0.85, 1.00]
Nyberg 2007	4	0	0	27	M	1.00 [0.40, 1.00]	1.00 [0.87, 1.00]
Sawicki 2010	6	0	0	28	M	1.00 [0.54, 1.00]	1.00 [0.88, 1.00]



- Blue dye only: 0.92, 95% confidence interval (CI) 0.82 to 0.97 (five studies; 290 groins).
- Technetium only: 0.91, 95% CI 0.87 to 0.94 (eight studies; 296 groins).
- Combined tests: 0.94, 95% CI 0.88 to 0.97 (13 studies; 1039 groins).
- Mixed tests: 0.87, 95% CI 0.77 to 0.93 (seven studies; 1030 groins).

Negative predictive values (NPVs) for the above tests were 98% for blue dye and the combined tests, 97% for mixed tests and 95% for technetium only.

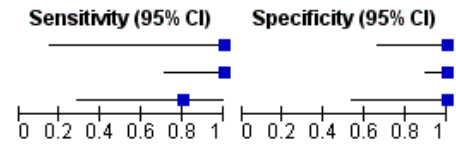
Per woman data

For each index test, the pooled estimates for sensitivity 'per woman' were as follows (Figure 7):

Figure 7. Forest plot of tests: 4 Per woman: Blue dye only, 6 Per woman: Tc-99m only, 8 Per woman: Combined test, 10 Per woman: Mixed tests.

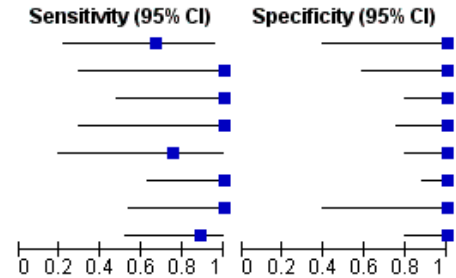
Per woman: Blue dye only

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Echt 1999	2	0	0	9	1.00 [0.16, 1.00]	1.00 [0.66, 1.00]
Levenback 2001	11	0	0	35	1.00 [0.72, 1.00]	1.00 [0.90, 1.00]
Rob 2007	4	0	1	6	0.80 [0.28, 0.99]	1.00 [0.54, 1.00]



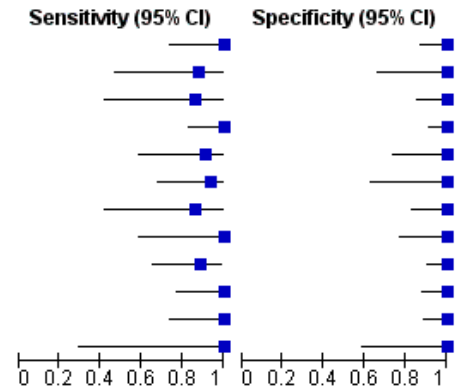
Per woman: Tc-99m only

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Boran 2003	4	0	2	4	0.67 [0.22, 0.96]	1.00 [0.40, 1.00]
DeCesare 1997	3	0	0	7	1.00 [0.29, 1.00]	1.00 [0.59, 1.00]
Goni 2011	5	0	0	16	1.00 [0.48, 1.00]	1.00 [0.79, 1.00]
Klar 2011	3	0	0	13	1.00 [0.29, 1.00]	1.00 [0.75, 1.00]
Merisio 2005	3	0	1	16	0.75 [0.19, 0.99]	1.00 [0.79, 1.00]
Sideri 2000	8	0	0	29	1.00 [0.63, 1.00]	1.00 [0.88, 1.00]
Trifiro 2010	6	0	0	4	1.00 [0.54, 1.00]	1.00 [0.40, 1.00]
Zekan 2012	8	0	1	16	0.89 [0.52, 1.00]	1.00 [0.79, 1.00]



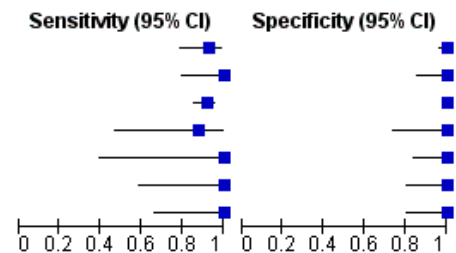
Per woman: Combined test

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Basta 2005	12	0	0	26	1.00 [0.74, 1.00]	1.00 [0.87, 1.00]
Camara 2009	7	0	1	9	0.88 [0.47, 1.00]	1.00 [0.66, 1.00]
Crosbie 2010	6	0	1	24	0.86 [0.42, 1.00]	1.00 [0.86, 1.00]
de Hullu 2000	20	0	0	39	1.00 [0.83, 1.00]	1.00 [0.91, 1.00]
Johann 2008	10	0	1	12	0.91 [0.59, 1.00]	1.00 [0.74, 1.00]
Klat 2009	14	0	1	8	0.93 [0.68, 1.00]	1.00 [0.63, 1.00]
Martinez-Palones 2006	6	0	1	20	0.86 [0.42, 1.00]	1.00 [0.83, 1.00]
Moore 2003a	7	0	0	14	1.00 [0.59, 1.00]	1.00 [0.77, 1.00]
Morotti 2011	16	0	2	37	0.89 [0.65, 0.99]	1.00 [0.91, 1.00]
Rob 2007	14	0	0	29	1.00 [0.77, 1.00]	1.00 [0.88, 1.00]
Vidal-Sicart 2007	12	0	0	31	1.00 [0.74, 1.00]	1.00 [0.89, 1.00]
Zambo 2002	3	0	0	7	1.00 [0.29, 1.00]	1.00 [0.59, 1.00]



Per woman: Mixed tests

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
HAMPL 2008	36	0	3	88	0.92 [0.79, 0.98]	1.00 [0.96, 1.00]
Hauspy 2007	16	0	0	23	1.00 [0.79, 1.00]	1.00 [0.85, 1.00]
Levenback 2012	121	0	11	286	0.92 [0.86, 0.96]	1.00 [0.99, 1.00]
Li 2009	7	0	1	12	0.88 [0.47, 1.00]	1.00 [0.74, 1.00]
Nyberg 2007	4	0	0	21	1.00 [0.40, 1.00]	1.00 [0.84, 1.00]
Sawicki 2010	7	0	0	17	1.00 [0.59, 1.00]	1.00 [0.80, 1.00]
Sliutz 2002	9	0	0	17	1.00 [0.66, 1.00]	1.00 [0.80, 1.00]



- Blue dye only: 0.94, 95% CI 0.69 to 0.99 (three studies; 68 women).
- Technetium only: 0.93, 95% CI 0.89 to 0.96 (eight studies; 149 women).
- Combined tests: 0.95, 95% CI 0.89 to 0.97 (12 studies; 390 women).
- Mixed tests: 0.91, 95% CI 0.71 to 0.98 (seven studies; 679 women).

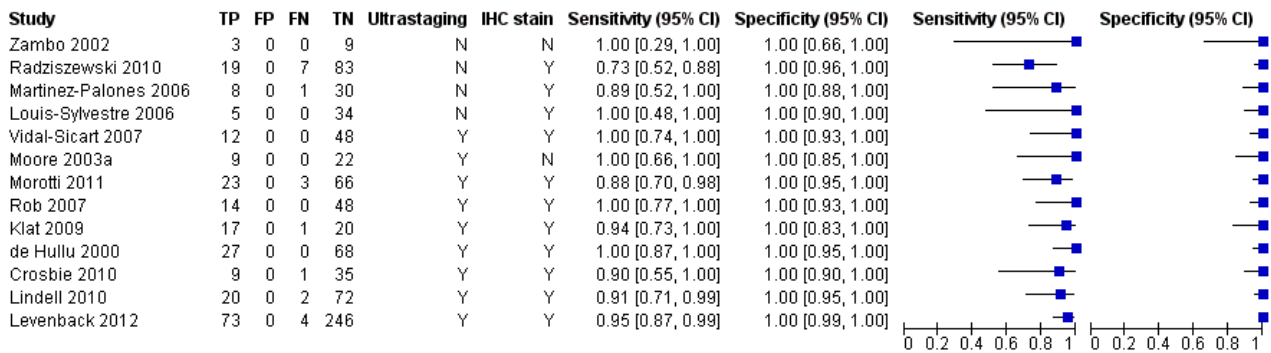
NPVs of the above tests ranged from 96% to 98%. The rate of groin node metastases in women across all included studies and index tests ((true positives + false negatives)/total number of women evaluated) was 32% (29 studies; 411/1286 women).

2.2 Test accuracy for combined tests (blue dye and technetium) according to histological methods

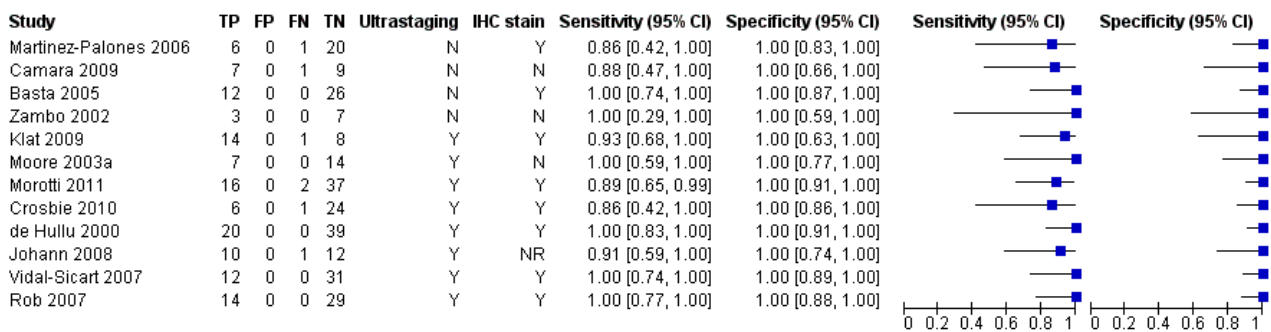
Pooled estimates of sensitivity for the combined tests according to histological methods were as follows (Figure 8):

Figure 8. Forest plot of tests: 7 Per groin: Combined test, 8 Per woman: Combined test. Covariate: ultrastaging and/or IHC

Per groin: Combined test



Per woman: Combined test



Per groin data

- Ultrastaging only: 0.95, 95% CI 0.91 to 0.97 (nine studies; 840 groins). Four studies (Louis-Sylvestre 2006; Martinez-Palones 2006; Radziszewski 2010; Zambo 2002), which did not report or use ultrastaging, had a pooled sensitivity estimate of 0.81, 95% CI 0.67 to 0.90 (four studies; 199 groins).
- Ultrastaging and/or immunohistochemistry (IHC): 0.94, 95% CI 0.88 to 0.97 (12 studies; 828 groins). Only one study did not use ultrastaging or IHC (Zambo 2002).

Per woman data

- Ultrastaging only: 0.95, 95% CI 0.89 to 0.98 (eight studies; 300 women). Four studies (Basta 2005; Camara 2009; Martinez-Palones 2006; Zambo 2002), which did not report or use ultrastaging, had a pooled sensitivity estimate of 0.93, 95% CI 0.77 to 0.98 (four studies; 92 women).
- Ultrastaging and/or IHC: 0.95, 95% CI 0.90 to 0.98 (10 studies; 363 women). Only Camara 2009 and Zambo 2002 did not report or

use ultrastaging or IHC; pooled sensitivity was 0.91, 95% CI 0.56 to 0.99 (two studies; 27 women).

2.3. Test accuracy for combined tests according to surgeons' experience

Only one study using the combined tests reported that the surgeons had performed more than 10 procedures prior to the study (Morotti 2011), therefore, meta-analysis according to this covariate was not possible.

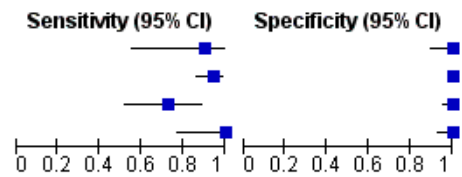
2.4. Test accuracy for combined tests according to tumour size

Four studies evaluated test accuracy for tumours of less than 4 cm for 'per groin' data (Crosbie 2010; Levenback 2012; Radziszewski 2010; Rob 2007) (Figure 9), with two of these studies also reporting 'per woman' data (Crosbie 2010; Rob 2007). The pooled estimate for sensitivity for 'per groin' data was 0.91, 95% CI 0.75 to 0.97 (four studies; 485 groins) and for 'per woman' data was 0.99, 95% CI 0.73 to 0.99 (two studies; 74 women).

Figure 9. Forest plot of tests: 11 Per groin: Tumour size < 4 cm, 12 Per woman: Tumour size < 4 cm.

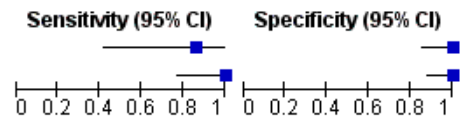
Per groin: Tumour size < 4 cm

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Crosbie 2010	9	0	1	35	0.90 [0.55, 1.00]	1.00 [0.90, 1.00]
Levenback 2012	67	0	4	198	0.94 [0.86, 0.98]	1.00 [0.98, 1.00]
Radziszewski 2010	19	0	7	83	0.73 [0.52, 0.88]	1.00 [0.96, 1.00]
Rob 2007	14	0	0	48	1.00 [0.77, 1.00]	1.00 [0.93, 1.00]



Per woman: Tumour size < 4 cm

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Crosbie 2010	6	0	1	24	0.86 [0.42, 1.00]	1.00 [0.86, 1.00]
Rob 2007	14	0	0	29	1.00 [0.77, 1.00]	1.00 [0.88, 1.00]



2.5. Sensitivity analyses for combined tests

Type of study

Per groin data

Only one study in the combined test meta-analysis was retrospective (Lindell 2010). When we excluded this study from the meta-analysis, the pooled estimate for sensitivity was 0.95, 95% CI 0.87 to 0.98 (12 studies; 945 groins).

Per woman data

Only one study in the combined test meta-analysis was retrospective (Johann 2008). When we excluded this study from the meta-analysis, the pooled estimate for sensitivity was 0.95, 95% CI 0.90 to 0.98 (11 studies; 367 women).

Other potential sources of heterogeneity

Three studies in the combined test meta-analysis reported the use of pre-operative imaging procedures (CT or ultrasound) using 'per groin' data (Morotti 2011; Radziszewski 2010; Rob 2007), and two of these additionally included 'per woman' data (Morotti 2011; Rob 2007).

Per groin data

The pooled sensitivity estimate of these three studies 'per groin' was 0.88, 95% CI 0.67 to 0.96 (three studies; 206 groins), compared with 0.95, 95% CI 0.91 to 0.98 (10 studies; 833 groins) for the studies in which it was not clear whether pre-operative imaging had been used.

Per woman data

The pooled estimate for sensitivity for 'per woman' data was 0.94, 95% CI 0.78 to 0.98 (two studies; 98 women), compared with 0.95, 95% CI 0.89 to 0.98 (10 studies; 292 women) for the studies which did not report whether pre-operative imaging had been used.

DISCUSSION

Summary of main results

The rate of groin node metastases in women across all included studies was approximately 32%. All index tests were associated

with pooled sensitivity estimates of greater than 90% for 'per woman' data and 'per groin' data (with exception of mixed tests where the pooled sensitivity estimate was 87% for 'per groin' data) (Summary of findings 1). The negative predictive value (NPV) for all index tests 'per woman' was greater than 95%. The combined tests were associated with the best pooled sensitivity estimate, and a narrow confidence interval, however the estimate was not much higher than that of technetium alone. Pooled sensitivity estimates for blue dye alone and mixed tests were associated with wide confidence intervals, which suggests that blue dye may not be sufficiently accurate when used on its own.

Crude mean detection rates across included studies were calculated to be 98%, 96%, 82% and 95% for the combined test, technetium, blue dye and mixed tests, respectively (per woman data). We used these data with the sensitivity data to estimate the clinical consequences of the test results (Summary of findings 1). Based on the 'per woman' pooled estimates, and assuming that 30 out of 100 women with FIGO grade IB or higher vulvar cancer without suspicious nodes will have groin metastases (30%), one and two women with groin metastases may be 'missed' with the combined tests and technetium only test alone, with a confidence interval (CI) of one to three women for both tests. With blue dye and mixed tests, however, the upper limit of the confidence interval is eight and nine women, respectively. This suggests that utilising blue dye alone or mixed tests may 'miss' as many as nine women out of 30 with groin metastases.

We found little evidence showing the influence of other factors in further reducing the number of missed women. However, pooled sensitivity estimates for the combined test were probably enhanced by the use of ultrastaging (lower CI limits were 0.91 versus 0.88, with and without ultrastaging, respectively (per groin data)). Most studies used these additional techniques for sentinel nodes that were negative on routine H&E staining.

Four studies using the combined test (with ultrastaging and/or IHC staining) evaluated sensitivity data for vulvar lesions less than 4 cm in diameter. The pooled estimate for this subgroup of women was slightly lower than the overall sensitivity estimate (0.91, 95% CI 0.75 to 0.97), probably due to insufficient data.

Strengths and weaknesses of the review

Strengths

To our knowledge, this is the most meticulous review of sentinel node test accuracy in vulval cancer to date. Previous reviews have included studies reporting women with suspicious lymph nodes in their samples (e.g. [Molpus 2001](#); [Tavares 2001](#)), studies where the reference standard was not consistent for all women (e.g. [Molpus 2001](#)), and data from different reports of the same study (e.g. [Sideri 2000](#) and [De Cicco 2000](#), and the [Levenback 2001](#) series). By applying clearly defined, pre-specified inclusion criteria we classified potentially eligible studies in a consistent way, thereby attempting to minimise heterogeneity across studies. We also assessed the methodological quality of each study as well as making judgements on its risk of bias. We excluded studies where the sample size was fewer than 10 women, and those in which women with clinically suspicious nodes accounted for 10% or more of the sample, if we were unable to separate these data from the study results. Although several studies had relatively small samples, the review comprises a large number of studies ($n = 34$) and participants ($n = 1614$), increasing the power of the analyses (with meta-analyses ranging from two to 13 included studies). Not all included studies reported their data in the same way or used the same unit of analysis therefore, where possible, we contacted authors for clarification and/or additional data. We pooled and analysed data separately for each index test, and separately for 'per groin' and 'per woman' data. Variations in methodological quality did not appear to have any impact on the overall findings. We therefore consider the resulting evidence to be of a moderate quality.

Analysing both 'per groin' and 'per woman' data is a strength of this review: 'per groin' data may be more precise for assessing test accuracy, however, 'per woman' data are useful for clinical decision-making. For example, with midline vulval lesions, as long as a metastatic node in one groin is identified, additional treatment (usually bilateral IFL) will be clinically indicated. Therefore, even if the metastatic node in the opposite groin is not identified ('false negative' according to 'per groin' data), the *woman* will have been identified as needing additional treatment.

Weaknesses

There are inherent weaknesses in DTA studies where the reference standard incorporates the index test result, giving an associated specificity of 100%. Increasing test sensitivity can cause a corresponding drop in specificity, but this would not be detected in these studies due to the absence of false positive results. The clinical consequence of a false positive index test would be a greater extent of surgery (complete inguofemoral lymphadenectomy (IFL)) for a woman without groin node involvement. However, as complete IFL is also the standard management for a positive test result, the clinical consequences cannot be estimated.

Only one study in this review reported assessor blinding and in most studies the assessment procedure was not clearly described. In unblinded studies, knowledge of the reference standard results might have affected the interpretation of the index test results. This overall lack of blinding would most likely have impacted the results in the direction of overestimating sensitivity, as unblinded assessors may have been tempted to alter their assessment of the index test findings in light of the reference standard results,

for example, if an original index test result was equivocal or inconclusive.

It is possible that test accuracy and detection rates are affected in situations where the tumour has already been excised (where agents are injected around the scar), in multi-focal tumours and in women with a previous history of vulval surgery. We were unable to evaluate test accuracy data for these variables, or midline versus lateralised lesions, as these baseline data were not consistently reported. Similarly, we were unable to evaluate test accuracy data according to the depth of invasion of the primary lesion.

There were insufficient data on surgeons' experience, therefore we were unable to evaluate the impact of this variable on the test results. However, in several studies the early part of the study was used as a learning curve; thus it is likely that, in general, detection and test accuracy may improve over time with increasing specialist expertise. Therefore, a lack of surgeons' experience would be likely to impact the results in the direction of underestimating test sensitivity.

