FULL PAPER

# BJC

# Sentinel lymph node biopsy in vulval cancer: systematic review and meta-analysis

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

[1]Centre for Primary Care and Public Health, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK; [2]Unit of Health Economics, University of Birmingham, Birmingham, UK; [3]Barts Cancer Institute, London, UK; [4]Arcana Institute, Kracow, Poland; [5]Addenbrooke's Hospital NHS Trust, Cambridge, UK and [6]Pan Birmingham Gynaecological Cancer Centre, City Hospital and School of Cancer Sciences, University of Birmingham, Birmingham B15 2TT, UK

**Background:** The purpose of this study was to determine the accuracy of sentinel lymph node (SLN) biopsy with technetium 99 (99mTc) and/or blue dye-enhanced lymphoscintigraphy in vulval cancer.

**Methods:** Sensitive searches of databases were performed upto October 2013. Studies with at least 75% of women with FIGO stage IB or II vulval cancer evaluating SLN biopsy with 99mTc, blue dye or both with reference standard of inguinofemoral lymphadenectomy (IFL) or clinical follow-up were included. Meta-analyses were performed using Meta-Disc version 1.4.

**Results:** Of the 2950 references, 29 studies (1779 women) were included; most of them evaluated 99mTc combined with blue dye. Of these, 24 studies reported results for SLN followed by IFL, and 5 reported clinical follow-up only for SLN negatives. Pooling of all studies was inappropriate because of heterogeneity. Mean SLN detection rates were 94.0% for 99mTc, 68.7% for blue dye and 97.7% for both. SLN biopsy had pooled sensitivity of 95% (95% CI 92–98%) with negative predictive value (NPV) of 97.9% in studies using 99mTc/blue dye, ultrastaging and immunohistochemistry with IFL as reference. Pooled sensitivity for SLN with clinical follow-up for SLN-negatives was 91% (85–95%) with NPV 95.6%. Patients undergoing SLN biopsy experienced less morbidity than those undergoing IFL.

**Conclusions:** Sentinel lymph node biopsy using 99mTC, blue dye and ultrastaging with immunohistochemistry is highly accurate when restricted to carefully selected patients, within a rigorous protocol, with close follow-up and where sufficient numbers for learning curve optimisation exist. Patients must make an informed choice between the slightly higher groin recurrence rates of SLN biopsy *vs* the greater morbidity of IFL.

Vulval cancer accounts for ∼3–5% of all gynaecological malignancies and 1% of all cancers in women, with an estimated 27 000 women diagnosed each year (Berek and Hacker, 2005). Standard treatment for squamous cell carcinoma of the vulva involves excision of the primary tumour and inguinofemoral lymphadenectomy (IFL) in all but FIGO stage Ia or superficially invasive disease. Groin lymph node status has been identified as the most important factor in predicting mortality attributable to vulval cancer (Royal College of Obstetricians and Gynaecologists, 1999). The efficacy of this treatment is good, with reported groin recurrence rates varying between 1% and 10% (Burger *et al*, 1995; Bell *et al*, 2000). However, only a third of patients with early-stage disease will have lymph node metastases, and the remainder will not benefit from elective IFL while risking significant morbidity (de Hullu *et al*, 2006; van der Zee *et al*, 2008). Complications affect over 50% of patients having IFL, including infection of groin wounds, wound breakdown, lymphocyst formation, lymphoedema and cellulitis (Gould *et al*, 2001; Pereira de Godoy *et al*, 2002; Gaarenstroom *et al*, 2003; Beesley *et al*, 2007).

*Correspondence: Dr S Sundar; E-mail: s.s.sundar@bham.ac.uk

Inguinofemoral lymphadenectomy is the standard of care because unrecognised disease in the inguinofemoral lymph nodes is usually fatal. A sentinel lymph node (SLN) refers to the first lymph node that receives drainage directly from the primary tumour and therefore has the highest probability of containing metastatic disease. The SLNs can be identified by lymphoscintigraphy using the radioactive tracer Technetium 99 (99mTc) and/ or with blue dye. The lymph node obtained can be examined using standard histopathology with haematoxylin and eosin (H&E), frozen section or enhanced testing (ultrastaging) with serial sectioning of the lymph node and immunohistochemistry for cytokeratins. Accurate identification of the sentinel node in early-stage vulval cancer may potentially spare the patient from undergoing IFL with its associated morbidity. The diagnostic performance of SLN biopsy in the 'real-world' setting to guide omission of IFL where the SLN is negative is also not fully established, but is the subject of on-going multicentre studies (GOG-0270 and GROINSS V11). We conducted a systematic review to evaluate the accuracy of SLN biopsy in vulval cancer.

## MATERIALS AND METHODS

**Protocol development and overview.** A protocol was developed for undertaking systematic reviews of test accuracy, diagnostic and therapeutic impact. Scoping searches for relevant systematic reviews were conducted in MEDLINE, EMBASE and the Cochrane Library. Systematic reviews were carried out using established methods (Higgins and Green, 2011; Diagnostic Test Accuracy Working Group, 2012). Presentation of results is according to the PRISMA guidelines (Moher et al, 2009). Inclusion of studies, data extraction and quality assessment were carried out in duplicate using predesigned and piloted data extraction sheets with differences resolved by consensus and/or arbitration involving a third reviewer. A two-stage process was used, firstly by screening titles and abstracts. For all references categorised as 'include' or 'uncertain' by both reviewers, full text was retrieved wherever possible and final inclusion decisions were made on the full paper.

**Search strategy, inclusion and exclusion criteria and quality assessment.** Comprehensive searches from the inception of database to 25 October 2013 were conducted in MEDLINE, Embase, Science Citation Index, the Cochrane Library, MEDION, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, the Health Technology Assessment Database, Clinical Trials.com as well as a search of internet resources (UK Clinical Research Network Portfolio, specialist search gateways (OMNI and the National Cancer Institute), Google and Copernic) to identify relevant published and unpublished studies and studies in progress. Electronic searches were supplemented by checking reference lists, handsearching the journal *Gynecologic Oncology* and contact with authors of included studies for information on any relevant published or unpublished studies. No language restrictions were applied. Search strategies were designed from a series of test searches and discussions of the results of searches among the review team. Both MESH terms and text words were used and included 'vulva cancer', 'sentinel lymph node biopsy' and 'lymphoscintigraphy'.

The population of interest was women with early stages of vulval cancer: at least 75% of population with FIGO stage I and II or TNM categories T1-2 N0 M0. We excluded studies on patients with vulval melanomas, advanced cancer – FIGO stage IV, inoperable tumours, tumours unsuitable for primary surgery, patients with clinical suspicion of metastases, that is, with palpable inguinofemoral lymph nodes, enlarged lymph nodes (>1.5 cm) on imaging or cytologically proven inguinofemoral lymph node metastases at the start of the study. The index testing strategies were SLN biopsy with 99mTc, blue dye or combined technique

(99mTc with blue dye), with histopathology by H&E either on formalin-fixed or frozen sections or enhanced testing with thinner sections and/or immunohistochemistry. Where studies reported any of ultrastaging, serial sections, multiple slices, additional sections or step sections, these were all classified as 'ultrastaging'. Studies on other imaging modalities and novel metastasis detection techniques were excluded. Reference standard was histology of IFL or clinical follow-up for SLN-negative patients. Outcomes of interest were diagnostic accuracy, morbidity following SLN biopsy, mortality and disease-free survival, quality of life, and impact on surgeon's and team's skills and experience (learning curve). Studies with nonclinical outcomes and those that reported outcomes per groin only were excluded. Any prospective or retrospective test accuracy study designs, studies investigating the diagnostic and therapeutic impact with or without concurrent assessment of test accuracy and prospective cohort studies of outcomes of patients tested with 99mTc, blue dye or combined technique for SLN biopsy were included. Case studies were excluded.

Study quality was assessed using standard guidelines for test accuracy (QUADAS) and diagnostic and therapeutic impact studies (Meads and Davenport, 2009; Whiting et al, 2010).