The number of included studies in the meta-analyses that reported use of pre-operative imaging was small (three and two for 'per groin' and 'per woman' analyses, respectively) and it is unclear whether this was standard procedure in the other studies; therefore it is difficult to make any inferences with regard to this variable.

We did not anticipate the substantial variation in the timing of administering technetium across included studies. Most studies administered technetium on the day before surgery; however 10 studies administered it within six hours of surgery. These studies reported detection rates ranging from 75% to 100% and, when all technetium studies were considered, the mean detection rate was 97% (282/290 women detected). The technetium study with the lowest detection rate (75%) administered the nanocolloid intra-operatively after general anaesthesia ([Klar 2011](#)). Such timing may be more convenient and more comfortable for patients; however, more evidence is needed on the safety and accuracy of this method.

It would be valuable to know the relative costs of the different tests, including their clinical consequences. However, we did not specify economic outcomes a priori and these data were not reported in any of the included studies.

Applicability of findings to the review question

Sentinel node assessment is a technique designed for use in women with early-stage vulval cancer (grade IB or higher) without clinically suspicious nodes. Almost all studies included in this review restricted participants to this group of women. The overall rate of groin node metastases across included studies (32%) is robust to a large observational study of early-stage vulval cancer ([GROINSS-V 2008](#)), in which groin metastases were identified in 33% of women. Therefore, we consider the review findings to be highly applicable.

Index test detection rates have a substantial impact on the clinical pathway of women undergoing sentinel node assessment: for women in whom a sentinel node is not detected, more extensive surgery (complete IFL) is usually indicated (see [Figure 1](#)). As IFL is associated with significantly greater morbidity than sentinel node dissection ([GROINSS-V 2008](#)), index tests with lower detection rates will be associated with a greater risk of morbidity. For example, index tests with detection rates of 82% and 98% will result in 18

versus two women per 100 requiring IFL without evidence of groin metastases, respectively (see [Summary of findings 1](#)). Therefore, index test accuracy cannot be evaluated without considering the index test detection rates. According to the results of this review, the combined test, with a sensitivity of 0.95 and a detection rate of 98% (per woman data), offers the best option for women with early-stage vulval cancer. The technetium test alone had detection and sensitivity rates that were very similar to the combined test estimates. It is possible that other tracer agents, e.g. fluorescent indocyanine green (ICG), may further enhance sentinel node detection rates; however, we did not find any test accuracy studies of this agent to include in this review.

For women undergoing sentinel node assessment with the combined or technetium tests, an estimated one or two women (2% or less) may be 'missed' (see [Summary of findings 1](#)). These women would not receive additional treatment and would be likely to experience a shorter survival than those who were identified as having groin metastases and who received additional treatment (IFL). This finding is consistent with [GROINSS-V 2008](#), in which 3% of sentinel node-negative women experienced groin recurrences within 16 months of the procedure. The three-year survival rate for sentinel node-negative women in [GROINSS-V 2008](#) was also 97%.

The largest study of sentinel node assessment in vulval cancer to date ([Levenback 2012](#)), found that women with vulval lesions greater than 4 cm in diameter were at greater risk of experiencing false negatives on sentinel node assessment. In this study, the NPV of sentinel node assessment for tumours less than 4 cm was 98% compared with 93% for larger tumours. [GROINSS-V 2008](#) only enrolled women with primary tumours less than 4 cm and revealed a similar NPV to [Levenback 2012](#) for this risk group (97%). The NPV for tests evaluated per women in this review ranged from 96% to 98% overall. Few studies contributed data to the meta-analysis according to vulval lesion size of less than 4 cm; therefore the review is unable to provide much evidence in this regard. However, based on the data from [Levenback 2012](#), it is prudent to restrict sentinel node assessment to women with vulval lesions less than 4 cm in diameter. Similarly, this review was unable to clarify the role of sentinel node assessment in women with multifocal lesions. However, limited evidence from [GROINSS-V 2008](#) suggests that women with multifocal lesions may not be suitable candidates for sentinel node assessment.

AUTHORS' CONCLUSIONS

Implications for practice

Sentinel node assessment performed at specialist oncology centres can accurately diagnose groin metastases in women with early vulval cancer and unknown groin node status. In practice, either the combined tests or technetium alone may be employed. Women undergoing sentinel node assessment can be counselled that the risk of the combined or technetium test missing the spread of vulval cancer to the groin lymph nodes is 1% to 3%. This means that one to three women with groin node metastases out of 100 women undergoing the procedure (or out of 30 women with groin lymph node involvement) may be 'missed'. The combined test may miss fewer women than technetium alone. Ultrastaging probably further enhances test accuracy. Using blue dye on its own may increase the number of 'missed' cases to nine per 100 women undergoing the procedure. Both the combined and technetium only tests will

reduce the need for complete inguinofemoral lymphadenectomy (IFL) by approximately 70% and, therefore, reduce the risk of surgical morbidity for women with early vulval cancer.

Implications for research

It is not yet clear how the survival of women with negative sentinel nodes compares to those undergoing standard surgery (IFL). A recent observational study of vulval cancer trends conducted in The Netherlands suggests that the introduction of less radical surgery has not affected survival rates ([Schuurman 2013](#)). In order to prove this definitively, one would have to design a study that randomised women with negative sentinel nodes to IFL or no additional treatment, with survival as the endpoint. Given the rarity of vulval cancer, such a study would be a challenge which would require worldwide co-operation including multiple centres. Furthermore, there are ethical issues in subjecting women with a very small risk of groin metastases to an operation associated with significant morbidity, and which may be unnecessary. It may be possible to design a trial of IFL compared with sentinel node assessment with or without IFL. This would still require huge numbers and therefore may not be feasible, but would be ethically acceptable as the node positivity in both arms would be about 30%, and one of the arms is the current standard treatment.

For this review, it was not possible to determine whether pre-operative radiology has an important role to play in patient selection and management and future research should address this question. Further DTA studies of existing and new technologies should employ assessor blinding to reduce the risk of detection bias. Further studies to evaluate the optimal timing of technetium administration for sentinel node assessment and patient satisfaction with the procedure, may be of value.

ACKNOWLEDGEMENTS

We thank the following:

- Co-ordinating Editor of the Cochrane Gynaecological Cancer Group (CGCG), Jo Morrison, for her clinical expertise;
- CGCG Managing Editors, Gail Quinn and Clare Jess, for their contribution to the editorial process;
- Trial Search Co-ordinator, Anne Eisinga, for designing the search strategy;
- Yemisi Takwoingi, for statistical advise and supplying Stata code for the univariate sensitivity analyses - her input was invaluable to the timely delivery of this review;
- the librarians at the Royal United Hospital in Bath for sourcing the essential reference material; and
- Sambor Sawicki ([Sawicki 2010](#)), Lukas Rob ([Rob 2007](#)), Matteo Morotti ([Morotti 2011](#)), Giuseppe Trifiro ([Trifiro 2010](#)), Reita Nyberg ([Nyberg 2007](#)), Charles Levenback and Shamshad Ali ([Levenback 2012](#)) and Mario Sideri ([Sideri 2000](#)), who provided us with additional information and/or data.

The National Institute for Health Research (NIHR) is the largest single funder of the Cochrane Gynaecological Cancer Group.

The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the NIHR, NHS or the Department of Health.

REFERENCES

References to studies included in this review
Akrivos 2011 {published data only}

* Akrivos N, Rodolakis A, Vlachos G, Sotiropoulou M, Papantoniou V, Biliatis I, et al. Detection and credibility of sentinel node in vulvar cancer: a single institutional study and short review of literature. *Archives of Gynecology and Obstetrics* 2011;**284**:1551-6. [ID 5]

Akrivos N, Rodolakis A, Vlachos G, Sotiropoulou M, Papantoniou V, Biliatis I, et al. Sentinel node in vulvar cancer. To use or not to use technetium?. *International Journal of Gynecological Cancer*. 2011; Vol. 21(Suppl 3):S905.

Ansink 1999 {published data only}

Ansink AC, Sie-Go DM, van der Velden J, Sijmons EA, de Barros Lopes A, Monaghan JM, et al. Identification of sentinel lymph nodes in vulvar carcinoma patients with the aid of a patent blue V injection: a multicenter study. *Cancer* 1999;**86**:652-6.

Basta 2005 {published data only}

* Basta A, Pitynski K, Basta P, Hubaiewska-Hola A, Oplawski M, Przeszlakowski D. Sentinel node in gynaecological oncology. *Reports of Practical Oncology and Radiotherapy* 2005;**10**:91-4.

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 [Znakowanie węzłów limfatycznych i poszukiwanie węzła wartowniczego w raku szyjki macicy, raku endometrium i raku sromu]. *Ginekologia Polska* 2003;**74**:830-5.

Boran 2003 {published data only}

Boran N, Kayikcioglu F, Kir M. Sentinel lymph node procedure in early vulvar cancer. *Gynecologic Oncology* 2003;**90**:492-3.

Camara 2009 {published data only}

Camara O, Gonnert H, Herrmann J, Egbe A, Diebolder H, Gajda M, et al. Sentinel lymph node biopsy in vulvar cancer: a pilot study. *European Journal of Gynaecological Oncology* 2009;**30**:622-4.

Crosbie 2010 {published data only}

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. *Surgical Oncology* 2010;**19**:e150-4.

DeCesare 1997 {published data only}

DeCesare SL, Fiorica JV, Roberts WS, Reintgen D, Arango H, Hoffman MS, et al. A pilot study utilizing intraoperative lymphoscintigraphy for identification of the sentinel lymph nodes in vulvar cancer. *Gynecologic Oncology* 1997;**66**:425-8.

de Hullu 2000 {published data only}

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. *Journal of Nuclear Medicine* 1998;**39**:1381-5.

* 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. *Journal of Clinical Oncology* 2000;**18**:2811-6.

Echt 1999 {published data only}

Echt ML, Finan MA, Hoffman MS, Kline RC, Roberts WS, Fiorica JV. Detection of sentinel lymph nodes with lymphazurin in cervical, uterine, and vulvar malignancies. *Southern Medical Journal* 1999;**92**:204-8.

Goni 2011 {published data only}

Goni E, Estebanez C, Muruzabal J, Serra P, Guarch R, Camarero A, et al. Sentinel lymph node biopsy in vulvar cancer: our experience. *European Journal of Nuclear Medicine and Molecular Imaging* 2011;**38**:S368-9.

Hampl 2008 {published data only}

Hampl M, Hantschmann P, Michels W, Hillemanns P, German Multicenter Study Group. Validation of the accuracy of the sentinel lymph node procedure in patients with vulvar cancer: results of a multicenter study in Germany. *Gynecologic Oncology* 2008;**111**:282-8.

Hauspy 2007 {published data only}

Hauspy J, Beiner M, Harley I, Ehrlich L, Rasty G, Covens A. Sentinel lymph node in vulvar cancer. *Cancer* 2007;**110**:1015-23.

Johann 2008 {published data only}

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. *Gynecologic Oncology* 2008;**110**:324-8.

Klar 2011 {published data only}

* Klar M, Bossart M, Stickeler E, Brink I, Orlowska-Volk M, Denschlag D. Sentinel lymph node detection in patients with vulvar carcinoma: feasibility of intra-operative mapping with technetium-99m-labelled nanocolloid. *European Journal of Surgical Oncology* 2011;**37**:818-23.

Klat 2009 {published data only}

* 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. *Australian & New Zealand Journal of Obstetrics & Gynaecology* 2009;**49**:672-6.

Levenback 2001 {published data only}

Levenback C, Burke TW, Morris M, Malpica A, Lucas KR, Gershenson DM. Intraoperative lymphatic mapping in vulvar cancer. *Obstetrics and Gynecology* 1994;**84**:163-7.

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

* 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. *Gynecologic Oncology* 2001;**83**:276-81.

Levenback 2012 {published data only}

Coleman RL, Ali S, Gold MA, Fowler JM, Judson PL, Levenback CF. When is bilateral lymphadenectomy for midline squamous cell carcinoma of the vulva necessary? An ancillary analysis from GOG-173. *International Journal of Gynecological Cancer* 2011;**21**:S50.

Coleman RL, Ali S, Levenback CF, Gold MA, Fowler JM, Judson PL, et al. Is bilateral lymphadenectomy for midline squamous carcinoma of the vulva always necessary? An analysis from Gynecologic Oncology Group (GOG) 173. *Gynecologic Oncology* 2013;**128**:155-9.

* Levenback CF, Ali S, Coleman RL, Gold MA, 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. *Journal of Clinical Oncology* 2012;**30**:3786-91.

Levenback CF, Tian C, Coleman RL, Gold MA, Fowler JM, Judson PL. Sentinel node (SN) biopsy in patients with vulvar cancer: a Gynecologic Oncology Group (GOG) study. *Journal of Clinical Oncology* 2009;**27**:5505.

Li 2009 {published data only}

Li B, Wu L-y, Liu L, Zhang R, Zhang G-y, Yu G-z. An initial experience of the use of sentinel lymph node biopsy in squamous cell cancer of the vulva. *Zhonghua Fu Chan Ke za Zhi* 2009;**44**:364-8.

Lindell 2010 {published data only}

* 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. *European Journal of Obstetrics, Gynecology, and Reproductive Biology* 2010;**152**:91-5.

Louis-Sylvestre 2006 {published data only}

Louis-Sylvestre C, Evangelista E, Leonard F, Itti E, Meignan M, Paniel BJ. Interpretation of sentinel node identification in vulvar cancer [Interpretation de l'identification du ganglion sentinelle dans les cancers vulvaires]. *Gynecologie, Obstetrique & Fertilité* 2006;**34**:706-10.

* Louis-Sylvestre C, Evangelista E, Leonard F, Itti E, Meignan M, Paniel BJ. Sentinel node localization should be interpreted with caution in midline vulvar cancer. *Gynecologic Oncology* 2005;**97**:151-4.

Martinez-Palones 2006 {published data only}

* 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 vulvar cancer. *Gynecologic Oncology* 2006;**103**:865-70.

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 vulvar cancer: Commentary. *Obstetrical and Gynecological Survey* 2007;**62**:240-2.

Merisio 2005 {published data only}

Merisio C, Berretta R, Gualdi M, Pultrone DC, Anfuso S, Agnese G, et al. Radioguided sentinel lymph node detection in vulvar cancer. *International Journal of Gynecological Cancer* 2005;**15**:493-7.

Moore 2003a {published data only}

Moore RG, DePasquale SE, Steinhoff MM, Gajewski W, Steller M, Noto R, et al. Sentinel node identification and the ability to detect metastatic tumor to inguinal lymph nodes in squamous cell cancer of the vulva. *Gynecologic Oncology* 2003;**89**:475-9.

Morotti 2011 {published and unpublished data}

Morotti M, Valenzano Menada M, Motzo M, Vincentelli A, Casabona F, Ferrero S, et al. Role of routine preoperative lymphoscintigraphy in sentinel node biopsy for vulvar cancer. *International Journal of Gynecological Cancer*. 2011; Vol. 21:S917.

* Morotti M, Valenzano Menada M, Motzo M, Vincentelli A, Villa G, Casabona F, et al. Role of routine preoperative lymphoscintigraphy in sentinel node biopsy for vulvar cancer. Unpublished paper Received via email on 30 July 2013.

Nyberg 2007 {published and unpublished data}

Nyberg RH, Iivonen M, Parkkinen J, Kuoppala T, Maenpaa JU. Sentinel node and vulvar cancer: a series of 47 patients. *Acta Obstetrica et Gynecologica Scandinavica* 2007;**86**:615-9.

Radziszewski 2010 {published data only}

Radziszewski J, Kowalewska M, Jedrzejczak T, Kozłowicz-Gudzinska I, Nasierowska-Guttmejer A, Bidzinski M, et al. The accuracy of the sentinel lymph node concept in early stage squamous cell vulvar carcinoma. *Gynecologic Oncology* 2010;**116**:473-7.

Rob 2007 {published and unpublished data}

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 Gynekologie* 2006;**71**:298-301.

* 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. *International Journal of Gynecological Cancer* 2007;**17**:147-53.

Sawicki 2010 {published and unpublished data}

Sawicki S, Kobierski J, Wydra D. Sentinel lymph node detection in early vulvar cancer. *International Journal of Gynecological Cancer* 2011;**21**:S935.

* Sawicki S, Romanowicz G, Wydra D, Lass P. The usefulness of sentinel lymph node detection in vulvar cancer - a short communication. *Nuclear Medicine Review* 2010;**13**:81-3.

Wydra D, Sawicki S, Emerich J, Romanowicz G. Evaluation of sentinel node detection in vulvar cancer. *Nuclear Medicine Review* 2005;**8**:128-30.

Sideri 2000 {published and unpublished data}

De Cicco C, Sideri M, Bartolomei M, Grana C, Cremonesi M, Fiorenza M, et al. Sentinel node biopsy in early vulvar cancer. *British Journal of Cancer* 2000;**82**:295-9.

* Sideri M, De Cicco C, Maggioni A, Colombo N, Boccione L, Trifiro G, et al. Detection of sentinel nodes by lymphoscintigraphy and gamma probe guided surgery in vulvar neoplasia. *Tumori* 2000;**86**:359-63.

Sliutz 2002 {published data only}

Sliutz G, Reinthaller A, Lantsch T, Mende T, Sinzinger H, Kainz C, et al. Lymphatic mapping of sentinel nodes in early vulvar cancer. *Gynecologic Oncology* 2002;**84**:449-52.

Trifiro 2010 {published data only}

Trifiro G, Travaini LL, Sanvito F, Pacifici M, Mallia A, Ferrari ME, et al. Sentinel node detection by lymphoscintigraphy and sentinel lymph node biopsy in vulvar melanoma. *European Journal of Nuclear Medicine and Molecular Imaging* 2010;**37**:736-41.

Vidal-Sicart 2007 {published data only}

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. *Gynecologic Oncology* 2003;**88**:29-34.

Vidal-Sicart S, Domenech B, Luján B, Pahisa J, Torné A, Martínez-Román S, et al. Sentinel node in gynaecological cancers. Our experience [Ganglio centinela en cánceres ginecológicos: Nuestra experiencia]. *Revista Española de Medicina Nuclear* 2009;**28**:221-8.

* Vidal-Sicart S, Puig-Tintore LM, Lejarcegui JA, Paredes P, Ortega ML, Muñoz A, et al. Validation and application of the sentinel lymph node concept in malignant vulvar tumours. *European Journal of Nuclear Medicine and Molecular Imaging* 2007;**34**:384-91.

Vidal-Sicart S, Puig-Tintore LM, Ordi J, Lejarcegui JA. Localization of sentinel lymph nodes in cancer of the vulva. *Ginecologia y Obstetricia Clinica* 2002;**3**:67-71.

Zambo 2002 {published data only}

Schmidt E, Zambo K, Hartmann T, Dehghani B, Bodis J. Preliminary experiences with sentinel node detection in cases of vulvar malignancy. *Magyar Noorvosok Lapja* 2003;**66**:177-80.

Zambo K, Schmidt E, Hartmann T, Kornya L, Dehghani B, Tinneberg HR, et al. Preliminary experiences with sentinel lymph node detection in cases of vulvar malignancy. *European Journal of Nuclear Medicine and Molecular Imaging* 2002;**29**:1198-200.

Zekan 2012 {published and unpublished data}

Corusic A, Babic D, Zekan J, Mutvar A, Planinic P, Lesin J, et al. Safety of sentinel node biopsy in vulvar cancer FIGO staging IB and II. *International Journal of Gynecological Cancer* 2011;**21**:S900.

* Zekan J, Mutvar A, Huic D, Petrovic D, Karelavic D, Mitrovic L. Reliability of sentinel node assay in vulvar cancer: the first Croatian validation trial. *Gynecologic Oncology* 2012;**126**:99-102.

References to studies excluded from this review
Achimas-Cadariu 2009 {published data only}

Achimas-Cadariu P, Harter P, Fisseler-Eckhoff A, Beutel B, Traut A, Du Bois A. Assessment of the sentinel lymph node in patients with invasive squamous carcinoma of the vulva. *Acta Obstetrica et Gynecologica Scandinavica* 2009;**88**:1209-14.