**Statistical analysis.** RevMan version 5.2 (The Nordic Cochrane Centre, Copenhagen, Denmark) was used for statistical analyses and Meta-Disc 1.4 (Unit of Clinical Biostatistics team of the Ramón y Cajal Hospital, Madrid, Spain) was used for meta-analysis. Sensitivity, specificity, true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) were taken directly from the source papers. If that was not possible, values were calculated from data provided. Based on an investigation of heterogeneity, summary estimates of sensitivity, specificity and likelihood ratios (LRs) were derived as appropriate. Results were displayed graphically on Forest and receiver operating curve (ROC) plots. Summary SLN detection rates and their 95% confidence intervals were calculated using Meta-Disc.

## RESULTS

**Characteristics of included studies.** There were 2950 citations identified from searches, of which 82 full papers were obtained and 29 relevant studies (38 publications) were included. Figure 1 displays the PRISMA diagram. Most studies were small with <50 patients, but there were 3 larger studies with 127 patients (Hampl et al, 2008), 452 patients (269 with tumours under 4 cm; Levenback et al, 2012) and 403 patients (van der Zee et al, 2008). The characteristics of the studies are presented in Supplementary Tables 1 and 2. Patients with early-stage vulval cancer varied between 86% and 100% of subjects, with 18 out of 29 (62%) studies having all patients at early stage. Where reported, tumour locations were evenly spread between midline or lateral positions. The most commonly reported tumour types were squamous cell carcinoma. Either TNM and FIGO staging alone or a combination of both were given in all studies.

Index tests and histopathological techniques used for SLN biopsy, and reference standards used in each of the studies are summarised in Table 1. Out of 29 studies, 24 presented results for both blue dye and 99mTc tests for SLN identification, although not all patients underwent both tests in every study. Presentation of results varied considerably. In 21 studies, detection rates per groin were presented for each test separately, and both tests combined. It is worth noting that SLN's were always subject to rigorous examination, whereas histopathological techniques for corresponding IFL nodes were less detailed and were assumed to be H&E unless otherwise stated.

**Quality of included studies.** Quality assessment is reported in Supplementary Table 3. Of the 29 included studies, 4 had no
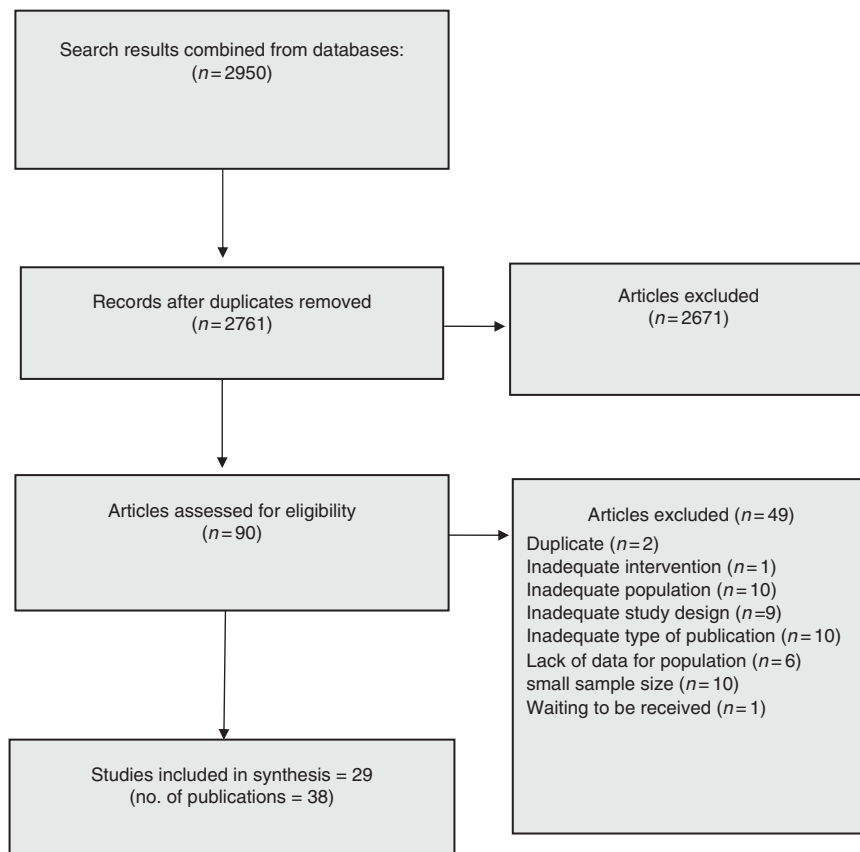
Figure 1. The PRISMA diagram for diagnostic review.