Ansink 2003 {published data only}

Ansink AC, de Hullu JA, van der Zee AGJ. Re: Further data on the usefulness of sentinel lymph node identification and ultrastaging in vulvar squamous cell carcinoma (in *Gynecol Oncol* 2003 Jan;**88**(1):29-34). *Gynecologic Oncology* 2003;**90**:688-90.

Armstrong 2011 {published data only}

Armstrong F, Reidy F, Walsh T, Boyd W, Jarvenkari N. The role of sentinel node biopsy in vulvar carcinoma. *Irish Journal of Medical Science* 2011;**180**:S156.

Baiocchi 2011 {published data only}

Baiocchi G, Cestari FMS, Cestari LA, Rocha RM, Kumagai LY, Oliveira RAR, et al. Patterns of inguinal lymph node metastasis in patients with vulvar cancer. *International Journal of Gynecological Cancer* 2011;**21**:S895.

Barton 1992 {published data only}

Barton DP, Berman C, Cavanagh D, Roberts WS, Hoffman MS, Fiorica JV, et al. Lymphoscintigraphy in vulvar cancer: a pilot study. *Gynecologic Oncology* 1992;**46**:341-4.

Beneder 2008 {published data only}

Beneder C, Fuechsel FG, Krause T, Kuhn A, Mueller MD. The role of 3D fusion imaging in sentinel lymphadenectomy for vulvar cancer. *Gynecologic Oncology* 2008;**109**:76-80.

Bibi 2011 {published data only}

Bibi G, Grisaru D, Lessing J, Schneebaum S, Ben-Arie A. Accuracy and recurrence-free survival after sentinel lymph node biopsy and lymphadenectomy in patients vulvar cancer. *International Journal of Gynecological Cancer* 2011;**21**:S896.

Brunner 2008 {published data only}

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 Research* 2008;**28**:4091-4.

Bus 2011 {published data only}

Bus K, Uracs I, Wisotzki C, Derlin T, Mester J, Klutmann S, et al. Preliminary experiences with sentinel lymph node scintigraphy in vulvar carcinoma patients. *Nuclear Medicine Review* 2011;**14**:A6.

Cady 2000 {published data only}

* Cady B. Sentinel lymph node procedure in squamous cell carcinoma of the vulva. *Journal of Clinical Oncology* 2000;**18**:2795-7.

Carcopino 2005 {published data only}

Carcopino X, Houvenaeghel G, Buttarelli M, Charaffe-Jauffret E, Gonzague L, Rossi I. Feasibility and morbidity of sentinel lymph node detection in patients with vulvar carcinoma [Faisabilité et morbidité du prelevement du ganglion sentinelle pour les carcinomes épidermoïdes vulvaires]. *Bulletin du Cancer* 2005;**92**:489-97.

Choudhary 2003 {published data only}

Choudhary P. Sentinel lymph node detection in cases of vulvar malignancy. *European Journal of Nuclear Medicine and Molecular Imaging* 2003;**30**:184.

Crane 2011 {published data only}

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 vulvar cancer: first clinical results. *Gynecologic Oncology* 2011;**120**:291-5.

de Hullu 2001 {published data only}

de Hullu JA, Piers DA, Hollema H, Aalders JG, van der Zee AG. Sentinel lymph node detection in locally recurrent carcinoma of the vulva. *British Journal of Obstetrics and Gynaecology* 2001;**108**:766-8.

de Hullu 2002 {published data only}

de Hullu JA, Hollema H, Hoekstra HJ, Piers DA, Mourits MJE, Aalders JG, et al. Vulvar melanoma: is there a role for sentinel lymph node biopsy?. *Cancer* 2002;**94**:486-91.

Devaja 2011 {published data only}

* Devaja O, Mehra G, Coutts M, Adamson S, Montalto SA, Donaldson J, et al. A prospective study of sentinel lymph node detection in vulval carcinoma: is it time for a change in clinical practice?. *International Journal of Gynecological Cancer* 2011;**21**:559-64.

Dhar 2005 {published data only}

* Dhar KK, Woolas RP. Lymphatic mapping and sentinel node biopsy in early vulvar cancer. *British Journal of Obstetrics and Gynaecology* 2005;**112**:696-702.

Ennik 2011 {published data only}

Ennik TA, Allen DG, Bekkers RLM, Hyde SE, Grant PT. Effects of previous surgery on the detection of sentinel nodes in women with vulvar cancer. *International Journal of Gynecological Cancer* 2011;**21**:1679-83.

Farrell 2010 {published data only}

* Farrell C, Lee ST, Grant MP, Rowe C. Rapid localisation of Sentinel Lymph nodes in vulvar lymphoscintigraphy. *Internal Medicine Journal* 2010;**40**:26-7. [64]

Freudenberg 2010 {published data only}

Freudenberg LS, Gortz E, Hagen C, Harms E, Koska WW, Marlowe RJ, et al. Lymphatic mapping using SPECT/CT in vulvar carcinoma. *Clinical Nuclear Medicine* 2010;**35**:950-2.

Garcia-Iglesias 2012 {published data only}

Garcia-Iglesias A, Rodriguez-Martin MO, Ruano R, Beltran D, Penalosa L, Hernandez-Barreiro B, et al. Sentinel node dissection in the treatment of early stages of vulvar cancer. *European Journal of Gynaecological Oncology* 2012;**33**:151-4.

GROINSS-V 2008 {published data only}

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. *The Lancet Oncology* 2010;**11**:646-52.

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 vulvar cancer mandates adjuvant groin treatment, independent of size. *Gynecologic Oncology* 2009;**112**:S14-5.

* 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 vulvar cancer. *Journal of Clinical Oncology* 2008;**26**:884-9.

Hefler 2008 {published data only}

Hefler LA, Grimm C. Inguinal sentinel lymph node dissection vs complete inguinal lymph node dissection in patients with vulvar cancer. *Anticancer Research* 2008;**28**:515-7.

Kraft 2012 {published data only}

Planar scintigraphy, SPECT/CT of SLN in various types of tumours. Estimation of some factors influencing detection. *European Journal of Nuclear Medicine and Molecular Imaging* 2012;**39**:S493.

Levenback 2000 {published data only}

* Levenback C, Coleman RL. Re: Terada et al. Sentinel node dissection and ultrastaging in squamous cell cancer of the vulva in *Gynecol Oncol* 76:40-44, 2000. *Gynecologic Oncology* 2000;**77**:484-5. [#44]

Makowski 2010 {published data only}

Makowski L, Tjahjadi M, Grubner S, Friedrich Gratz K, Hillemanns P, Hertel H. SPECT-CT improved the detection of inguinal sentinel lymph nodes in vulvar cancer. *Archives of Gynecology and Obstetrics* 2010;**282**:S199.

Maza 2007 {published data only}

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. *European Journal of Nuclear Medicine and Molecular Imaging* 2007;**34**:378-83.

Molpus 2001 {published data only}

* Molpus KL, Kelley MC, Johnson JE, Martin WH, Jones HW. Sentinel lymph node detection and microstaging in vulvar carcinoma. *Journal of Reproductive Medicine* 2001;**46**:863-9.

Moore 2003b {published data only}

Moore RG, Granai CO, Gajewski W, Gordinier M, Steinhoff MM. Pathologic evaluation of inguinal sentinel lymph nodes in vulvar cancer patients: a comparison of immunohistochemical staining versus ultrastaging with hematoxylin and eosin staining. *Gynecologic Oncology* 2003;**91**:378-82.

Moore 2008 {published data only}

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. *Gynecologic Oncology* 2008;**109**:65-70.

Nickles Fader 2012 {published data only}

Nickles Fader A, Ducie J, Abu-Rustum N, Escobar P, Levenback C, Moore K, et al. The utility of sentinel lymph node biopsy in the surgical management of women with vulvar melanoma. *Gynecologic Oncology* 2012;**125**:S84.

Penson 2001 {published data only}

Penson RT, Fuller AF Jr. Nodal metastasis is highly consistent in squamous cell carcinoma of the vulva. *Journal of Clinical Oncology* 2001;**19**:2364-5.

Puig Calvo 2011 {published data only}

Puig Calvo O, Bajen M, Rodriguez-Gasen A, Notta P, Ricart Y, Sabate-Llobera A, et al. Reliability of sentinel node biopsy in early stage vulvar carcinoma. *European Journal of Nuclear Medicine and Molecular Imaging* 2011;**38**:S369.

Radziszewski 2003 {published data only}

Radziszewski J, Bidzinski M, Panek G, Sobiczewski P, Derlatka P, Nasierowska-Guttmejer A, et al. Sentinel lymph node in vulvar cancer - a pilot study to identify and assess the diagnostic value. *Nowotwory* 2003;**53**:270-4.

Robison 2006 {published data only}

Robison K, Steinhoff MM, Granai CO, Brard L, Gajewski W, Moore RG. Inguinal sentinel node dissection versus standard inguinal node dissection in patients with vulvar cancer: a comparison of the size of metastasis detected in inguinal lymph nodes. *Gynecologic Oncology* 2006;**101**:24-7.

* Robison K, Steinhoff MM, Granai CO, Brard L, Gajewski W, Moore RG. Inguinal sentinel node dissection versus standard inguinal node dissection in patients with vulvar cancer: a comparison of the size of metastasis detected in inguinal lymph nodes. *Obstetrical and Gynecological Survey* 2006;**61**:518-9.

Rodier 1999 {published data only}

* Rodier JF, Janser JC, Routiot T, David E, Ott G, Schneegans O, et al. Sentinel node biopsy in vulvar malignancies: a preliminary feasibility study. *Oncology Reports* 1999;**6**:1249-52.

Soliman 2012 {published data only}

Soliman AA, Heubner M, Kimmig R, Wimberger P. Morbidity of inguinofemoral lymphadenectomy in vulvar cancer. *The Scientific World Journal* 2012;**2012**:341253.

Sun 2009 {published data only}

Sun Y, Li B, Wu L, Zhang R, Yu G, Zhang G. Role of sentinel lymph node biopsy in the management of vulvar melanoma. *Chinese Journal of Clinical Oncology* 2009;**36**:253.

Tavares 2001 {published data only}

* Tavares MGM, Sapienza MT, Galeb NA, Belfort FA, Costa RR, Osorio CA, et al. The use of 99mTc-phytate for sentinel node mapping in melanoma, breast cancer and vulvar cancer: a study of 100 cases. *European Journal of Nuclear Medicine* 2001;**28**:1597-604.

Tenney 2011 {published data only}

Tenney M, Bishop E, Nugent E, Landrum L, Moore K, Mannel R, et al. Positive vulvar sentinel lymph node biopsy: a single-institution retrospective review. *Gynecologic Oncology* 2011;**120**:S119.

Terada 2000 {published data only}

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. *Gynecologic Oncology* 1998;**70**:65-9.

Terada KY, Shimizu DM, Wong JH. Sentinel node dissection and ultrastaging in squamous cell cancer of the vulva. *Gynecologic Oncology* 2000;**76**:40-4.

Terada 2006 {published data only}

Terada KY, Shimizu DM, Jiang CS, Wong JH. Outcomes for patients with T1 squamous cell cancer of the vulva undergoing sentinel node biopsy. *Gynecologic Oncology* 2006;**102**:200-3.

Tjin Asjoe 2008 {published data only}

Tjin Asjoe FM, van Bekkum E, Ewing P, Burger CW, Ansink AC. Sentinel node procedure in vulvar squamous cell carcinoma: a histomorphologic review of 32 cases. The significance of anucleate structures on immunohistochemistry. *International Journal of Gynecological Cancer* 2008;**18**:1032-6.

Vakselj 2007 {published data only}

Vakselj A, Bebar S. The role of sentinel lymph node detection in vulvar carcinoma and the experiences at the Institute of Oncology Ljubljana. *Radiology and Oncology* 2007;**41**:167-73.

Van Den Eynden 2003 {published data only}

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

van der Velden 2006 {published data only}

van der Velden J, Ansink AC. Re: "Outcomes of stage I/II vulvar cancer patients after negative superficial inguinal lymphadenectomy". *Gynecologic Oncology* 2006;**100**:219-20.

Wechter 2004 {published data only}

Wechter ME, Gruber SB, Haefner HK, Lowe L, Schwartz JL, Reynolds KR, et al. Vulvar melanoma: a report of 20 cases and review of the literature. *Journal of the American Academy of Dermatology* 2004;**50**:554-62.

Additional references
Abang Mohammed 2000

Abang Mohammed DK, Uberoi R, de B Lopes A, Monaghan JM. Inguinal node status by ultrasound in vulva cancer. *Gynecologic Oncology* 2000;**77**:93-6.

Andreasson 1985

Andreasson B, Nyboe J. Predictive factors with reference to low-risk of metastases in squamous cell carcinoma in the vulvar region. *Gynecologic Oncology* 1985;**21**:196-206.

Boyce 1985

Boyce J, Fruchter RG, Kasambilides E, Nicastrì AD, Sedlis A, Remy JC. Prognostic factors in carcinoma of the vulva. *Gynecologic Oncology* 1985;**20**:364-77.

Burger 1995

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. *Gynecologic Oncology* 1995;**57**:327-34.

Canavese 2010

Canavese G, Bruzzi P, Catturich A, Vecchio C, Tomei D, Carli F, et al. Intra-operative evaluation of the sentinel lymph node for T1-N0 breast-cancer patients: always or never? A risk/benefit and cost/benefit analysis. *European Journal of Surgical Oncology* 2010;**36**:737-44.

Cohn 2002

Cohn DE, Dehdashti F, Gibb RK, Mutch DG, Rader JS, Siegel BA, et al. Prospective evaluation of positron emission tomography for the detection of groin node metastases from vulvar cancer. *Gynecologic Oncology* 2002;**85**:179-84.

Creasman 1997

Creasman WT, Phillips JL, Menck HR. The National Cancer Data Base report on early stage invasive vulvar carcinoma. The American College of Surgeons Commission on Cancer and the American Cancer Society. *Cancer* 1997;**80**:505-13.

Curry 1980

Curry SL, Wharton JT, Rutledge F. Positive lymph nodes in vulvar squamous carcinoma. *Gynecologic Oncology* 1980;**9**:63-7.

De Cicco 2000

De Cicco C, Sideri M, Bartolomei M, Grana C, Cremonesi M, Fiorenza M, et al. Sentinel node biopsy in early vulvar cancer. *British Journal of Cancer* 2000;**82**:295-9.

de Hullu 1998

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. *Journal of Nuclear Medicine* 1998;**39**(8):1381-5.

de Hullu 1999

de Hullu JA, Pruim J, Que TH, Aalders JG, Boonstra H, Vaalburg W, et al. Noninvasive detection of inguino-femoral lymph node metastases in squamous cell cancer of the vulva by L-[1-¹¹C]-tyrosine positron emission tomography. *International Journal of Gynecological Cancer* 1999;**9**:141-6.

Endnote 2012 [Computer program]

Thomson Reuters. EndNote: Reference Manager. Version X5. Thomson Reuters, 2012.

Fotiou 1996

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. *European Journal of Gynaecological Oncology* 1996;**17**:67-73.

Gaarenstroom 2003

Gaarenstroom KN, Kenter GG, Trimboos JB, Agous I, Amant F, Peters AA, et al. Postoperative complications after vulvectomy and inguino-femoral lymphadenectomy using separate groin incisions. *International Journal of Gynecological Cancer* 2003;**13**:522-7.

Hacker 1981

Hacker NF, Leuchter RS, Berek JS, Castaldo TW, Lagasse LD. Radical vulvectomy and bilateral inguinal lymphadenectomy through separate groin incisions. *Obstetrics and Gynecology* 1981;**58**:574-9.

Hacker 1993

Hacker NF, Van der Velden J. Conservative management of early vulvar cancer. *Cancer* 1993;**71**:1673-7.

Hall 2003

Hall TB, Barton DP, Trott PA, Nasiri N, Shepherd JH, Thomas JM, et al. The role of ultrasound-guided cytology of groin lymph nodes in the management of squamous cell carcinoma of the vulva: 5-year experience in 44 patients. *Clinical Radiology* 2003;**58**:367-71.

Hawnaur 2002

Hawnaur JM, Reynolds K, Wilson G, Hillier V, Kitchener HC. Identification of inguinal lymph node metastases from vulval carcinoma by magnetic resonance imaging: an initial report. *Clinical Radiology* 2002;**57**:995-1000.

Homesley 1991

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). *American Journal of Obstetrics and Gynecology* 1991;**164**:997-1003; discussion 1003-4.

Hudson 2004

Hudson CN, Shulver H, Lowe DC. The surgery of 'inguino-femoral' lymph nodes: is it adequate or excessive?. *International Journal of Gynecological Cancer* 2004;**14**:841-5.

Iversen 1981

Iversen T. Squamous cell carcinoma of the vulva. Localization of the primary tumor and lymph node metastases. *Acta Obstetrica et Gynecologica Scandinavica* 1981;**60**:211-4.

Katz 2003

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. *International Journal of Radiation Oncology, Biology, Physics* 2003;**57**:409-18.

Knopp 2005

Knopp S, Holm R, Trope C, Nesland JM. Occult lymph node metastases in early stage vulvar carcinoma patients. *Gynecologic Oncology* 2005;**99**:383-7.

Krag 2007

Krag DN, Anderson SJ, Julian TB, Brown AM, Harlow SP, Ashikaga T, et al. Technical outcomes of sentinel-lymph-node resection and conventional axillary-lymph-node dissection in patients with clinically node-negative breast cancer: results from the NSABP B-32 randomised phase III trial. *Lancet Oncology* 2007;**8**:881-8.

Land 2006

Land R, Herod J, Moskovic E, King M, Sohaib SA, Trott P, et al. Routine computerized tomography scanning, groin ultrasound with or without fine needle aspiration cytology in the surgical management of primary squamous cell carcinoma of the vulva. *International Journal of Gynecological Cancer* 2006;**16**:312-7.

Levenback 1995

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

Louis-Sylvestre 2005

Louis-Sylvestre C, Evangelista E, Leonard F, Itti E, Meignan M, Paniel BJ. Sentinel node localization should be interpreted with caution in midline vulvar cancer. *Gynecologic Oncology* 2005;**97**:151-4.

Makela 1993

Makela PJ, Leminen A, Kaariainen M, Lehtovirta P. Pretreatment sonographic evaluation of inguinal lymph nodes in patients with vulvar malignancy. *Journal of Ultrasound in Medicine* 1993;**12**:255-8.

Morton 1990

Morton DL. Current management of malignant melanoma. *Annals of Surgery* 1990;**212**:123-4.

Morton 2006

Morton DL, Thompson JF, Cochran AJ, Mozzillo N, Elashoff R, Essner R, et al. Sentinel-node biopsy or nodal observation in melanoma. *New England Journal of Medicine* 2006;**355**:1307-17.

Moskovic 1999

Moskovic EC, Shepherd JH, Barton DP, Trott PA, Nasiri N, Thomas JM. The role of high resolution ultrasound with guided cytology of groin lymph nodes in the management of squamous cell carcinoma of the vulva: a pilot study. *British Journal of Obstetrics and Gynaecology* 1999;**106**:863-7.

ONS 2009

Office for National Statistics (ONS). Mortality Statistics: Deaths Registered in England and Wales 2009 (Series DR). <http://www.ons.gov.uk/ons/rel/vsob1/mortality-statistics--deaths-registered-in-england-and-wales--series-dr-/2009/index.html> (accessed 29 January 2013).

Parker 1975

Parker RT, Duncan I, Rampone J, Creasman W. Operative management of early invasive epidermoid carcinoma of the vulva. *American Journal of Obstetrics and Gynecology* 1975;**123**:349-55.

Pecorelli 2009

Pecorelli S. Revised FIGO staging for carcinoma of the vulva, cervix, and endometrium. *International Journal of Gynaecology and Obstetrics* 2009;**105**(2):103-4.

Podratz 1983

Podratz KC, Symmonds RE, Taylor WF, Williams TJ. Carcinoma of the vulva: analysis of treatment and survival. *Obstetrics and Gynecology* 1983;**61**:63-74.

Puig-Tintore 2003

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. *Gynecologic Oncology* 2003;**88**:29-34.

Reitsma 2005

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

RevMan 2012 [Computer program]

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

Rodier 2007

Rodier JF, Velten M, Wilt M, Martel P, Ferron G, Vaini-Elies V, et al. Prospective multicentric randomized study comparing periareolar and peritumoral injection of radiotracer and blue dye for the detection of sentinel lymph node in breast sparing procedures: FRANSENODE trial. *Journal of Clinical Oncology* 2007;**25**:3664-9.

Rouzier 2003

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. *Journal of the American College of Surgery* 2003;**196**:442-50.

Sankaranarayanan 2006

Sankaranarayanan R, Ferlay J. Worldwide burden of gynaecological cancer: the size of the problem. *Best Practice & Research Clinical Obstetrics & Gynaecology* 2006;**20**:207-25.

Saraiya 2008

Saraiya M, Watson M, Wu X, King JB, Chen VW, Smith JS, et al. Incidence of in situ and invasive vulvar cancer in the US, 1998-2003. *Cancer* 2008;**113**:2865-72.

Schuurman 2013

Schuurman MS, van den Einden LC, Massuger LF, Kiemeny LA, van der Aa MA, de Hullu JA. Trends in incidence and survival of Dutch women with vulvar squamous cell carcinoma. *European Journal of Cancer* 2013;**49**(18):3872-80.

Simonsen 1984

Simonsen E. Invasive squamous cell carcinoma of the vulva. *Annales Chirurgiae et Gynaecologiae* 1984;**73**:331-8.

Smyczek-Gargya 1997

Smyczek-Gargya B, Volz B, Geppert M, Dietl J. A multivariate analysis of clinical and morphological prognostic factors in squamous cell carcinoma of the vulva. *Gynecologic and Obstetric Investigation* 1997;**43**:261-7.

Sohaib 2002

Sohaib SA, Richards PS, Ind T, Jeyarajah AR, Shepherd JH, Jacobs IJ, et al. MR imaging of carcinoma of the vulva. *American Journal of Roentgenology* 2002;**178**:373-7.

Stata 2012 [Computer program]

StataCorp LP. Intercooled Stata 12.0 for Windows. Version 12.0. Texas: StataCorp LP, 2012.

Stehman 1992

Stehman FB, Bundy BN, Dvoretzky 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. *Obstetrics and Gynecology* 1992;**79**(4):490-7.

Stehman 2007

Stehman FB. Chapter 8: Invasive cancer of the vulva. In: Disaia P, Creasman W editor(s). *Clinical Gynaecological Oncology*. 7th Edition. Mosby, Elsevier, 2007:240-69.

Sutton 1991

Sutton GP, Miser MR, Stehman FB, Look KY, Ehrlich CE. Trends in the operative management of invasive squamous carcinoma of the vulva at Indiana University, 1974 to 1988. *American Journal of Obstetrics and Gynecology* 1991;**164**:1472-8; discussion 1478-81.

Thompson 2007

Thompson JF, Shaw HM. Sentinel node mapping for melanoma: results of trials and current applications. *Surgical Oncology Clinics of North America* 2007;**16**:35-54.

Van der Zee 2008

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. *Journal of Clinical Oncology* 2008;**26**(6):884-9.

Wang 2011

Wang Z, Wu LC, Chen JQ. Sentinel lymph node biopsy compared with axillary lymph node dissection in early breast cancer: a meta-analysis. *Breast Cancer Research and Treatment* 2011;**129**(3):675-89.

Whiting 2011

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

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies [ordered by study ID]

Akrivos 2011
Study characteristics

Patient sampling	Prospective study; consecutive enrolment
Patient characteristics and setting	34 women with vulval SCC (64 groins); 4 women had lateralised tumours and 30 had mid-line or near midline disease Excluded women with clinically suspicious lymph nodes. Included if tumour depth > 1 mm and size < 8 cm

Akrivos 2011 (Continued)

Median age: 69 years (35 to 86)

Setting: tertiary institutions in the UK and the Netherlands

Index tests	Blue dye only (27 women) Blue dye with Tc-99m (7 women) Histological methods: ultrastaging with IHC staining was performed for SNs that were negative with standard H&E stains; standard sections with H&E stains were performed for other nodes (reference standard (RS)).
Target condition and reference standard(s)	TC: groin lymph node involvement RS: complete IFL
Flow and timing	Tc-99m (0.5 mCi) was injected intradermally 24 hours pre-operative at 4 sites around the primary lesion. LSG was performed and the position of the SNs were marked on the skin. At the time of the operation, after general anaesthesia, blue dye (1 ml blue dye + 1 ml normal saline) was injected around the tumour at 4 sites. The node was considered 'hot' if the activity was > 5% of that at the injection site or > 10% of the 'hottest' SN SND and RS were performed during the same operation No withdrawals occurred
Comparative	
Notes	We were unable to extract separate 2 x 2 data for the combined technique only. Additional data were requested 28 May 2013 Investigators observed a difference in NPV by tumour size: NPV was 100% for tumours ≤ 4 cm (25/25 women) and 50% (4/9 women) for tumours > 4 cm Tc-99m was not superior to blue dye in detecting SN (42/50 versus 50/64, P value = 0.65)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	Low
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Low	Low
DOMAIN 3: Reference Standard			

Akrivos 2011 (Continued)

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear

Unclear

Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Were other imaging tests performed prior to the index test to rule out groin lymph node metastases? Unclear

Unclear

Ansink 1999
Study characteristics

Patient sampling Prospective study

Patient characteristics and setting 51 women with vulval SCC (93 groins); 9 women had lateralised vulval cancer and 42 had midline or near midline disease. It is unclear whether at least 90% of lesions were \geq stage IB

Excluded women with clinically suspicious lymph nodes or other metastases

Median age: 70 years (34 to 90)

Setting: tertiary institutions in the UK and the Netherlands

Index tests SN detection by blue dye

The histological methods used were not reported

Target condition and reference standard(s) TC: groin lymph node involvement

RS: complete IFL

Flow and timing After GA, 1 to 2 ml blue dye was injected intradermally, circumferentially, around the tumour at the start of surgery. Blue dye injection was repeated every 20 minutes during surgery until the SN was identified and dissected

Ansink 1999 (Continued)

SND and RS were performed during the same operation
 No withdrawals occurred

Comparative

Notes Midline/near midline was defined as within 1 cm of the midline

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Unclear	Unclear
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear		
		Unclear	

Basta 2005
Study characteristics

Patient sampling	Prospective study; sampling methods are unclear. This study also included women with endometrial and cervical cancer but data were reported separately
Patient characteristics and setting	39 women with vulval cancer FIGO stage I/II were included; number of groins was not reported. Number of midline/lateral lesions and tumour size was not reported Median age not reported Setting: a teaching hospital in Poland; enrolment period not reported
Index tests	Tc-99m and blue dye Histological methods included standard sections stained with H&E and IHC stains
Target condition and reference standard(s)	TC: groin lymph node involvement RS: complete IFL
Flow and timing	A total of 2.5 mCi Tc-99m was injected peritumourally 14 to 18 hours pre-operatively. 2 hours later, LSG was done. Blue dye (2 to 4 ml) was injected intra-operatively 20 to 30 minutes before mapping SND and IFL were performed during the same operation Withdrawals, if any, were not described
Comparative	
Notes	Limited details reported; no data per groin (requested 21 June 2013)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
		Unclear	Unclear
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			

Basta 2005 (Continued)

Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
	Unclear Low
DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear
	Unclear

Boran 2003
Study characteristics

Patient sampling	Prospective pilot study; consecutive enrolment
Patient characteristics and setting	<p>10 women with vulval cancer stage T1/T2 (17 groins). It is unclear whether any women with T1a (stage IA) lesions were included</p> <p>3 had lateralised tumours and 7 had midline tumours</p> <p>Excluded women with clinically suspicious lymph nodes</p> <p>Median age: NR</p> <p>Setting: a tertiary institution in Turkey between April 2000 and April 2002</p>
Index tests	<p>Tc-99m only</p> <p>Histological methods: IHC stain performed on SNs if H&E was negative</p>
Target condition and reference standard(s)	<p>TC: groin lymph node involvement</p> <p>RS: complete IFL</p>
Flow and timing	<p>Approximately 45 to 60 minutes pre-operatively, 0.4 to 0.6 ml Tc-99m was injected circumferentially intradermally. A hand-held gamma counter was used to identify 'hot' nodes. After the first SN was removed, the groin was re-examined and dissection was continued if more 'hot' nodes were identified (defined as > 10% of the 'hottest' SN)</p> <p>SND and RS were performed during the same operation</p> <p>Withdrawals, if any, were not described</p>
Comparative	

Boran 2003 (Continued)

Notes

A brief report (letter to the editor). Investigators state that being in the "learning curve" of performing the procedure may have affected the results

ICH staining did not reveal additional nodal involvement

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
		Unclear	Unclear
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	No		
		High	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear		
		Unclear	

Camara 2009
Study characteristics

Patient sampling	Prospective pilot study
Patient characteristics and setting	<p>17 women with vulval cancer stage I/II: 16 with SCC, 1 with melanoma; number of groins dissected was not reported. It is unclear whether any women with stage IA were included</p> <p>Median age: 75 years (37 to 83)</p> <p>Setting: a tertiary institution in Germany from February 2003 to March 2007</p>
Index tests	<p>Tc-99m and blue dye</p> <p>Histological methods and ultrastaging not described.</p>
Target condition and reference standard(s)	<p>TC: groin lymph node involvement</p> <p>RS: complete IFL</p>
Flow and timing	<p>Tc-99m and blue dye injected intradermally at 4 sites around the tumour. Timing and other details were not reported</p> <p>SND and RS were performed during the same operation</p> <p>Withdrawals, if any, were not described</p>
Comparative	
Notes	Brief report. Results were not reported per groin

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
		Unclear	Unclear
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Camara 2009 (Continued)

Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear

Low
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Were other imaging tests performed prior to the index test to rule out groin lymph node metastases? Unclear

Unclear
Crosbie 2010
Study characteristics

Patient sampling Prospective study; sampling method not described; mean follow-up was 5 years

Patient characteristics and setting 32 women with clinical stage I/II vulval SCC < 4 cm; depth > 1 mm; histologically confirmed; 17 midline and 15 lateralised tumours

No nodal involvement evident clinically or radiologically

Median age: 67 years (34 to 94)

Setting: tertiary referral hospital in the UK; recruitment from 2002 to 2006

Index tests Tc-99m and blue dye

Histological methods: ultrastaging with IHC staining was performed for SNs that were negative with standard section and H&E stains

Target condition and reference standard(s) TC: groin lymph node involvement

RS: complete IFL

Flow and timing 24 hours pre-operatively Tc-99m (40 MBq in a total volume of 0.2 ml) was injected as 4 intradermal peritumoural injections (or around scar) and LSG acquired immediately for 20 to 30 minutes, up to 1 to 2 hours if nodes were not visualised. SNs were marked on the skin. Pre-operatively 3 ml blue dye was injected at the same perilesional locations

Radioactivity was detected intra-operatively by gamma probe and blue dye was detected by visual identification of blue stained node and draining lymphatics

SND and IFL were performed during the same procedure

Comparative

Notes 'Midline' lesions were not defined

Crosbie 2010 (Continued)

Approximately half of the women in this cohort underwent excisional biopsy of the primary lesion prior to inclusion in the study. Significantly fewer SNs were detected in these women (2.6 versus 1.8; P value = 0.03). 2 women were excluded as they were unfit for surgery

The only false negative case occurred without an obvious explanation/association, such as obesity, nodal enlargement or complete replacement of the node by tumour. The investigators postulated that excision of the primary lesion may compromise SN detection and predispose to false negatives. They advocate performing incisional biopsy only

There were no groin recurrences or distant metastases in women with negative SNs during clinical follow-up (mean 5 years, range 33 to 84 months)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Unclear	Low
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Crosbie 2010 (Continued)

Were all patients included in the analysis? Yes

Were other imaging tests performed prior to the index test to rule out groin lymph node metastases? Unclear

Unclear

de Hullu 2000
Study characteristics

Patient sampling	Prospective study; consecutive enrolment; 7 women refused to participate. No post-test exclusions
Patient characteristics and setting	59 women with vulval SCC stages T1/T2 (107 groins). It is unclear whether at least 90% of lesions were \geq stage IB 11 women had lateralised tumours and 48 had midline or near midline disease No obvious nodes Median age: 69 years (33 to 92) Setting: 2 centres in the Netherlands from July 1996 to July 1999
Index tests	SN detection by Tc-99m and blue dye Histological methods: standard sections stained with H&E. Ultrastaging with IHC stain performed if H&E was negative
Target condition and reference standard(s)	TC: groin lymph node involvement RS: complete IFL
Flow and timing	Tc-99m (0.2 to 0.6 ml) was injected intradermally and circumferentially around the tumour 1 day before surgery. LSG was performed and the position of the SNs were marked on the skin. 2 ml blue dye was injected around the tumour at the time of surgery SND and RS were performed during the same operation Withdrawals explained
Comparative	
Notes	The extent of the procedure (bilateral versus unilateral) was determined by the proximity to the midline: if within 1 cm of midline, bilateral lymphadenectomy was performed

Methodological quality

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

de Hullu 2000 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Yes	
Was a case-control design avoided?	Yes	
Did the study avoid inappropriate exclusions?	Yes	
		Low Unclear

DOMAIN 2: Index Test Test group

Had the test operator performed 10 or more procedures?	Unclear	
		Unclear Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes	
		Low Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear	
		Unclear

DeCesare 1997
Study characteristics

Patient sampling	Prospective study; consecutive sampling; 11 women enrolled but only 10 women underwent IFL (RS) (1 withdrawal due to refusal)
Patient characteristics and setting	10 women with vulval SCC; 6 T1, 2 T2 and 2 T3. T1b not specified 1 woman (10%) had clinically suspicious nodes Median age: NR

DeCesare 1997 (Continued)

Setting: a tertiary institution in the USA; enrolment dates not specified

Index tests	Tc-99m only Histological methods were not reported
Target condition and reference standard(s)	TC: groin lymph node involvement RS: complete IFL
Flow and timing	Intra-operatively, after GA, 400 mCi Tc-99m injected peritumourally or at the prior tumour scar. Vulvectomy was performed first, followed by the SN assessment/biopsy to allow time for clearance of the background gamma counts from the nodal basin. SND and RS were performed during the same operation 1 withdrawal was described
Comparative	
Notes	Limited clinical characteristics reported

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	Unclear
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

DeCesare 1997 (Continued)

Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear
Unclear	

Echt 1999
Study characteristics

Patient sampling	Prospective study
Patient characteristics and setting	<p>Included women with endometrial (8), cervical (13) and vulval cancer (12). Of the women with vulval SCC, 11 had bilateral LND and 1 had unilateral LND (23 groin nodes)</p> <p>Unclear whether women with suspicious nodes were excluded</p> <p>Median age: NR</p> <p>Setting: tertiary hospitals affiliated with 2 institutions in the USA from January 1993 to March 1995</p>
Index tests	<p>Blue dye (lymphazurin) only</p> <p>Histological methods and ultrastaging not described</p>
Target condition and reference standard(s)	<p>TC: groin lymph node involvement</p> <p>RS: complete IFL</p>
Flow and timing	<p>Blue dye (1 ml) was injected around lesion or scar after general anaesthesia</p> <p>SND and RS were performed during the same operation</p> <p>Withdrawals, if any, were not described</p>
Comparative	
Notes	A brief report

Methodological quality

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

Echt 1999 (Continued)

Unclear Unclear

DOMAIN 2: Index Test Test group

Had the test operator performed 10 or more procedures? Unclear

Unclear Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear

Unclear Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Were other imaging tests performed prior to the index test to rule out groin lymph node metastases? Unclear

Unclear

Goni 2011
Study characteristics

Patient sampling Prospective study

Patient characteristics and setting 24 women with vulval cancer (22 SCC; 2 other), clinical stage I and II. Number of groins evaluated is unclear
 Excluded women with clinically suspicious groin nodes
 Mean age: 71 (43 to 85)
 Setting: a tertiary institution in Spain; enrolment from 2001 to 2011

Index tests Tc-99m only
 Standard histological methods were used for SN assessment and IHC staining was performed if H&E stain was negative

Target condition and reference standard(s) TC: groin lymph node involvement
 RS: complete IFL

Goni 2011 (Continued)

Flow and timing

 Procedure: 24 hours pre-operatively, Tc-99m was injected and LSG performed (details not reported). Radioactivity was detected intra-operatively by hand-held gamma probe. SND and RS were performed during the same operation
 Withdrawals were not described
 Mean follow-up was 23 months

Comparative

Notes

Conference abstract only

The number of women with midline/lateralised tumours and who underwent bilateral or unilateral IFL is not stated

Midline lesions were not defined

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
		Unclear	Low
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		

Goni 2011 (Continued)

Were other imaging tests performed prior to the index test to rule out groin lymph node metastases? Unclear

Unclear

HAMPL 2008

Study characteristics

Patient sampling	Prospective study; sampling methods unclear; 13 withdrawals not described
Patient characteristics and setting	127 women with primary T1-3 vulval SCC; 230 groins. One woman had adenocarcinoma. 66 lesions were 'midline', 49 were lateral Excluded if clinically suspicious nodes (detected by ultrasound), proven lymph node metastasis or unresectable tumours Median age: 61.4 years (range not reported) Setting: 7 university or teaching hospitals in Germany; recruitment from 2003 to 2006
Index tests	Tc-99m and blue dye (72 women) Tc-99m only (47) Blue dye only (8 women) Histological methods included ultrastaging with IHC staining if standard method was negative
Target condition and reference standard(s)	TC: groin lymph node involvement RS: complete IFL
Flow and timing	On day of operation, 0.2 ml 60 to 120 MBq Tc-99m was injected intradermally at 4 sites around the tumour. LSG imaging was performed immediately and every 30 minutes thereafter. If detection failed, imaging was repeated the next day. Blue dye was optional (0.5 to 1 ml) and was injected at 4 sites 5 to 10 minutes before the skin incision. SNs were detected intra-operatively by gamma probe and visually (blue nodes) SND and RS were performed at the same operation
Comparative	
Notes	Midline was defined as within 1 cm of the midline 21 unilateral and 103 bilateral IFL performed; for 3 women there were no data All 5 false negatives (4 women) had midline tumours close to the urethra; 2 had tumours ≥ 4 cm Mixed tests used; combined results reported

Methodological quality

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

HAMPL 2008 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Unclear	
Was a case-control design avoided?	Yes	
Did the study avoid inappropriate exclusions?	Unclear	
	Unclear	Low

DOMAIN 2: Index Test Test group

Had the test operator performed 10 or more procedures?	Unclear	
	Unclear	Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	
	Unclear	Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Yes	
	Unclear	

Hauspy 2007
Study characteristics

Patient sampling	Prospective study; consecutive enrolment; 42 recruited; 1 exclusion
Patient characteristics and setting	41 women with vulval cancer (39 SCC, 2 melanomas) at clinical stages T1/2 (68 groins). It is unclear whether any stage IA lesions were included 11 women had lateralised lesions and 30 had midline or near midline lesions No palpable nodes

Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer (Review)

Hauspy 2007 (Continued)

Median age: 65 years (34 to 94)

Setting: a tertiary institution in Canada from April 2004 to September 2006

Index tests	SN detection by Tc-99m with or without blue dye Histological methods: frozen section and ultrastaging was performed for SNs with H&E and IHC stains; standard sections with H&E stains were performed for other nodes (RS)
Target condition and reference standard(s)	TC: groin lymph node involvement RS: complete IFL
Flow and timing	Tc-99m (0.1 to 0.2 mCi) was injected intradermally 2 to 4 hours pre-operatively in 2 to 4 sites around the primary lesion, with or without up to 4 ml blue dye injected intradermally at the start of surgery. The node was considered 'hot' if radioactivity was 5 x background activity SND and RS were performed during the same operation 1 withdrawal described
Comparative	
Notes	At the beginning of the study, all women received both Tc-99m and blue dye, but later on blue dye was only used if the initial Tc-99m test failed to identify SNs. Thus, combined tests and Tc-99m only tests were used in 30 and 11 women, respectively. We were unable to extract separate 2 x 2 data for the combined and single techniques Participants were stratified by the proximity of the lesion to the midline (midline; ≤ 1 cm of midline; lateralised)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	Unclear
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Hauspy 2007 (Continued)

Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear

Unclear

Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Were other imaging tests performed prior to the index test to rule out groin lymph node metastases? Unclear

Unclear

Johann 2008
Study characteristics

Patient sampling Retrospective study; all patients who underwent SND and/or IFL between 1990 and 2007 were systematically reviewed. For the accuracy analysis, 23 women were included. The remainder were excluded as only SND or IFL had been performed (11)

Patient characteristics and setting 34 women with vulval SCC who had undergone SND and/or IFL. We extracted data for the group who had undergone both (23 women, 45 groins). It is unclear whether any women had clinically suspicious groin lymph nodes pre-operatively or whether T1 cases included T1a
 Median age: 68.4 years (34 to 87)
 Setting: university hospital in Switzerland from 1990 to 2007

Index tests Tc-99m and blue dye
 Histological methods not described in detail, except that 'step sectioning' was done

Target condition and reference standard(s) TC: groin lymph node involvement
 RS: complete IFL

Flow and timing 1 day pre-operatively, 4 x 15 MBq Tc-99m injected intradermally around the tumour. Blue dye was injected 2 to 10 minutes prior to skin incision
 Detected intra-operatively by gamma probe and visually (blue nodes)
 SND and RS were performed during the same operation

Comparative

Notes Unilateral IFL was performed in 1 woman; the rest underwent bilateral IFL

Johann 2008 (Continued)

Insufficient SN detection/accuracy data per groin available for 2 x 2 table. Authors emailed 19 June 2013

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	Unclear
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear		
		Unclear	

Klar 2011
Study characteristics
Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer (Review)

Klar 2011 (Continued)

Patient sampling	Prospective; consecutive enrolment; 4 exclusions
Patient characteristics and setting	<p>20 women with vulval SCC, FIGO stage IB; 29 groins; 3 unilateral and 13 bilateral LNDs</p> <p>Excluded women with no obvious nodes or women with other histology</p> <p>Median age: 66 years (36 to 88)</p> <p>Setting: tertiary institution in Germany; September 2007 to March 2010</p>
Index tests	<p>Tc-99m only</p> <p>Histological methods: ultrastaging with IHC staining was performed for SNs if standard section with H&E stain was negative; standard sections with H&E stains were performed for other nodes (RS)</p>
Target condition and reference standard(s)	<p>TC: groin lymph node involvement</p> <p>RS: complete IFL</p>
Flow and timing	<p>After GA, 0.3 ml of 10 MBq (4 x 2.5) Tc-99m was injected intradermally at 4 sites around the tumour. Detected immediately thereafter by gamma probe. Dissection was continued if more 'hot' nodes were identified (> 10% of the 'hottest' SN).</p> <p>SND and RS were performed during the same operation</p> <p>4 withdrawals described including 1 woman with vaginal melanoma and 3 women with obvious groin node metastases</p>
Comparative	
Notes	<p>This study evaluates the feasibility of a more convenient and less invasive way of administering the index test (Tc-99m only).</p> <p>Midline defined as within 1 cm of the midline</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	Low
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Yes		
		Low	Low

Klar 2011 (Continued)

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
Unclear	
Low	

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear
Low	

Klat 2009
Study characteristics

Patient sampling	Prospective study; sampling method unclear
Patient characteristics and setting	23 women with early-stage vulval SCC; 41 groins; clinically T1B/2; 5 had lateral and 18 had midline lesions Excluded women with clinically suspicious groin nodes Median age: 68.5 years (38 to 92) Setting: a tertiary institution in the Czech Republic; enrolment from June 2004 to November 2007
Index tests	Tc-99m and blue dye Histological methods: ultrastaging and staining with H&E and IHC were performed for SNs
Target condition and reference standard(s)	TC: groin lymph node involvement RS: complete IFL
Flow and timing	Tc-99m (50 MBq) was injected peritumourally on the day of surgery; LSG was performed 30 to 60 minutes later and SNs marked on the skin. After GA and groin skin incisions, patent blue dye was injected around the tumour. Radioactivity was detected intra-operatively by gamma probe and blue-stained nodes and draining lymphatics were identified visually. SND and RS were performed during the same operation

Klat 2009 (Continued)

No withdrawals described
 Followed up for 8 to 46 months

Comparative

Notes

1 woman had blue dye only due to failure of the gamma probe

All women (100%) with lesions > 4 cm had LN involvement at SN assessment and 75% of women with T2 lesions (≥ 2 cm)

The proportion of positive groins in this study is quite high (15/23) but may be a reflection on the use of ultrastaging and IHC for all SNs

Micrometastatic involvement was present in 5 groins (including the 1 false negative case)

The 1 false negative case occurred at the beginning of the study and was considered by the investigators to be due to a 'learning curve' (possible failure to remove a SN that had, in fact, been present on the pre-op LSG)

4 recurrences (1 death) occurred within the first 2 years of follow-up

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Unclear	Low
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			

Klat 2009 (Continued)

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear
Unclear	

Levenback 2001
Study characteristics

Patient sampling	Prospective study; sampling methods and withdrawals, if any, were not described
Patient characteristics and setting	<p>52 women with vulval cancer of any stage; histology included SCC (35), melanoma (7) and adenocarcinoma (9)</p> <p>4 women (8%) had clinically suspicious groin nodes</p> <p>Median age: 58 years (18 to 92)</p> <p>Setting: tertiary institution in USA from 1993 to 1999</p>
Index tests	<p>Blue dye (isosulfan) only</p> <p>Histological methods: ultrastaging and IHC were not consistently performed. SNs and other nodes were examined in standard sections with H&E staining. Frozen section was performed for macroscopically suspicious nodes</p>
Target condition and reference standard(s)	<p>TC: groin lymph node involvement</p> <p>RS: superficial LND or complete IFL (at the discretion of the surgeon)</p>
Flow and timing	<p>Blue dye (1 to 4 ml) was injected at the leading edge of the tumour closest to the groin after GA. In women with midline tumours, injections were done bilaterally. If after 10 minutes no lymph channel was seen, further LND was performed.</p> <p>SND and RS were performed during the same operation</p> <p>No withdrawals were described</p>
Comparative	
Notes	<p>Includes data from Levenback 1995 (first 21 women)</p> <p>Lesions were considered to be midline if within 2 cm of the midline. For midline tumours, bilateral LND was performed.</p> <p>Of 27 women with midline lesions, 24 had bilateral LND, the others had unilateral LND. The extent of LND was at the attending surgeon's discretion and the surgeon "could choose to abort a case in the event of positive nodes"</p>

Levenback 2001 (Continued)

Superficial inguinal LND was performed in 42 women (81%), complete IFL in 6 women (12%) and no RS in 4 women (8%).

Detection failures were more frequent earlier on in the study. In a subset of 16 women with SCC stage T1/2 and clinically non-suspicious nodes, there were no detection failures (women or groins). Investigators concluded that conditions that disrupt lymphatic drainage, such as prior surgery, excisional biopsy or infection, compromise SN detection

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Unclear	High
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		

Levenback 2001 *(Continued)*

Were other imaging tests performed prior to the index test to rule out groin lymph node metastases? Unclear

High

Levenback 2012
Study characteristics

Patient sampling	Prospective study; 551 enrolled; 56 were excluded following GOG centralised reviews and 7 women did not undergo the SN procedure
Patient characteristics and setting	<p>452 women with vulval SCC at least 1 mm deep, and with tumour size 2 to 6 cm (772 groins); 132 lateralised, 320 midline/near midline lesions</p> <p>11% had undergone prior wide local excision of the primary tumour</p> <p>Women were excluded if groin involvement was suspected, or if they had undergone prior groin surgery or irradiation, had multifocal disease or a grossly inflamed tumour</p> <p>Median age: NR</p> <p>Setting: 47 centres in the USA</p>
Index tests	<p>SN detection by blue dye with or without Tc-99m</p> <p>Histological methods: ultrastaging with H&E staining were performed on all SNs. If these were negative, IHC staining with cytokeratin was performed</p> <p>H&E stains were used for all other nodes in the reference standard</p>
Target condition and reference standard(s)	<p>TC: groin lymph node involvement</p> <p>RS: complete IFL</p>
Flow and timing	<p>Intradermal injection of blue dye was given 5 minutes before groin incision. If a blue channel led to a lymph node, it was considered sentinel, whether or not the node was blue. If Tc-99m was used, the node was considered 'hot' if the radioactivity was 10 x greater than background activity</p> <p>SND and RS were performed during the same operation</p> <p>Withdrawals described</p>
Comparative	
Notes	<p>Registered ID: NCT00003325</p> <p>We were unable to extract separate 2 x 2 data for the combined and single techniques. Additional data were requested 28 May 2013</p> <p>The type of procedure (bilateral versus unilateral) was determined by the proximity of the lesion to the midline: if ≤ 2 cm of midline, bilateral IFL was performed</p> <p>Investigators analysed data by tumour size (< 4 cm and ≥ 4 cm): NPV was lower for the larger tumour size subgroup (98% versus 93%)</p>

Levenback 2012 (Continued)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Unclear	Low
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear		
		Unclear	

Li 2009
Study characteristics

Patient sampling	Prospective study
Patient characteristics and setting	21 women with vulval cancer; unclear whether women with clinically suspicious nodes were excluded Median age: not reported Setting: a tertiary hospital in China; recruitment from October 2004 to April 2008
Index tests	Blue dye only (11 women) Technetium and blue dye (10 women) Histological methods unclear
Target condition and reference standard(s)	TC: groin lymph node involvement RS: complete IFL
Flow and timing	Insufficient information available
Comparative	
Notes	We were unable to obtain the full article but were able to construct 2 x 2 tables from the abstract

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
		Unclear	Unclear
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		

Li 2009 (Continued)

	Unclear	Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Unclear	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Unclear	
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear	
Unclear		

Lindell 2010

Study characteristics	
Patient sampling	Retrospective study; all patients with T1-3 vulval cancer, without palpable lymph nodes, were included
Patient characteristics and setting	77 women with vulval SCC T1-T3, 130 groins; 55 midline lesions and 22 lateral lesions Women with palpable groin nodes were excluded Mean age: 71.2 years (40 to 92) Setting: university hospital in Sweden from 2000 to 2007
Index tests	Tc-99m and blue dye (60 women); blue dye only (17 women) Histological methods: ultrastaging done; if H&E was negative, IHC stain was used
Target condition and reference standard(s)	TC: groin lymph node involvement RS: complete IFL
Flow and timing	LSG performed on morning of surgery. 40 MBq Tc-99m injected at 4 sites around the tumour. Sites of hot nodes were marked on the skin and surgery followed 2 to 5 hours after the injection. 15 minutes prior to skin incision, blue dye was injected at the same 4 sites Detected intra-operatively by gamma probe and visually (blue nodes) where applicable SND and IFL were performed during the same operation
Comparative	
Notes	The 2 false negatives occurred in women with midline lesions ≥ 4 cm; 1 had a multifocal lesion In the 32 undetected groins, 8 were positive for metastases

Lindell 2010 (Continued)

All 77 procedures were performed by the same 2 surgeons

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	Low
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear		
		Unclear	

Louis-Sylvestre 2006
Study characteristics

Louis-Sylvestre 2006 (Continued)

Patient sampling	Prospective study; consecutive sampling
Patient characteristics and setting	<p>38 women with operable vulval cancer; 64 groins; depth of invasion > 1 mm; histology not specified; 12 lateralised and 26 median lesions</p> <p>Excluded if prior excisional biopsy. Included 6 women with clinically suspicious nodes (N1)</p> <p>Median age: 66 years (34 to 90)</p> <p>Setting: a tertiary institution in France; enrolled between June 2002 and December 2005</p>
Index tests	<p>Tc-99m and blue dye</p> <p>Blue dye was optional and not used in 8 women</p> <p>Histological methods: ultrastaging was not done. IHC staining was performed on SNs if H&E stain was negative</p>
Target condition and reference standard(s)	<p>TC: groin lymph node involvement</p> <p>RS: complete IFL</p>
Flow and timing	<p>1 day pre-operatively, Tc-99m was injected in 3 x intradermal peritumoural injections (30 MBq) and immediately followed by LSG. SN site/s were located with hand-held gamma probe and marked on the skin. Blue dye was injected in the same way intra-operatively</p> <p>SND and RS were performed during the same operation</p> <p>Withdrawals not described</p>
Comparative	
Notes	<p>Considered midline if within 1 cm of the midline</p> <p>This study incorporates the data for 17 women from an earlier publication (Louis-Sylvestre 2005)</p> <p>Due to the high risk of bias from the inclusion of 6 women with obvious groin involvement, we excluded these women in our data extraction. This resulted in a groin detection rate of 39/52 (compared with the published rate of 47/64) and no false negatives. (The only false negative in this study occurred in a woman with obvious bilateral groin node involvement)</p>

Methodological quality

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

Louis-Sylvestre 2006 (Continued)

DOMAIN 2: Index Test Test group

Had the test operator performed 10 or more procedures?	Unclear
--	---------

Unclear

Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	Yes
---	-----

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
--	---------

Unclear

Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes
--	-----

Did all patients receive the same reference standard?	Yes
---	-----

Were all patients included in the analysis?	Unclear
---	---------

Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear
---	---------

Unclear

Martinez-Palones 2006
Study characteristics

Patient sampling	Prospective; consecutive enrolment
------------------	------------------------------------

Patient characteristics and setting	Women with early vulval cancer, stage I/II; mostly SCC (2 women with melanoma); depth of invasion > 1 mm; excluded previously treated vulval cancer Excluded if clinically suspicious nodes. 16 lateralised and 12 median lesions Mean age: 71 (30 to 84) Setting: tertiary institution in Spain; recruited between January 2002 and July 2005
-------------------------------------	---

Index tests	Tc-99m and blue dye Histological methods: standard sections with H&E were done for SNs and repeated with IHC stains if negative. Methods for non-SN nodes unclear
-------------	--

Target condition and reference standard(s)	TC: groin lymph node involvement
--	----------------------------------

Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer (Review)

Martinez-Palones 2006 (Continued)

RS: complete IFL

Flow and timing

1 day pre-operatively, Tc-99m was injected as 4 x intradermal peritumoural injections and immediately followed by LSG. Blue dye (2 to 4 ml) was injected in the same way intra-operatively. Hot SNs detected intra-operatively by gamma probe and dissection was continued if more 'hot' nodes were identified (> 10 x background levels). Blue nodes and draining lymphatics identified visually

SND and RS were performed during the same operation
 Withdrawals not described

Comparative

Notes

This study also compared these 28 women with SN assessment to 27 women (retrospective, series) without SN assessment

'Midline' not defined

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	Low
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Martinez-Palones 2006 *(Continued)*

Were all patients included in the analysis?	Yes
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear
Unclear	

Merisio 2005
Study characteristics

Patient sampling	Prospective study. Sampling not described
Patient characteristics and setting	<p>20 women with vulval SCC, clinically T1/2; 32 groins. Unclear whether any women with stage IA were included</p> <p>Excluded if clinically suspicious groin nodes, T3-4 lesions or prior chemo/radiotherapy</p> <p>9 lateral and 11 had midline lesions; 2 excisional biopsies</p> <p>Mean age: 75 (49 to 92)</p> <p>Setting: 2 tertiary institutions in Italy; enrolment from May 1999 to May 2003</p>
Index tests	<p>Tc-99m only</p> <p>Histological methods: ultrastaging performed. IHC staining done for SNs if H&E stains were negative</p>
Target condition and reference standard(s)	<p>TC: groin lymph node involvement</p> <p>RS: complete IFL (8 had unilateral and 12 had bilateral LND)</p>
Flow and timing	<p>16 hours pre-operatively, Tc-99m (10 to 20 MBq) was injected as 4 x intradermal peritumoural injections and immediately followed by LSG</p> <p>Early and late scans done and the SN/s location marked on skin. Detected intra-operatively by hand-held gamma probe</p> <p>SND and RS were performed during the same operation</p> <p>Withdrawals, if any, were not described</p>
Comparative	
Notes	Lesions were considered midline if within 2 cm of the midline. 1 woman with a lateral lesion had bilateral IFL

Methodological quality

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

Merisio 2005 *(Continued)*

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Unclear	Unclear

DOMAIN 2: Index Test Test group

Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear		
		Unclear	

Moore 2003a
Study characteristics

Patient sampling	Prospective study; sampling methods not described. Appear to have recruited 23 women and excluded 2 due to palpable groin nodes at surgery
Patient characteristics and setting	<p>21 women with biopsy-proven vulval SCC > 1 mm deep; 31 groins. No patients had undergone incisional biopsy before enrolment</p> <p>Excluded women with clinically suspicious nodes. Location and other details of the tumours was not reported</p> <p>Median age: NR</p> <p>Setting: 2 tertiary care hospitals in USA; recruitment dates not stated</p>

Moore 2003a (Continued)

Index tests	Tc-99m and blue dye Ultrastaging with H&E staining was performed for SNs. Standard sections for other non-SN nodes
Target condition and reference standard(s)	TC: groin lymph node involvement RS: complete IFL
Flow and timing	A total of 2 mCi Tc-99m in 1 ml was injected at 2 sites around the tumour 90 to 180 minutes prior to surgery. LSG was performed. Intra-operatively, 3 ml blue dye was injected at the same 2 sites. Detected intra-operatively by gamma probe and visually (blue nodes). LNs with increased activity (> 5% of the injection site) were also removed. Nodes with a count of at least 10% of the hottest SN were defined as 'hot' SND and IFL were performed during the same operation
Comparative	
Notes	10 bilateral IFL and 11 unilateral procedures were performed 3 women were found to have unilateral, palpable groin nodes, therefore the investigators excluded these 3 groins from the SN assessment. Tc-99m detected SNs in all 31 groins but blue dye only detected 19/31 groins (61%) and 3/9 metastatic groins Metastases were also present in the non-sentinel nodes of 5/9 SN positive groins

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
		Unclear	Low
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		

Moore 2003a (Continued)

Unclear
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear

Unclear
Morotti 2011
Study characteristics

Patient sampling	Prospective study; consecutive recruitment. 5 women were excluded from study as they did not consent to the procedure (1), did not undergo SN mapping (2) or did not undergo surgery at the study hospital (2)
Patient characteristics and setting	56 women with vulval cancer FIGO stage Ib/II Mean age: 73.6 years (54 to 92) Setting: university hospital in Italy; recruitment from February 2007 to February 2011
Index tests	Tc-99m and blue dye Histological methods: ultrastaging and IHC stains performed if H&E stain was negative
Target condition and reference standard(s)	TC: groin lymph node involvement RS: complete IFL
Flow and timing	24 hours pre-operatively, Tc-99m (40 MBq) was injected as 4 x intradermal peritumoural injections and followed 30 to 60 minutes later by LSG. The hottest SN was marked on the skin. At surgery, 2 ml blue dye was injected intradermally at the same 4 sites. Tc-99m was detected intra-operatively using a hand-held gamma probe. SND was followed by IFL during the same operation
Comparative	
Notes	We obtained this unpublished paper from the investigators. Additional per groin data were requested and obtained by email

Methodological quality

Morotti 2011 (Continued)

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	Low
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Yes		
		Low	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Yes		
		Low	

Nyberg 2007
Study characteristics

Patient sampling	Retrospective study; consecutive cases
Patient characteristics and setting	Included 47 women with vulval cancer (46 with SCC) who underwent SN mapping and IFL. Stage I/II (25 women) comprised only 55%; we extracted these data separately. Median lesions (11), unilateral lesions (9), bilateral lesions (5)*

Nyberg 2007 (Continued)

	Median age: not reported separately for stage I/II women Setting: university hospital in Finland
Index tests	Tc-99m and blue dye (20 women) Blue dye only (5 women) Histological methods not described
Target condition and reference standard(s)	TC: groin lymph node involvement RS: complete IFL
Flow and timing	Tc-99m was injected either pre-operatively or the day before and detected via hand-held gamma probe intra-operatively. No LSG was done pre-operatively Blue dye was injected perilesionally and the SN was the first node traced by the afferent lymph channel irrespective of whether the node itself was blue SND and IFL were performed during the same operation
Comparative	
Notes	Investigators state that "To familiarise with the (SN) method, the procedure was performed in most of the patients who underwent lymphadenectomy regardless of the clinical staging". 'Per patient' data only were reported for the 25 women with stage I/II. 'Per groin' data were obtained for these women from the authors via email. 'Midline' was not defined. *Unpublished information/groin data obtained from the authors