information about histopathological methods (Pitynski et al, 2003; Nyberg et al, 2007; Vakselj and Bebar, 2007; Camara et al, 2009;). One study used frozen section as reference standard (Camara et al, 2009). In 19 studies, upon negative H&E, immunohistochemistry using antibodies such as AE1, AE3, S-100, HMB-45, Mab, CKMNF, CK-88 and EMA was performed. In others, additional sections/ ultrastaging was used if samples were negative by H&E staining and standard sectioning. Thickness of slices varied between studies. Only de Hullu et al (2000) achieved blinding of pathologists.

**Test accuracy results.** Reporting of results was frequently ambiguous, making it difficult to distinguish between patients who had no SLN detected from those with negative SLN biopsy on histology. Results of test accuracy are presented on the basis of detected SLN. In all, 24 studies evaluated the test accuracy of SLN with IFL for all, and 5 studies evaluated SLN with clinical follow-up for test-negative patients and IFL for patients with malignancy detected in SLN biopsy (Van den Eynden et al, 2003; Terada et al, 2006; Moore et al, 2008; van der Zee et al, 2008; Achimas-Cadariu et al, 2009). For calculation of sensitivity and specificity, studies have been categorised into groups by the reference standards used, the index test used and the histopathological techniques used as follows:

1. Inguinofemoral lymphadenectomy for all

   ○ Technetium 99 with blue dye (Supplementary Table 4)
      ■ Haematoxylin and eosin only or insufficient details to determine whether immunohistochemistry or ultrastaging were used
      ■ Immunohistochemistry
      ■ Frozen section only
      ■ Immunohistochemistry with ultrastaging

   ○ Technetium 99 only (Supplementary Table 5)
      ■ Haematoxylin and eosin only or insufficient details to determine whether immunohistochemistry or ultrastaging were used
      ■ Immunohistochemistry

   ○ Blue dye only (Supplementary Table 6)
      ■ Immunohistochemistry with ultrastaging

2. Inguinofemoral lymphadenectomy for SN positive and clinical follow-up for SLN negative (Supplementary Table 7)

   ○ Technetium 99 and blue dye
      ■ Immunohistochemistry
      ■ Ultrastaging

Point estimates of specificity are 100%. The ROC plane was unhelpful and not presented. Although the point estimates of sensitivity are close to 100%, confidence intervals were wide, reflecting the small sample sizes available.

The SLN detection rates for each of the analysed techniques (blue dye, 99mTc and blue dye/99mTc) are presented in Table 2. The detection rate calculated per patient was available in all included studies. Combined blue dye and 99mTc testing had the highest rate of SLN detection. Pooled rates are 94.0% for 99mTc (95% CI 90.5–96.4), 68.7% for blue dye alone (95% CI 63.1–74.0) and 97.7 (95% CI 96.6–98.5) for 99mTc and blue dye combined.

**Training and experience.** Studies commonly specified the first 10 cases as learning curve (Hampl et al, 2008; van der Zee et al, 2008; Achimas-Cadariu et al, 2009), after which SLN biopsy without IFL could be performed. Only Levenback et al (2001) calculated that

| Number | Author and year of publication | 99mTc | Blue dye | Both together | Histopathological techniques – SLN | Histopathological techniques – remaining nodes | Type of surgery (or radiotherapy) |
|---|---|---|---|---|---|---|---|
| 1 | Achimas-Cadariu et al (2009) | | | X | H&E, ultrastaging | NR | Radical vulvectomy (58%), 'modified' (41%) |
| 2 | Basta et al (2005) | Xa | Xa | X | SLN immunochemical stain for micrometastases | NR | NR |
| 3 | Brunner et al (2008) | X (91%) | | X (9%) | Frozen sections, H&E and, if −ve, immunohistochemistry for cytokeratins | Routine techniques | NR |
| 4 | Camara et al (2009) | X | X | X | Frozen section | NR | NR |
| 5 | Crosbie et al (2010) | | | X | H&E and, if −ve, with additional sections and immunohistochemistry for cytokeratins AE1-3 | NR | Radical excision (47%), unclear (53%) |
| 6 | de Cicco et al (2000) | X | | | H&E | H&E | Wide radical excision, hemivulvectomy or radical vulvectomy |
| 7 | de Hullu et al (2000) | | X | X | H&E and, if −ve, with additional sections and immunohistochemistry for cytokeratins AE1-3 | H&E | Radical excision (100%) |
| 8 | Hampl et al (2008) | X | X | X | H&E and, if −ve, with additional sections and immunohistochemistry for panycytokeratin antibody | NR | Hemivulvectomy (35%), vulvectomy (35%), local tumour resection (30%) |
| 9 | Hauspy et al (2007) | X | | X | Frozen section then serial sections H&E and immunohistochemistry for cytokeratins AE1-3 for some sections | H&E | Wide local excision (76%), radical vulvectomy (20%), radiotherapy (5%) |
| 10 | Johann et al (2008) | | | X | Step sectioning | Step sectioning | Radical vulvectomy (30%), hemivulvectomy (57%), wide excision (13%) |
| 11 | Klat et al (2009) | | | X | H&E, ultrastaging and immunohistochemistry for cytokeratins AE1-3 | NR | Radical surgery (100%) |
| 12 | Levenback et al (2001) | | X | | Frozen section if suspicious, step sectioning and some immunohistochemistry using several protocols | NR | NR |
| 13 | Lindell et al (2010) | | X (22%) | X (78%) | Step sections, H&E and, if −ve, immunohistochemistry for cytokeratin MNF116 | H&E | Vulvectomy (47%), hemivulvectomy (31%), wide local excision (22%) |
| 14 | Louis-Sylvestre et al (2006) | X (21%) | | X (79%) | Serial sections, H&E and immunohistochemistry for cytokeratins AE1 and AE3 | NR | NR |
| 15 | Martinez-Palonez et al (2006) | | | X | 0.2 mm sections, H&E and, if −ve, immunohistochemistry for cytokeratin and membrane epithelial antigen | NR | NR |
| 16 | Merisio et al (2005) | X | | | H&E, ultrastaging, immunohistochemistry for cytokeratins in 50% | Standard techniques | Radical vulvectomy or radical vulval excision (percentages NR) |

Table 1. Studies included in the systematic review showing details of index test and reference standard

**Table 1.** (Continued)

| Number | Author and year of publication | 99mTc | Blue dye | Both together | Histopathological techniques – SLN | Histopathological techniques – remaining nodes | Type of surgery (or radiotherapy) |
|---|---|---|---|---|---|---|---|
| 17 | Moore et al (2008) | | | X | H&E and ultrastaging | NR | Radical vulvectomy or radical vulval excision (percentages NR) |
| 18 | Nyberg et al (2007) | X (20%) | X (80%) | | Histopathology | NR | NR |
| 19 | Pitynski et al (2003) | X (14%) | X (86%) | | NR | NR | NR |
| 20 | Radziszewski et al (2010) | | | X | Multiple slices, H&E in 50% slices, H&E and immunohistochemistry in other 50% slices | H&E in 50% slices, H&E and immunohistochemistry in other 50% slices | NR |
| 21 | Rob et al (2007) | | X (27%) | X (73%) | Frozen section then serial sections, H&E and immunohistochemistry on every third slide | H&E | NR |
| 22 | Terada et al (2006) | | | X | Multiple slices, H&E and, if −ve, immunohistochemistry with cytokeratin antigen | NR | NR |
| 23 | Vakselj and Bebar (2007) | | | X | NR | NR | Tumour excised (100%) |
| 24 | Van den Eynden et al (2003) | | | X | H&E and, if −ve, ultrastaging, and immunohistochemistry for cytokeratins AE1 and AE3 | NR | NR |
| 25 | van der Zee et al (2008) | | | X | Frozen section or routine histopathology, ultrastaging | H&E | Radical excision (100%) |
| 26 | Vidal-Sicart et al (2007) | | | X | Multiple slices, H&E and, if −ve, H&E with immunohistochemistry | H&E | Radical vulvectomy or radical vulval excision (percentages NR) |
| 27 | Klar et al (2011) | X | | | Frozen section or routine histopathology and, if −ve, then ultrastaging and immunohistochemistry with cytokeratin | H&E | Radical excision (100%) |
| 28 | Levenback et al (2012) | X | X | X | H&E and, if −ve, ultrastaging and immunohistochemistry with cytokeratin | NR | NR |
| 29 | Zekan et al (2012) | X | | | Multiple slices, H&E and, if −ve, immunohistochemistry with cytokeratin | H&E | Radical excision (100%) |