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
		Unclear	High
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Nyberg 2007 (Continued)

Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear

Unclear

Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Were other imaging tests performed prior to the index test to rule out groin lymph node metastases? Unclear

Unclear

Radziszewski 2010
Study characteristics

Patient sampling Prospective study; consecutive sampling; 6 withdrawals explained

Patient characteristics and setting 62 women with clinical stage I/II vulvar SCC \leq 4 cm; depth > 1 mm; histologically confirmed; no obvious nodal involvement; WHO performance status I-II

Median age: 68 years (37 to 94)

Setting: tertiary hospital in Poland; January 2002 to December 2006

Index tests Tc-99m and blue dye were used (results were compared and reported separately)

Histological methods: standard sections with H&E and IHC stains were performed for SN and non-SN assessment. In addition, all specimens were subjected to a RT-PCR test for CA9 marker expression

Target condition and reference standard(s) TC: groin lymph node involvement

RS: complete IFL

Flow and timing 24 hours pre-operatively, Tc-99m (1.2 mCi) was injected as 3 intradermal perilesional injections and followed 1 hour later by LSG. A 2.5% solution of patent blue was injected in a similar way, 10 minutes before the skin incision

Radioactivity was detected intra-operatively by hand-held gamma probe. SND and IFL were performed during the same operation

Comparative

Notes Detection rates were reported separately for blue dye and Tc-99m

Radziszewski 2010 (Continued)

6 women were excluded as they did not fulfil the inclusion criteria (e.g. lymph node metastasis on FNA)

SNs and non-SNs were also assessed for CA9 marker expression by RT-PCR. This method detected twice as many positive SNs (and 1 additional false negative) as the H&E/IHC method, however the clinical significance of these micrometastases is not known

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	Low
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	No		
		High	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Yes		
		Unclear	

Rob 2007

Study characteristics

Patient sampling	Prospective study; consecutive sampling (unpublished information). 9 women with 'bulky nodes' were excluded (unpublished information)
Patient characteristics and setting	59 women with vulval SCC stage IB/II; 86 groins Excluded lesions > 4 cm and women with clinically suspicious nodes Age range: 26 to 95 years Setting: a university hospital in the Czech Republic; recruited from December 2001 to December 2005
Index tests	Blue dye only (16 women) and combined technique (blue dye and Tc-99m; 43 women) Frozen section and ultrastaging with H&E and IHC were performed for SNs. Standard methods were used for non-SNs
Target condition and reference standard(s)	TC: groin lymph node involvement RS: complete IFL
Flow and timing	Tracer agents were injected intradermally around tumour at 3 sites. Tc-99m was injected 3 to 4 hours pre-operatively in a volume of 0.2 to 0.4 ml. Blue dye (1 to 2 ml) was injected 2 to 3 minutes pre-operatively. SN location was marked on skin. Tc-99m was detected intra-operatively by a gamma probe. First SN was identified and nodes exhibiting at least 10% of activity at the area of application were also removed. Blue stained lymphatics were identified visually SND and IFL were performed during the same operation
Comparative	
Notes	Investigators reported the blue dye and combined technique accuracy results separately. Additional methodological information and groin data (for the combined technique) were supplied to us by the investigators via email (4 July 2013). 16% of identified SNs were deep femoral groin nodes Frozen section was 98% accurate in this study and influenced the extent of the IFL The optimum timing of pre-operative LSG was 45 minutes after the Tc-99m injection

Methodological quality

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

Rob 2007 (Continued)

		Low	Low
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Yes		
		Unclear	

Sawicki 2010

Study characteristics	
Patient sampling	Prospective study; consecutive sampling*; no withdrawals*
Patient characteristics and setting	24 women (39 groins) with confirmed vulval SCC \geq stage IB; diagnosis by biopsy Excluded women with clinically suspicious nodes Mean age: 66.2 years (37 to 86) Setting: university hospital in Poland; recruited between 2003 and 2010
Index tests	Blue dye only (10 women) Tc-99m and blue dye (14 women)

Sawicki 2010 (Continued)

Histological methods not described

Target condition and reference standard(s)	TC: groin lymph node involvement RS: complete IFL
Flow and timing	Tc-99m (1 to 2 mCi) was injected 15 to 120 minutes pre-operatively in a volume of 1 to 2 ml. Blue dye (1 to 2 ml) was injected after GA 15 to 20 minutes pre-operatively. Both 'injected intradermally around tumour'. Blue dye without Tc-99m was used alone in 14 women Radioactivity detected intra-operatively by gamma probe at least 5 x background activity; blue nodes and lymphatic channels detected visually SND and IFL were performed during the same operation
Comparative	
Notes	*Unpublished information obtained via e-mail correspondence with the authors Accuracy data 'per groin' were not clearly reported, however these were also obtained via e-mail Investigators report that detection failures (5/39) occurred in the first half of the study, suggesting a 'learning curve'

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	Low
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	No		
		High	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low

Sawicki 2010 (Continued)

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear
Unclear	

Sideri 2000
Study characteristics

Patient sampling	Prospective study; consecutive enrolment
Patient characteristics and setting	<p>44 women with vulval SCC; 77 groins; 20 T1 and 23 T2 lesions; 1 vaginal SCC. It is unclear whether all T1 lesions were T1b.</p> <p>No clinically suspicious groin nodes; 11 had lateral and 33 had midline lesions</p> <p>Median age: NR</p> <p>Setting: tertiary institution in Italy; enrolled from May 1996 to July 1999</p>
Index tests	<p>SN detection by Tc-99m only</p> <p>Histological methods: standard sections stained with H&E</p>
Target condition and reference standard(s)	<p>TC: groin lymph node involvement</p> <p>RS: complete IFL</p>
Flow and timing	<p>1 day pre-operatively, Tc-99m (14 to 30 MBq) was injected as 2 (for small unilateral tumours) or 4 intradermal peritumoural injections and followed by LSG. LSG was also performed 3 hours later and SNs were marked on the skin. 'Hot' nodes were detected intra-operatively by gamma probe</p> <p>SND and RS were performed during the same operation</p> <p>Withdrawals were not described</p>
Comparative	
Notes	<p>Considered midline if the lesion was within 2 cm of the midline</p> <p>This is an extension of De Cicco 2000, which reported on SN assessment in 37 women (55 groins), however there are discrepancies between these reports that suggest that some women from the earlier report were excluded in the latter report. By our calculations, 8 women with unilateral tumours (8 groins dissected) may have been excluded from the original paper and 15 women with bilateral lesions (30 groins) added to the extended</p>

Sideri 2000 (Continued)

data set (Sideri 2000). We contacted the authors, however they no longer had access to these data. We therefore considered these data to be at a potentially high risk of bias

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	Low
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear		
		High	

Sliutz 2002
Study characteristics

Patient sampling	Prospective study; consecutive enrolment
Patient characteristics and setting	<p>26 women with T1/2 vulval cancer (24 with SCC, 2 other); groins not described (?46). 6 lateralised lesions, 20 midline lesions</p> <p>Women with clinically suspicious nodes, prior surgery/chemotherapy or radiotherapy were excluded</p> <p>Median age: 62.5 years (40 to 86)</p> <p>Setting: university hospital in Germany; recruited from May 1998 to November 2000</p>
Index tests	<p>Tc-99m with (8 women) or without blue dye (18 women)</p> <p>Histological methods: standard sections and H&E stains were used. If SNs were negative, ultrastaging with additional IHC stains was performed</p>
Target condition and reference standard(s)	<p>TC: groin lymph node involvement</p> <p>RS: complete IFL</p>
Flow and timing	<p>2 to 3 hours pre-operatively, Tc-99m (14 MBq in 0.4 ml saline) was injected intradermally at 4 sites perilesionally. Dynamic (immediately) and static images (2 hours later) were done and SNs were marked on the skin. The first 8 women also received blue dye (1 ml), injected perilesionally</p> <p>Radioactivity was detected intra-operatively by gamma probe. Only 3/8 women had potential blue nodes identified by afferent blue channels</p> <p>SND was followed by complete IFL in the same operation</p>
Comparative	
Notes	<p>Midline not defined. 'Per groin' data were not reported. Authors e-mailed on 2 July 2013 (unsuccessful)</p> <p>IHC/ultrastaging identified 1 additional micrometastasis</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	Low
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		

Sliutz 2002 (Continued)

	Unclear	Low
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	
	Unclear	Low

DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear	
	Unclear	

Trifiro 2010

Study characteristics	
Patient sampling	Retrospective study; consecutive enrolment
Patient characteristics and setting	12 women with vulval melanoma without clinically suspicious nodes or previous surgery of the primary tumour Midline/lateralised lesions not reported Median age: 64 years (29 to 79) Setting: tertiary institution in Italy; treated from April 1997 to May 2003
Index tests	Tc-99m only Histological methods: standard sections with H&E staining
Target condition and reference standard(s)	TC: groin lymph node involvement RS: complete IFL (only 10/12 women underwent this procedure)
Flow and timing	24 hours pre-operatively, Tc-99m (20 MBq) was injected as 2 to 4 x intradermal peritumoural injections and immediately followed by LSG. LSG was also performed 3 hours later and SNs were marked on the skin. Radioactivity was detected intra-operatively by gamma probe

Trifiro 2010 (Continued)

SND and IFL were performed during the same procedure in 10/12 cases. 2 women underwent SND only

Comparative

Notes

Midline was not defined. The number of bilateral and unilateral LNDs was not reported

'Per groin' data were requested via e-mail on 1 July 2013

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	Low
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear		
		High	

Vidal-Sicart 2007
Study characteristics

Patient sampling	Prospective study; enrolment appears to be consecutive but this is not clear. Comprised a validation study (50 women) and an applications study (20 women)
Patient characteristics and setting	<p>50 women with vulval SCC, including 7 women who had stage III lesions clinically. We excluded the data from these 7 women, resulting in 43 assessable women (64 groins)</p> <p>Mean age: not reported</p> <p>Setting: 3 tertiary hospitals in Spain; recruitment from June 1998 to July 2005</p>
Index tests	<p>Tc-99m and blue dye</p> <p>Histological methods: standard section with H&E staining was used for SNs; if negative, ultra-staging with IHC stains was performed</p>
Target condition and reference standard(s)	<p>TC: groin lymph node involvement</p> <p>RS: complete IFL</p>
Flow and timing	<p>1 day pre-operatively, 2 to 4 x 37 MBq Tc-99m (0.1 ml) was injected intradermally around the tumour. Dynamic and planar imaging was done (30 minutes and 2 hours later). SN locations were marked on the skin. Delayed imaging was obtained where necessary. Blue dye (1 ml) was injected 2 to 10 minutes prior to skin incision</p> <p>Detected intra-operatively by gamma probe and visually (blue nodes). SNs with 10 x background activity were removed. After removal, the field was scanned to check for 'other significant activity'</p> <p>SND and IFL were performed during the same operation</p>
Comparative	
Notes	<p>As the individual patient data were reported, we were able to exclude the data for 7 women with stage III at enrolment from our analyses</p> <p>Midline was not defined and the numbers of women with midline/lateralised lesions was not reported. Instead, the number of women considered to have bilateral or unilateral lymphatic drainage was reported</p> <p>We extracted groin SN assessment data from the individual patient data presented in Table 1 of the primary article</p> <p>Failed SN detection per groin was not reported, however, half (2/4) of the undetected nodes were node-positive in the opposite groin</p>

Methodological quality

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

Vidal-Sicart 2007 *(Continued)*

Did the study avoid inappropriate exclusions? Unclear

Unclear

Unclear

DOMAIN 2: Index Test Test group

Had the test operator performed 10 or more procedures? Unclear

Unclear

Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear

Unclear

Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Were other imaging tests performed prior to the index test to rule out groin lymph node metastases? Unclear

Unclear

Zambo 2002
Study characteristics

Patient sampling	Prospective pilot study; sampling methods and withdrawals (if any) were not described
Patient characteristics and setting	10 women with vulval cancer (8 SCC, 2 melanomas); 20 groins dissected; tumours included 8 lateralised lesions, 1 midline lesion and 1 bilateral lesion Excluded women with palpable lymph nodes

Zambo 2002 (Continued)

Median age: 59.5 years (32 to 77)

Setting: tertiary hospital in Hungary; April 1999 to March 2002

Index tests	Tc-99m and blue dye Standard histological methods with H&E stains were used
Target condition and reference standard(s)	TC: groin lymph node involvement RS: complete IFL
Flow and timing	24 hours pre-operatively, Tc-99m (100 MBq in 1 ml) was injected as 4 intradermal peritumoural injections and followed by LSG 1, 3 and 24 hours later. SN locations were marked on the skin. Patent blue was injected in a similar way during GA SND and bilateral IFL were performed for all women
Comparative	
Notes	Investigators concluded that SN detection by LSG "is highly successful in the management of patients with vulval cancer"

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Unclear	Low
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			

Zambo 2002 (Continued)

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear
Unclear	

Zekan 2012
Study characteristics

Patient sampling	Prospective study; consecutive enrolment; withdrawals unclear
Patient characteristics and setting	<p>25 women with vulval SCC (50 groins); clinical stage I and II; depth of invasion > 1 mm</p> <p>14 women had midline lesions and 11 had lateralised lesions</p> <p>Excluded women with clinically suspicious groin nodes, previously treated vulval SCC</p> <p>Median age: 69 years (48 to 79)</p> <p>Setting: university hospital in Croatia; December 2007 to May 2011</p>
Index tests	<p>Tc-99m only</p> <p>Histological methods: ultrastaging not described. Standard sections with H&E staining were performed on all SNs. If these were negative, IHC staining with cytokeratin was performed. H&E stains were used for all other RS nodes; if IHC staining of SNs was negative, IHC staining was done on other RS nodes</p>
Target condition and reference standard(s)	<p>TC: groin lymph node involvement</p> <p>RS: complete IFL</p>
Flow and timing	<p>3 hours pre-operatively, Tc-99m was injected peritumourally at 4 sites (15 to 20 MBq) and LSG was done. Detected intra-operatively by gamma probe. Dissection was continued if more 'hot' nodes were identified (> 10% of the 'hottest' SN).</p> <p>SND and RS were performed during the same operation</p> <p>Withdrawals were not described</p>
Comparative	
Notes	<p>In a preliminary report (Corousic 2011 conference abstract) 29 women underwent SN assessment. Zekan 2012 reports the results of only 25 women. The authors confirmed via e-mail that these data are from the same series, however we were unable to ascertain the reasons for the 4 exclusions, therefore we considered the potential risk of bias to be high for patient flow.</p> <p>The same surgeon performed all procedures although the first 10 patients included represent the 'learning curve'. Investigators reported the results separately for the first 10 and subsequent procedures</p>

Zekan 2012 (Continued)

Tumours were classified as 'midline' if within 2 cm of the midline. All women had bilateral IFL regardless of proximity to midline

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
		Unclear	Low
DOMAIN 2: Index Test Test group			
Had the test operator performed 10 or more procedures?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Were other imaging tests performed prior to the index test to rule out groin lymph node metastases?	Unclear		
		High	

Abbreviations: FIGO = International Federation of Gynecology and Obstetrics; FNA = fine needle aspiration; H&E = haematoxylin and eosin; IFL = inguino-femoral lymphadenectomy; IHC = immunohistochemical; LN = lymph node; LND = lymph node dissection; LSG = lymphoscintigraphy; NPV = negative predictive value; NR = not reported; RS = reference standard; RT-PCR = reverse transcriptase polymerase chain reaction; Tc-99m = radiolabelled technetium

Characteristics of excluded studies [ordered by study ID]

Study	Reason for exclusion
Achimas-Cadariu 2009	A retrospective case series of 59 women with vulval cancer, of whom 28 women underwent SND only and 31 women underwent IFL with or without SND (precise numbers not specified). 2 x 2 tables could not be constructed
Ansink 2003	Letter in response to Puig-Tintore 2003
Armstrong 2011	Construction of a 2 x 2 table was not possible from this abstract of a retrospective study of 15 SNDs
Baiocchi 2011	Not a study of SN assessment. Retrospective cohort study of groin metastasis in 146 women with vulval SCC who underwent bilateral full IFL
Barton 1992	A small pilot study of the SN procedure in women with vulval cancer. Most cases (6/10) had clinically suspicious groin lymph nodes and it was not possible to construct a 2 x 2 table with the available data
Beneder 2008	Pilot study of fusion imaging (SPECT/CT/MRI) for detection of SNs in 10 women with vulval cancer, not a study of SN assessment with full IFL
Bibi 2011	Conference abstract; 21 women with vulval SCC underwent SN detection with Tc-99m and blue dye. It was not possible to construct a 2 x 2 table with the available data
Brunner 2008	An Austrian study evaluating the accuracy of frozen section for SN assessment in women who did not undergo IFL
Bus 2011	Conference abstract; 28 women with vulval cancer underwent SN assessment with Tc-99m only and IFL. Groin data not reported. Insufficient detail to construct a 2 x 2 table (per woman or per groin)
Cady 2000	Editorial
Carcopino 2005	A study in which 15 women with vulval SCC underwent SN detection with Tc-99m and blue dye, but did not undergo complete IFL. If node was positive, the women received inguinal irradiation
Choudhary 2003	A letter to the editor with reference to Zambo 2002
Crane 2011	Feasibility study of NIRF technique for detecting SNs in 10 women with early vulval cancer, compared with Tc-99m and blue dye techniques. Full IFL was not performed as the reference standard. 2 x 2 tables could not be constructed
de Hullu 2001	Case report
de Hullu 2002	A retrospective cohort study. 9 women with vulval melanoma underwent the SN procedure but only 6 women underwent IFL (reference standard). Number of groins was not reported. 2 x 2 tables could not be constructed
Devaja 2011	Prospective study of SN in women with vulval cancer. Construction of a 2 x 2 table was not possible. IFL was only performed in 41/60 participants. Included women with clinically suspicious nodes
Dhar 2005	A review

Study	Reason for exclusion
Ennik 2011	Retrospective study of 64 women who underwent an SN procedure using mixed methods over a 10-year period to evaluate the effects of previous surgery on detection rates (27/64 women had had previous vulval surgery). Most women did not undergo IFL (RS), The sample included women with clinically suspicious nodes "to investigate whether firm nodal metastases can cause SNs to be by-passed by lymph flow"
Farrell 2010	Not a study of SN assessment. Retrospective study of SN lymphoscintigraphy in 13 women with vulval cancer to determine time to SN localisation (SAT)
Freudenberg 2010	A case report of lymphatic mapping using SPECT/CT in vulval cancer
Garcia-Iglesias 2012	A retrospective study of 76 women with vulval SCC who underwent SND. IFL was selectively performed in women with metastases-positive SNs. It was not possible to construct 2 x 2 tables
GROINSS-V 2008	A study evaluating the safety of SND alone in 403 women with vulval SCC T1/2 (< 4 cm). The relationship between size of SN metastasis and chances of non-SN involvement in was also evaluated. IFL was only performed for women in whom SNs were positive (115). 3-year survival rate was 97% (95% CI 91% to 99%). Morbidity with SND versus IFL was compared
Hefler 2008	A retrospective comparative study of SN dissection versus complete IFL in women with vulval cancer. Investigators reported a significant reduction in postoperative morbidity with SN dissection only
Kraft 2012	A conference abstract of a study comparing planar scintigraphy and SPECT/CT imaging techniques in various types of tumours including 7 vulval cancers
Levenback 2000	Letter to the editor regarding Terada 2000
Makowski 2010	Conference abstract: this study of 16 women with vulval cancer (including stage IA) underwent SN assessment with SPECT/CT and planar LSG pre-operatively. Complete IFL was performed in half of the women only. 5 women had LN metastases, including 1 false negative result. Pre-operative SN detection was 100% for SPECT/CT compared with 94% for LSG
Maza 2007	A study evaluating a novel multimodal fusion imaging approach using a mixture of Tc-99m and supramagnetic iron oxide injected perilesionally in 14 patients with different tumour entities, including 1 women with vulval SCC
Molpus 2001	A study of sentinel node detection and microstaging in 11 women with vulval cancer. Only 7 women underwent SND plus IFL; 2 had clinical suspicious groin nodes
Moore 2003b	A prospective study comparing IHC staining versus ultrastaging with H&E in women who underwent SND (completion IFL not reported)
Moore 2008	A prospective study evaluating the recurrence rates in 31 women with vulval cancer and SN negative nodes, who underwent conservative management (no IFL). Groin recurrence rate was 4.3% per groin and 6.4% per woman
Nickles Fader 2012	A retrospective multi-site study of 45 women with vulval melanoma and clinically negative nodes who underwent SND. Only 11/45 of the cohort underwent completion IFL
Penson 2001	A letter to the editor in response to de Hullu 2000
Puig Calvo 2011	A retrospective study of 30 women who underwent sentinel node biopsy. IFL was not routinely performed; 7/11 women with positive sentinel nodes did not undergo IFL