Abbreviations: 99mTc = technetium 99; H&E = haematoxylin and eosin staining; NR = not recorded; SLN = sentinel lymph node.
[a]The 99mTc and blue dye discrepant results are shown in text and table.

the rate of SLN detection was worse in the first 2 years of the study (failure rate 16% vs 7%).

**Recurrence rates following SLN biopsy.** Two groups presented recurrences at follow-up. The first group used full IFL at initial operation to establish diagnostic accuracy, but also presented follow-up data (Martinez-Palonez et al, 2006; Vakselj and Bebar, 2007; Vidal-Sicart et al, 2007; Klat et al, 2009; Crosbie et al, 2010). The second group used clinical follow-up for SLN-negative patients to establish diagnostic accuracy (Van den Eynden et al, 2003; Terada et al, 2006; Moore et al, 2008; van der Zee et al, 2008; Achimas-Cadariu et al, 2009). In the first group, number of recurrences seen (18) was less than the number of SLN-positive patients (36) in the 4 studies that present follow-up data by SLN status (Martinez-Palonez et al, 2006; Vakselj and Bebar, 2007;

Klat et al, 2009; Crosbie et al, 2010). The SLN-negative patients developed recurrence in 3 of these 4 studies (Martinez-Palonez et al, 2006; Vakselj and Bebar, 2007; Klat et al, 2009). Of these, two studies showed a higher recurrence rate in SLN-negative patients than in patients who underwent IFL after false-negative SLN biopsies (Martinez-Palonez et al, 2006; Vakselj and Bebar, 2007). This may imply a therapeutic benefit to IFL or the confounding effect of adjuvant radiotherapy.

In the second group with clinical follow-up for SLN-negative patients, recurrence rates for groin and distant recurrence were calculated (Supplementary Table 7). Pooled sensitivity from SLN with clinical follow-up (91% CI 85%–95%) is comparable to estimates where patients received IFL as the gold standard (pooled sensitivity of 95% (92–98%), but with a lower NPV (95.3% vs NPV of 97.9%; see Figures 2 and 3).