Study	Reason for exclusion
Radziszewski 2003	A prospective pilot study of women with vulval SCC stage T1/2 N0-2 who underwent SND and complete IFL. The study included an unspecified number of women with obvious nodal metastases
Robison 2006	A retrospective study of standard IFL compared with SND comparing the size of the metastases detected in inguinal nodes. SND with ultrastaging enabled the detection of smaller nodal metastases compared with standard examination
Rodier 1999	A prospective study of 8 women with vulval cancer who underwent SND only (no RS)
Soliman 2012	A retrospective study of the complications associated with IFL in women with vulval cancer
Sun 2009	A study of 3 women with vulval melanoma to investigate the feasibility of SND
Tavares 2001	A prospective study of 100 SN procedures in people with cancer, including 18 women with vulval cancer. SNs were detected in 15/18 women. 2 of the 15 women (13%) had clinically suspicious nodes and 2 x 2 tables could not be constructed
Tenney 2011	A retrospective study of 50 women with stage II vulval cancer who underwent SN biopsy and IFL. The aim was to evaluate the risk of positive non-SNs when positive SNs are identified. This conference abstract contains insufficient data to construct a 2 x 2 table for the purpose of this review. 12/50 women had positive SNs; 3/12 of these had positive non-SNs. False negatives were not described. Women with positive (12) and negative (40) SNs do not add up to 50 as reported
Terada 2000	A prospective study of 9 women with stage T1 vulval SCC (12 groins) who underwent SND without completion IFL in the first instance. IFL was performed only for 2 women with metastases-positive SNs
Terada 2006	A retrospective study of 21 women with T1 vulval SCC who underwent SND only. 3 women had positive SNs and received a full IFL. Women with negative nodes received no further treatment. 3-year survival for all women was 90%, and for women with negative nodes was 100%
Tjin Asjoe 2008	A histomorphologic review of 32 cases of vulval SCC to determine the significance of anucleate squamous cells on IHC
Vakselj 2007	A prospective study of SND alone in 35 women with vulval cancer
Van Den Eynden 2003	A retrospective study of 32 women with vulval cancer, including 8 women with clinically suspicious nodes (25%). IFL was not performed for all women
van der Velden 2006	Letter to the editor, not a study
Wechter 2004	A review of 20 patients with vulval melanoma and a review of the literature

CT = computed tomography; H&E = haematoxylin and eosin; IFL = inguinofemoral lymphadenectomy; LSG = haematoxylin and eosin; MRI = magnetic resonance imaging; RS = reference standard; SAT = scintigraphic appearance time; SCC = squamous cell carcinoma; SN = sentinel node; SND = sentinel node dissection; SPECT = single photon emission computed tomography; Tc-99m = radiolabelled technetium

DATA

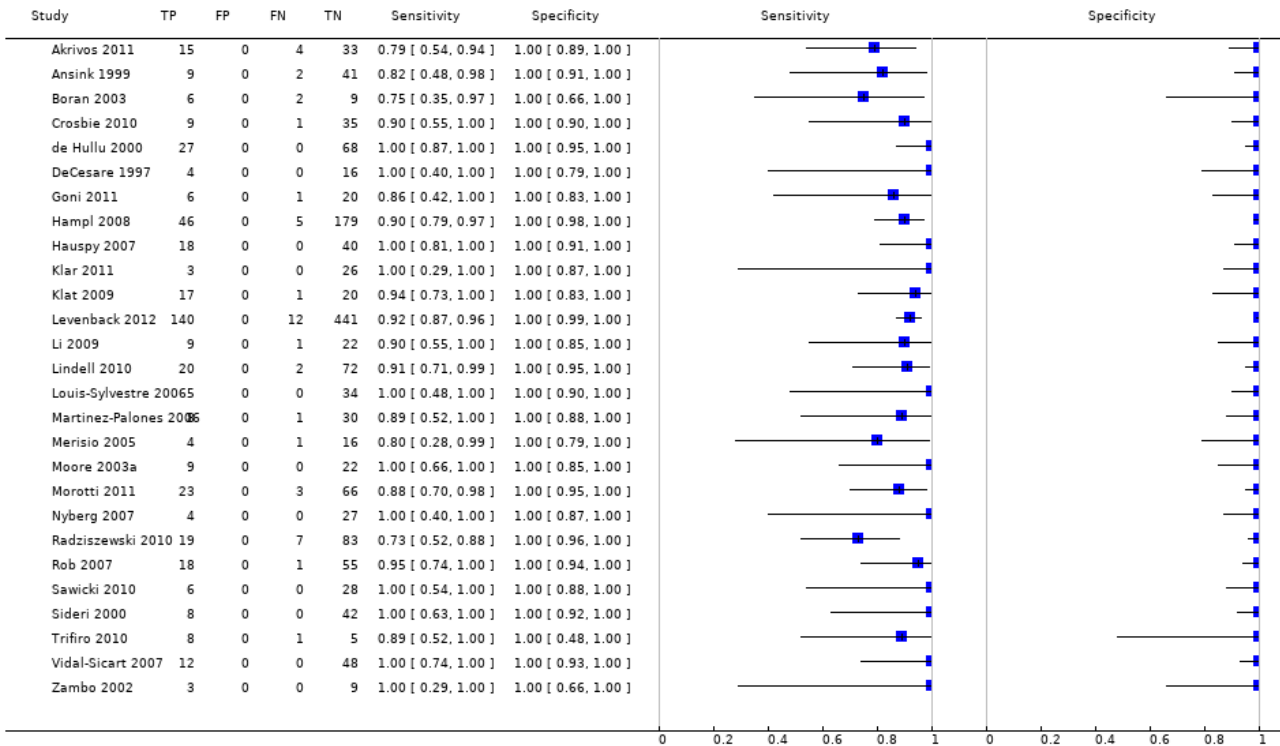
Presented below are all the data for all of the tests entered into the review.

Table Tests. Data tables by test

Test	No. of studies	No. of participants
1 Per groin: All tests	27	1988
2 Per woman: All tests	29	1286
3 Per groin: Blue dye only	5	290
4 Per woman: Blue dye only	3	68
5 Per groin: Tc-99m only	8	296
6 Per woman: Tc-99m only	8	149
7 Per groin: Combined test	13	1039
8 Per woman: Combined test	12	390
9 Per groin: Mixed tests	7	1030
10 Per woman: Mixed tests	7	679
11 Per groin: Tumour size < 4 cm	4	485
12 Per woman: Tumour size < 4 cm	2	74

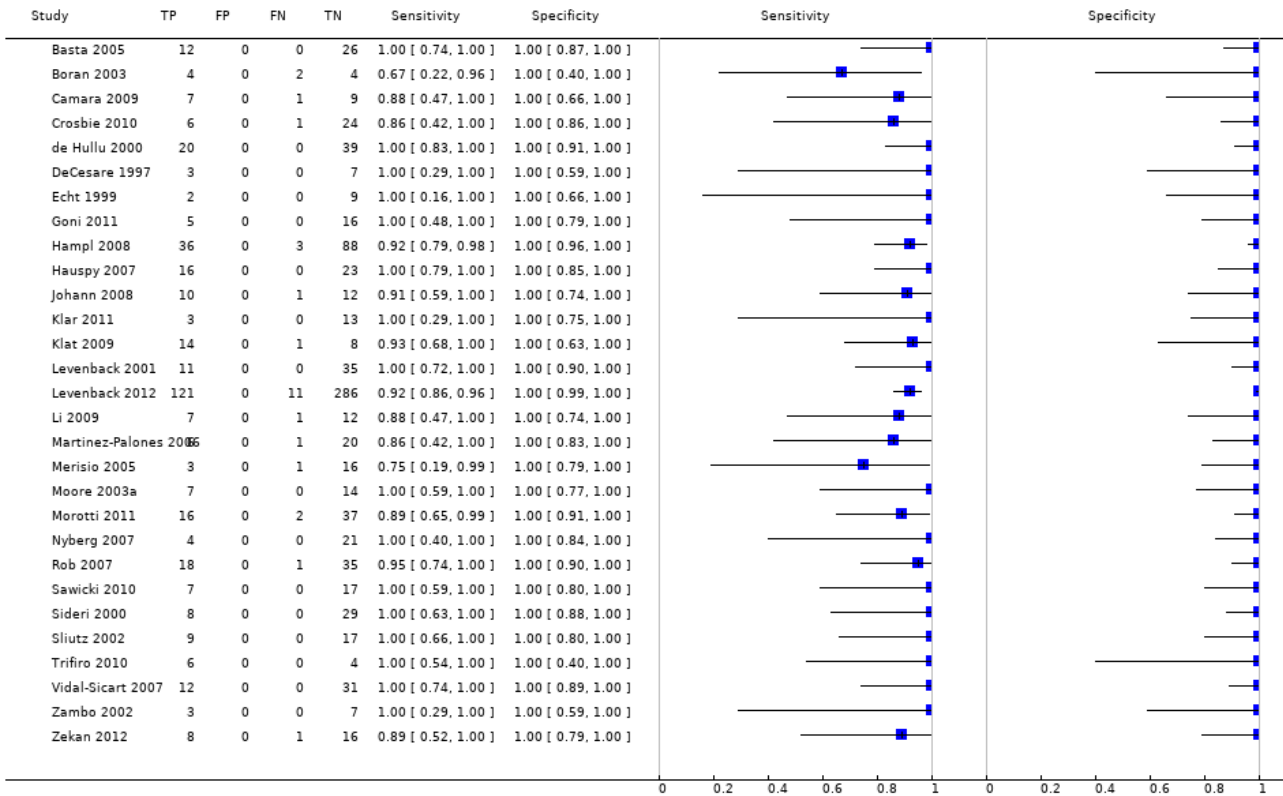
Test 1. Per groin: All tests.

Review: Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer
Test: 1 Per groin: All tests



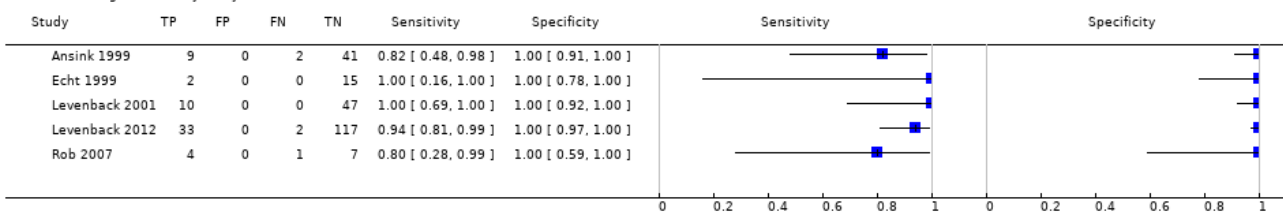
Test 2. Per woman: All tests.

Review: Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer
Test: 2 Per woman: All tests



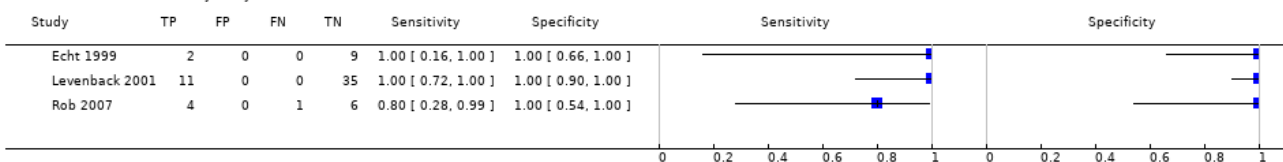
Test 3. Per groin: Blue dye only.

Review: Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer
Test: 3 Per groin: Blue dye only



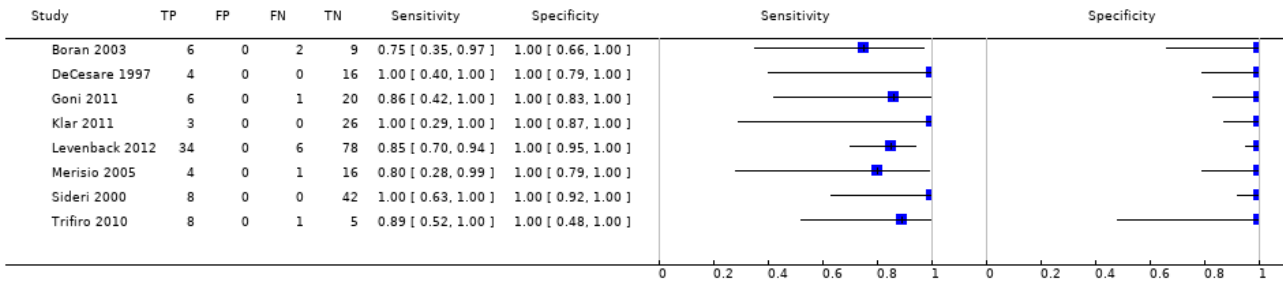
Test 4. Per woman: Blue dye only.

Review: Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer
Test: 4 Per woman: Blue dye only



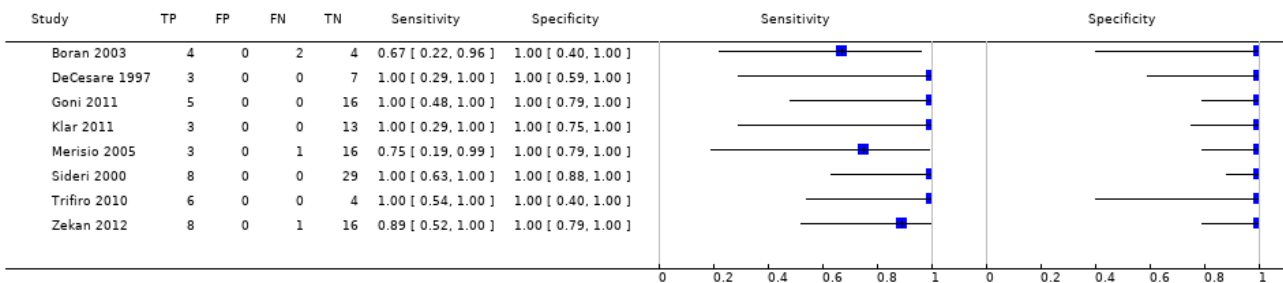
Test 5. Per groin: Tc-99m only.

Review: Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer
Test: 5 Per groin: Tc-99m only



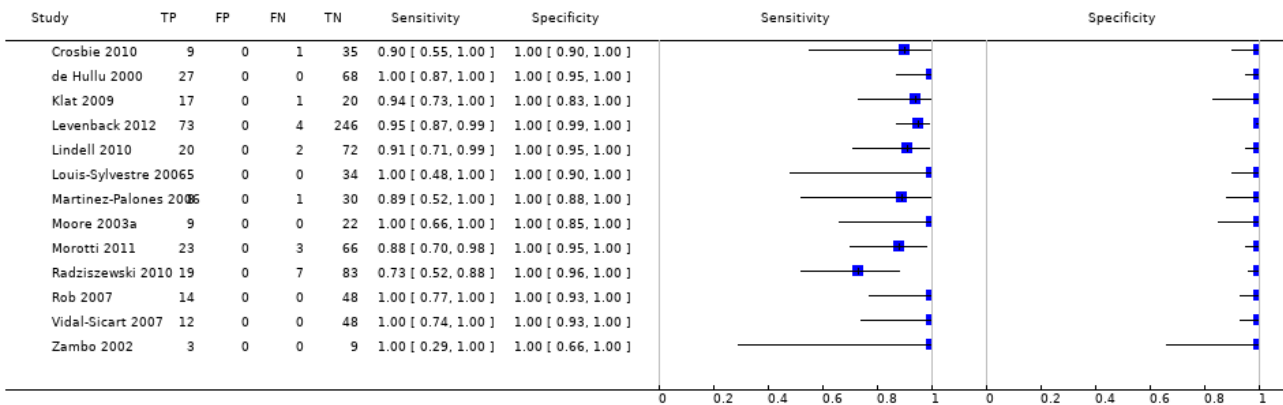
Test 6. Per woman: Tc-99m only.

Review: Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer
Test: 6 Per woman: Tc-99m only



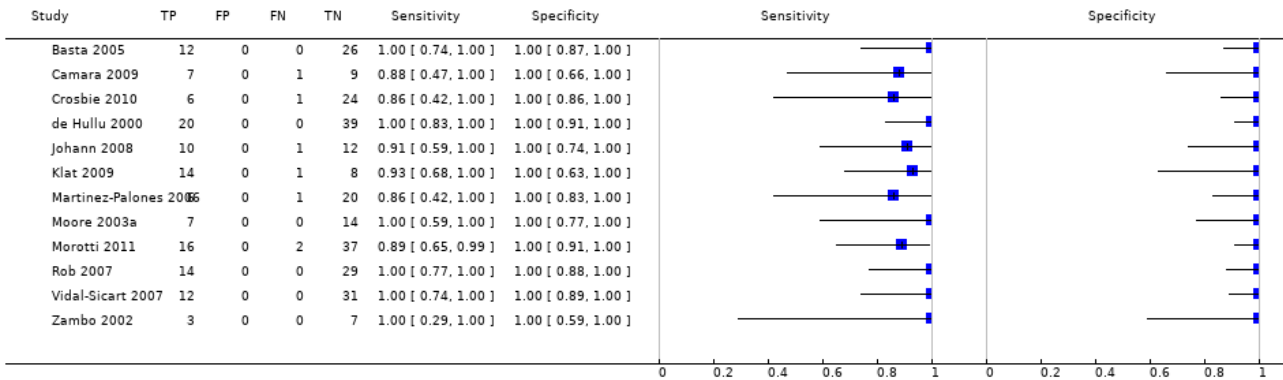
Test 7. Per groin: Combined test.

Review: Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer
Test: 7 Per groin: Combined test



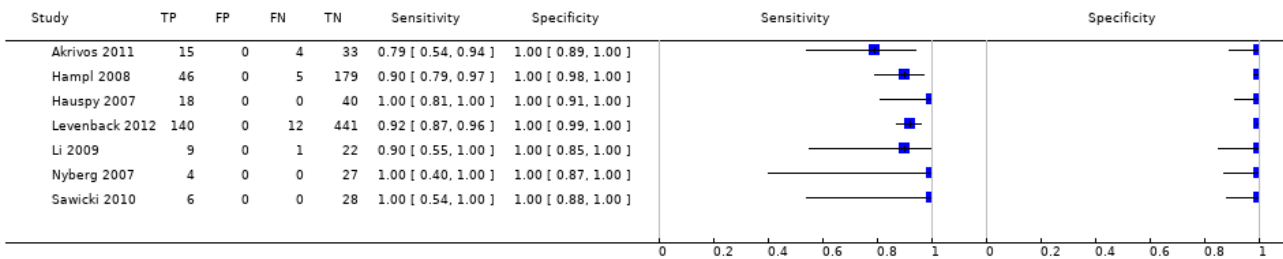
Test 8. Per woman: Combined test.

Review: Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer
Test: 8 Per woman: Combined test



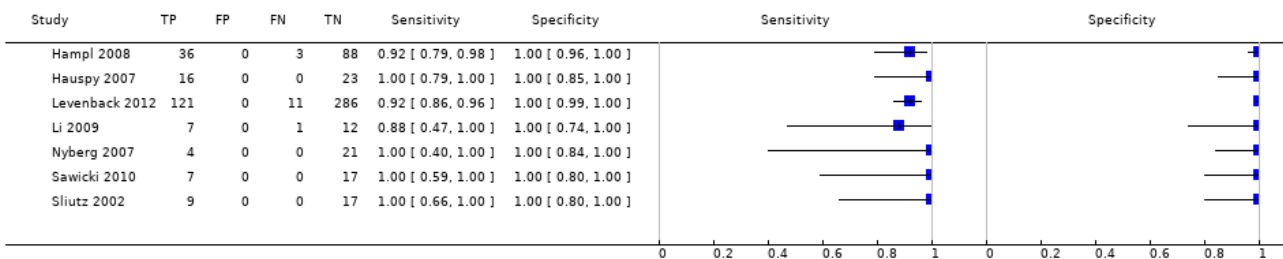
Test 9. Per groin: Mixed tests.

Review: Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer
Test: 9 Per groin: Mixed tests



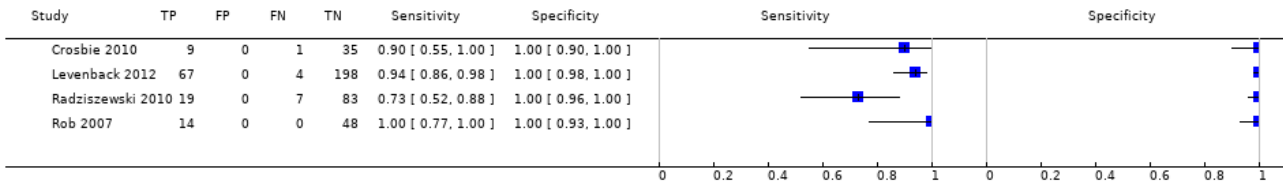
Test 10. Per woman: Mixed tests.

Review: Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer
Test: 10 Per woman: Mixed tests



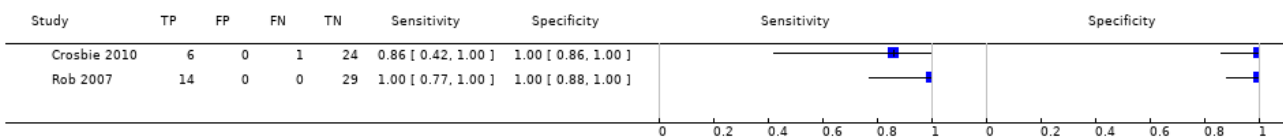
Test 11. Per groin: Tumour size < 4 cm.

Review: Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer
Test: 11 Per groin: Tumour size < 4 cm



Test 12. Per woman: Tumour size < 4 cm.

Review: Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer
Test: 12 Per woman: Tumour size < 4 cm



ADDITIONAL TABLES

Table 1. FIGO staging of vulval cancer *

Stage I	Tumour confined to the vulva
1A	Lesions ≤ 2 cm in size, confined to the vulva or perineum and with stromal invasion ≤ 1.0 mm**, no nodal metastasis
1B	Lesions > 2 cm in size or with stromal invasion > 1.0 mm*, confined to the vulva or perineum, with negative nodes
Stage II	Tumour of any size with extension to adjacent perineal structures (1/3 lower urethra, 1/3 lower vagina, anus) with negative nodes
Stage III	Tumour of any size with or without extension to adjacent perineal structures (1/3 lower urethra, 1/3 lower vagina, anus) with positive groin lymph nodes
IIIA	(i) With 1 lymph node metastasis (≥ 5 mm), or (ii) 1 to 2 lymph node metastasis(es) (< 5 mm)
IIIB	(i) With 2 or more lymph node metastases (≥ 5 mm), or (ii) 3 or more lymph node metastases (< 5 mm)
IIIC	With positive nodes with extracapsular spread
Stage IV	Tumour invades other regional (2/3 upper urethra, 2/3 upper vagina), or distant structures
IVA	(i) upper urethral and/or vaginal mucosa, bladder mucosa, rectal mucosa, or fixed to pelvic bone, or (ii) fixed or ulcerated groin lymph nodes

Table 1. FIGO staging of vulval cancer * (Continued)

IVB Any distant metastasis including pelvic lymph nodes

*Pecorelli 2009.

*The depth of invasion is defined as the measurement of the tumour from the epithelial-stromal junction of the adjacent, most superficial dermal papilla to the deepest point of invasion.

Table 2. Sentinel node detection rates of included studies

Study ID	SN detection (per woman) n/N (%)	SN detection (per groin) n/N (%)	Undetected +ve groins n/N (%)
Blue dye only	108/131 (82)	148/228 (65)	
1. Ansink 1999	42/51 (82)	52/93 (56)	3/41 (7)
2. Akrivos 2011*	NR	10/14 (71)	0/4 (0)
3. Echt 1999	9/12 (75)	17/23 (74)	2/6 (33)
4. Levenback 2001	46/52 (88)	57/76 (75)	2/19 (10)
5. Rob 2007*	11/16 (69)	12/22 (55)	1/10 (10)
Tc-99m only	152/159 (96)	158/189 (84)	
1. Boran 2003	10/10 (100)	17/17 (100)	0/0 (0)
2. Zekan 2012	25/25 (100)	NR	NR
3. Merisio 2005	20/20 (100)	21/31 (68)	1/10 (10)
4. Goni 2011	21/24 (88)	NR	NR
5. Sideri 2000	44/44 (100)	61/77 (79)	NR
6. DeCesare 1997	10/10 (100)	20/20 (100)	0/0 (0)
7. Klar 2011	12/16 (75)	25/29 (86)	0/4 (0)
8. Trifiro 2010	10/10 (100)	14/15 (93)	NR
Combined tests	365/371 (98)	562/607 (93)	
1. Radziszewski 2010	NR	106/107 (99)	NR
2. Basta 2005	38/39 (97)	NR	NR
3. Camara 2009	15/17 (88)	NR	NR

Table 2. Sentinel node detection rates of included studies (Continued)

4. Crosbie 2010	31/32 (97)	NR	0/1 (0)
5. de Hullu 2000	59/59 (100)	95/107 (89)	NR
6. Louis-Sylvestre 2006**	NR	39/52 (75)	3/13 (23)
7. Martinez-Palones 2006	27/28 (96)	39/40 (98)	0/1 (0)
8. Rob 2007*	43/43 (100)	62/64 (97)	0/2 (0)
9. Vidal-Sicart 2007**	43/43 (100)	60/64 (94)	NR
10. Moore 2003a	21/21 (100)	31/31 (100)	0/0 (0)
11. Morotti 2011	55/56 (98)	92/101 (91)	NR
13. Johann 2008	NR	NR	NR
14. Klat 2009	23/23 (100)	38/41 (93)	0/3 (0)
15. Zambo 2002	10/10 (100)	NR	NR
Mixed tests	786/827 (95)	1092/1349 (81)	
1. Levenback 2012	418/452 (92)	593/772 (77)	NR
2. Akrivos 2011	34/34 (100)	52/64 (81)	0/12 (0)
3. Hauspy 2007	39/41 (95)	58/68 (85)	NR
4. Sawicki 2010	24/24 (100)	34/39 (87)	1/5 (20)
5. Hampl 2008	125/127 (99)	228/230 (99)	NR
6. Lindell 2010	75/77 (97)	94/130 (72)	8/36 (22)
7. Nyberg 2007	25/25 (100)	NR	NR
8. Li 2009	20/21 (95)	NR	NR
9. Sliutz 2002	26/26 (100)	32/46 (70)	NR

*Data separated from total detection rates for mixed tests.

**Excluding women with suspicious nodes.

NR = not reported

APPENDICES

Appendix 1. MEDLINE search strategy

MEDLINE (OvidSP) (MEDLINE[®] In-Process & Other Non-Indexed Citations and Ovid MEDLINE[®] 1946 to present: February week 4 2013)

1. ((vulva* or clitoris or clitoral) adj5 (tumor* or tumour* or neoplas* or cancer* or carcinoma* or malignan* or metasta* or micrometasta* or carcinogen* or adenocarcinoma* or adenosquamous or growth*)).tw,ot.
2. Vulvar Neoplasms/
3. Neoplasm Invasiveness/ or Lymphatic Metastasis/ or Neoplasm Micrometastasis/ or Neoplasm Recurrence, Local/ or Neoplasm, Residual/
4. Neoplasm Staging/
5. Carcinoma, Squamous Cell/
6. ((lymphovascular or lympho-vascular) adj4 invasiv*).tw,ot.
7. (lymph* adj4 metasta*) or (detect* adj4 metasta*).tw,ot.
8. (vulva* or clitoris or clitoral).tw,ot.
9. Clitoris/
10. Vulva/
11. (3 or 4 or 5 or 6 or 7) and (8 or 9 or 10)
12. 1 or 2 or 11
13. Lymph Nodes/
14. Sentinel Lymph Node Biopsy/
15. ((groin or inguin*) adj3 (dissect* or status or node*1)).tw,ot.
16. Coloring Agents/
17. Technetium Tc 99m Sulfur Colloid/
18. Technetium Tc 99m Aggregated Albumin/
19. Radiopharmaceuticals/
20. Rosaniline Dyes/
21. Methylene Blue/
22. (technetium or tc 99m* or 99mtc* or blue dye*1 or patent blue or methylene blue or isosulfan or iso sulfan or lymphazurin blue or radiocolloid*).tw,ot,nm.
23. (sentin?l adj3 node*1).tw,ot.
24. (lymphoscintigraph* or lymphoscintigram*).tw,ot.
25. Lymphography/
26. (scintiphotograph* or scintigraph* or scintigram*).tw,ot.
27. (gamma adj3 (camera*1 or counter*1 or probe*1)).tw,ot.
28. radioisotope*.tw,ot.
29. (tracer or tracers).tw,ot.
30. ((trace or tracing or traceable) adj agent*1).tw,ot.
31. radiotracer*.tw,ot.
32. Fluorescence/
33. fluorescen*.tw,ot.

34. (microbubble or micro-bubble).tw,ot.
35. paraffin.tw,ot.
36. frozen section*.tw,ot.
37. (ultrastaging or ultra staging).tw,ot.
38. (microsection* or micro section*).tw,ot.
39. cytokeratin.tw,ot.
40. ((haematoxylin adj2 eosin) or (hematoxylin adj2 eosin)).tw,ot.
41. Gamma Cameras/
42. Frozen Sections/
43. (immuno-histo-chemistry or immunohistochemistry or immunohistochemical or immuno-histo-chemical).tw,ot.
44. (localisation or localization).tw,ot.
45. ((lymph node*1 or lymph nodal) adj3 (assess* or evaluat* or observ* or involve* or biopsy or biopsies)).tw,ot.
46. or/13-45
47. 13 and 46
48. exp animals/ not humans/
49. 47 not 48

Appendix 2. EMBASE search strategy

EMBASE (OvidSP) (1974 to 2013 March 12 Week 10)

1. ((vulva or clitoris or clitoral) adj5 (tumor* or tumour* or neoplas* or cancer* or carcinoma* or malignan* or metasta* or micrometasta* or carcinogen* or adenocarcinoma* or adenosquamous or growth*)).tw,ot.
2. exp vulva tumor/
3. cancer invasion/
4. exp lymph node metastasis/
5. micrometastasis/
6. tumor recurrence/
7. cancer recurrence/
8. cancer localization/
9. cancer staging/
10. squamous cell carcinoma/
11. ((lymphovascular or lympho-vascular) adj4 invasiv*).tw,ot.
12. (lymph* adj4 metasta*).tw,ot.
13. (vulva* or clitoris or clitoral).tw,ot.
14. clitoris/
15. vulva/
16. or/3-12

17. or/13-15
18. 16 and 17
19. 1 or 2 or 18
20. sentinel lymph node biopsy/
21. ((groin or inguin*) adj3 (dissect* or status or node*1)).tw,ot.
22. technetium 99m/
23. coloring agent/
24. fuchsine/
25. methylene blue/
26. radiopharmaceutical agent/
27. (technetium or tc 99m* or 99mtc* or blue dye*1 or patent blue or methylene blue or isosulfan or iso sulfan or lymphazurin blue or radiocolloid*).tw,ot.
28. (sentin?l adj3 node*1).tw,ot.
29. (lymphoscintigraph* or lymphoscintigram*).tw,ot.
30. lymphography/
31. (scintiphotograph* or scintigraph* or scintigram*).tw,ot.
32. (gamma adj3 (camera*1 or counter*1 or probe*1)).tw,ot.
33. exp scintillation camera/
34. radioisotope*.tw,ot.
35. (tracer or tracers).tw,ot.
36. ((trace or tracing or traceable) adj agent*1).tw,ot.
37. radiotracer*.tw,ot.
38. fluorescence/
39. fluorescen*.tw,ot.
40. (microbubble or micro-bubble).tw,ot.
41. paraffin.tw,ot.
42. frozen section*.tw,ot.
43. (ultrastaging or ultra staging).tw,ot.
44. (microsection* or micro section*).tw,ot.
45. cytokeratin.tw,ot.
46. ((haematoxylin adj2 esoin) or (hematoxylin adj2 eosin)).tw,ot.
47. frozen section/
48. (immuno-histo-chemistry or immunohistochemistry or immunohistochemical or immuno-histo-chemical).tw,ot.
49. (localisation or localization).tw,ot.
50. ((lymph node*1 or lymph nodal) adj3 (assess* or evaluat* or observ* or involve* or biopsy or biopsies)).tw,ot.

51. sentinel lymph node/

52. or/20-51

53. 52 and 19

54. (exp animal/ or exp invertebrate/ or animal.hw. or nonhuman/) not (exp human/ or human cell/ or (human or humans or man or men or wom?n).ti.)

55. 53 not 54

Appendix 3. Assessment of methodological quality criteria

We considered the following core QUADAS-2 domains in the current review.

DOMAIN 1: PATIENT SELECTION

A. Risk of bias

1. Was a consecutive or random sample of patients enrolled? Yes/No/Unclear
2. Did the study avoid inappropriate exclusions? Yes/No/Unclear
3. Was the spectrum of patients representative of the patients who will receive the test in practice?
 - Yes: at least 90% of patients had FIGO stage 1B or higher vulval cancer
 - No: less than 90% of patients had FIGO stage 1B or higher vulval cancer
 - Unclear: stage of disease was not reported or could not be clearly determined

OVERALL RISK OF BIAS: LOW/HIGH/UNCLEAR

B. Applicability

Is there concern that the included patients do not match the review question?

CONCERN: LOW/HIGH/UNCLEAR

DOMAIN 2: SENTINEL NODE ANALYSIS (INDEX TEST)

A. Risk of bias

1. Were the index test results interpreted without knowledge of the results of the reference standard? (Reference standard results blinded)
 - Yes: histological analysis of sentinel node was carried out without the knowledge of the histological analysis of groin lymph node
 - No: histological analysis of sentinel node was carried out with the knowledge of the histological analysis of groin lymph node
 - Unclear: it is unclear whether histological analysis of sentinel node was carried out with or without the knowledge of the histological analysis of groin lymph node

OVERALL RISK OF BIAS: LOW/HIGH/UNCLEAR

B. Applicability

Is there concern that the index test, its conduct or interpretation differ from the review question?

CONCERN: LOW /HIGH/UNCLEAR

DOMAIN 3: GROIN LYMPH NODE ANALYSIS (REFERENCE STANDARD)

A. Risk of bias

1. Is the reference standard likely to correctly classify the target condition
 - Yes: all the patients received standard groin lymph node removal
 - No: only some patients received standard groin lymph node removal
 - Unclear: surgical extent of lymph node removal is not reported or could not be clearly distinguished

2. Were the reference standard results interpreted without knowledge of the results of the index test? (Index test results blinded)

- Yes: histological analysis of groin lymph node removal was carried out without the knowledge of the histological analysis of sentinel node
- No: histological analysis of groin lymph node removal was carried out with the knowledge of the histological analysis of sentinel node
- Unclear: it is unclear whether the histological analysis of groin lymph node removal was carried out with or without the knowledge of histological analysis of sentinel node

OVERALL RISK OF BIAS: LOW/HIGH/UNCLEAR

B. Applicability

Is there concern that the target condition as defined by the reference standard does not match the review question?

CONCERN: LOW /HIGH/UNCLEAR

DOMAIN 4: FLOW AND TIMING

A. Risk of bias

1. Was there an appropriate interval between index test and reference standard?

- Yes: less than or equal to four weeks between index test and reference standard
- No: more than four weeks between index test and reference standard
- Unclear: not reported, variable or could not be clearly determined

2. Were other imaging tests performed prior to sentinel node test to rule out groin lymph node metastasis?

- Yes: participants underwent US, CT, MRI or PET-CT to detect groin metastasis prior to sentinel node procedure
- No: US, CT, MRI or PET-CT was not carried out prior to sentinel node procedure
- Unclear: insufficient information as to whether US, CT, MRI or PET-CT was carried out to detect groin metastasis prior to sentinel node procedure

3. Did all patients receive a reference standard?

- Yes: whole population received reference standard
- No: reference standard not carried out in the whole population
- Unclear: no clear information on what proportion of population received reference standard

4. Did patients receive the same reference standard?

- Yes: all the patients received same reference standard regardless of index test results
- No: not all the patients received same reference standard regardless of index test results
- Unclear: it is unclear whether all the patients received same reference standard regardless of index test results

5. Were un-interpretable/intermediate test results reported?

- Yes: all un-interpretable results (sentinel node detection and its histological analysis) were reported
- No: not all un-interpretable results were reported
- Unclear: it is not possible to determine if all un-interpretable results were reported

6. Were withdrawals from the study explained?

- Yes: all withdrawals from study were explained
- No: not all withdrawals from study were explained
- Unclear: insufficient information to determine if all withdrawals were explained

7. Had test operators had appropriate training?

- Yes: experience of at least 10 sentinel node procedures prior to taking part in the study
- No: experience of fewer than 10 sentinel node procedures prior to taking part in the study
- Unclear: insufficient information on training of surgeon taking part in the study

OVERALL RISK OF BIAS: LOW/HIGH/UNCLEAR

WHAT'S NEW

Date	Event	Description
21 September 2016	Amended	Contact details updated.

HISTORY

Protocol first published: Issue 2, 2013

Review first published: Issue 6, 2014

Date	Event	Description
1 April 2015	Amended	Contact details updated.
11 February 2015	Amended	Contact details updated.

CONTRIBUTIONS OF AUTHORS

Guarantor of the review: AP
 Conceiving the idea: AP
 Designing and co-ordinating the review: AP
 Data collection for the review: TAL, AP
 Designing search strategies: Anne Eisinga
 Undertaking searches: Anne Eisinga
 Screening search results: TAL, AP
 Organising retrieval of papers: TAL, AP
 Screening retrieved papers against inclusion criteria: TAL, AP
 Appraising quality of papers: TAL, AP
 Extracting data from papers: TAL, AP
 Writing to authors of papers for additional information: TAL
 Providing additional data about papers: TAL, AP, NR
 Obtaining and screening data on unpublished studies: TAL, AP
 Data management of the review: TAL, AP
 Entering data into RevMan 5.2: TAL, AP
 Analysis and interpretation of data: TAL, AP, AB
 Providing a methodological perspective: TAL, AB
 Providing a clinical perspective AP, NR
 Providing a consumer perspective: AP, NR, AB, AR, PMH, RN
 Writing the review: TAL, AP
 Providing general advice on the review: AP, TAL, NR, AR, PMH, RN
 Securing funding for the review: AP, NR, RN

DECLARATIONS OF INTEREST

None known.

SOURCES OF SUPPORT

Internal sources

- Department of Health, UK.

NIHR Cochrane Programme Grant support from 'Optimising care, diagnosis and treatment pathways to ensure cost effectiveness and best practice in gynaecological cancer. Improving the evidence for the NHS.' CPG-10/4001/12

External sources

- No sources of support supplied

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

In the protocol, we planned only to include studies for which 'per groin' data were available. However, for the review, we amended this to include studies that presented 'per woman' data only, and we analysed these 'per woman' data separately.

INDEX TERMS**Medical Subject Headings (MeSH)**

Coloring Agents; Groin; Lymph Node Excision; Lymph Nodes [*pathology]; Lymphatic Metastasis; Neoplasm Staging; Randomized Controlled Trials as Topic; Sentinel Lymph Node Biopsy [*methods]; Technetium; Vulvar Neoplasms [*pathology]

MeSH check words

Female; Humans