**Table 2.** The SLN detection rate of blue dye, 99mTc and both

| Number | Study | 99mTc | Blue dye | Both together | Study characteristics |
|---|---|---|---|---|---|
| 1 | Achimas-Cadariu et al (2009) | — | — | 43 Out of 46 (94%) | Only SCC, early and late stages |
| 2 | Basta et al (2005) | 38 Out of 39ª (97%) | 32 Out of 39ª (82%) | 38 Out of 39 (97%) | Not reported, early stage |
| 3 | Camara et al (2009) | 13 Out of 17 (76%) | 9 Out of 17 (53%) | 15 Out of 17 (88%) | Mostly SCC, early and late stages |
| 4 | Crosbie et al (2010) | | | 31 Out of 32 (97%) | Only SCC, early and late stages |
| 5 | de Cicco et al (2000) | 37 Out of 37 (100%) | | | Only SCC, early and late stages |
| 6 | de Hullu et al (2000) | | 35 Out of 59 (60%) | 59 Out of 59 (100%) | Only SCC, early stage |
| 7 | Hampl et al (2008) | 119 Out of 127 (94%) | 80Out of127 (63%) | 125 Out of 127 (98%) | Mostly SCC, early and late stages |
| 8 | Hauspy et al (2007) | | | 39 Out of 41 (95%) | Mostly SCC, early and late stages |
| 9 | Johann et al (2008) | | | 23 Out of 23 (100%) | Only SCC, early stage |
| 10 | Klar et al (2011) | 12 Out of 16 (75%) | | | Only SCC, early stage |
| 11 | Klat et al (2009) | | | 23 Out of 23 (100%) | Only SCC, early stage |
| 12 | Levenback et al (2001) | | 46 Out of 52 (88%) | | Mostly SCC, early and late stages |
| 13 | Martinez-Palonez et al (2006) | | | 27 Out of 28 (96%) | Mostly SCC, early stage |
| 14 | Merisio et al (2005) | 20 Out of 20 (100%) | | | Only SCC, early stage |
| 15 | Moore et al (2008) | | | 35 Out of 36 (97%) | Only SCC, early and late stages |
| 16 | Radziszewski et al (2010) | | | 56 Out of 62 (90%) | Only SCC, early stage |
| 17 | Terada et al (2006) | | | 21 Out of 21 (100%) | Only SCC, early stage |
| 18 | Vakselj and Bebar (2007) | | | 35 Out of 35 (100%) | Mostly SCC, early stage |
| 19 | Van den Eynden et al (2003) | | | 27 Out of 32 (84%) | Mostly SCC, early stage |
| 20 | van der Zee et al (2008) | | | 403 Out of 403 100% | Only SCC, early stage |
| 21 | Vidal-Sicart et al (2007) | | | 49 Out of 50 (98%) | Mostly SCC, early and late stages |
| 22 | Zekan et al (2012) | 25 Out of 25 (100%) | | | Only SCC, early stage |
| | Combined rates | 94.0% | 68.7% | 97.7% | |
| | 95% CI | 90.5–96.4 | 63.1–74.0 | 96.6–98.5 | |

Abbreviations: 99mTc = technetium 99; CI = confidence interval; SCC = squamous cell carcinoma; SLN = sentinel lymph node. Note that for Brunner et al (2008), Lindell et al (2010), Louis-Sylvestre et al (2006), Nyberg et al (2007), Pitynski et al (2003), Rob et al (2007) and Levenback et al, (2012) a single test was used for a proportion of patients and a combination of tests was used for the remainder, and hence the detection rate per patient is not specific to any single test or combination.
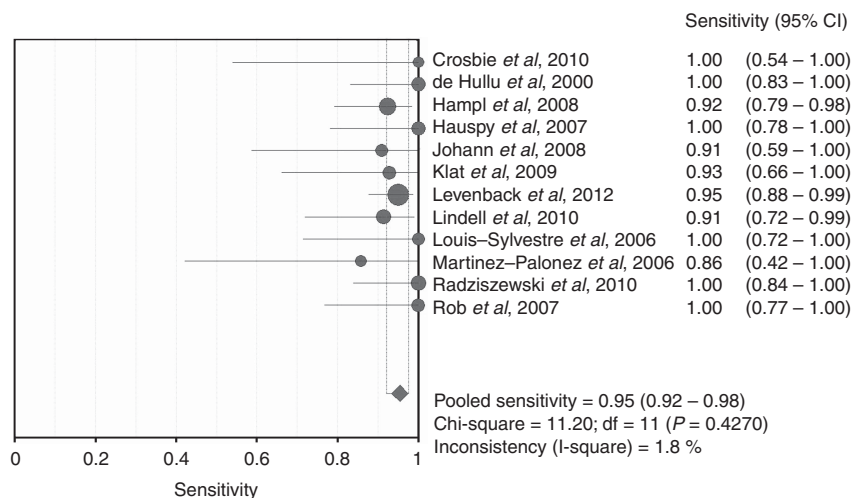[a]The 99mTc and blue dye discrepant results are shown in text and table.



Figure 2. Forest plot of sensitivity of SLN biopsy in group with IFL for all, 99mTc with blue dye – ultrastaging with immunohistochemistry.

**Survival rates.** Nine studies gave information about survival (Martinez-Palonez et al, 2006; Terada et al, 2006; Vakselj and Bebar, 2007; Moore et al, 2008; van der Zee et al, 2008; Achimas-Cadariu et al, 2009; Klat et al, 2009; Crosbie et al, 2010; Oonk et al, 2010). All studies were consistent with a relatively low survival rate for patients with groin relapse. Deaths were reported by Achimas-Cadariu et al (2009) (12 out of 46 patients, survival 61.2 months for whole cohort and16.2 months for 8 patients with relapse), Crosbie et al (2010) (2 out of 32 patients, follow-up 62 months), Klat et al (2009) (1out of 23 patients,
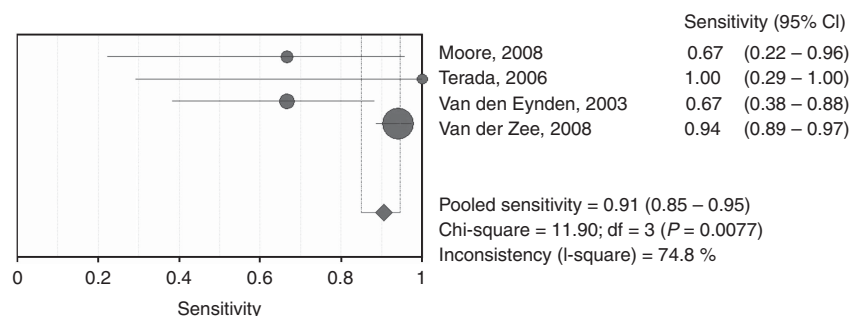
Figure 3. Forest plot of sensitivity of SLN biopsy in group with IFL for SN positive, clinical follow-up for SN negative, 99mTc and blue dye, ultrastaging, groin and distant recurrences only* (*data from Achimas-Cadariu et al, 2009 could not be included as there were insufficient data in paper).

| Table 3. Summary of adverse events from SLN or SLN biopsy with IFL | | | |
|---|---|---|---|
| **Author and year of publication** | **Complications** | **SLN alone** | **SLN biopsy + IFL** |
| van der Zee et al (2008); Brunner et al (2008); Achimas-Cadariu et al (2009); Moore et al (2008); Terada et al, 2006); Van den Eynden et al (2003); Johann et al (2008); Crosbie et al (2010) | Short-term adverse events[a] | Transient lymph oedema (13%), wound breakdown (11.7%) and wound cellulitis (4.5%) | Transient lymph oedema (39%), postoperative groin lymphocoele (5.5%), cellulitis arising in the labia majora (2.8%), wound cellulitis (9.5%) and seroma (4.3%), wound breakdown (34%) and wound cellulitis (21.3%), cellulitis (5.9%) and lymphocele (11.8%) |
| | Longer-term adverse events[a] | Lymphoedema (1.9%) and recurrent erysispelas (0.4%) | Wound infection (31%), wound dehiscence (25%), lymphocyst (22%) and chronic lymphoedema (16%), lymphoedema (25.2%) and recurrent erysipelas (16.2%) |
| Abbreviations: IFL = inguinofemoral lymphadenectomy; SLN = sentinel lymph node. [a]As defined by the papers. | | | |

follow-up 8–46 months), Moore et al (2008) (1 out of 35 patients died of intercurrent disease, follow-up 29 months), Terada et al (2006) (follow-up 55 months, 2/3 node-positive patients died of cancer), Vakselj and Bebar (2007) (6 out of 10 node-positive patients, 2 out of 25 node-negative patients, 1 died of disease at 49 months) and Vidal-Sicart et al (2007) (1 out of 50 patients died of disease, follow-up 20 months).

The largest study presented disease-specific survival for node-negative patients with a median follow-up of 35 months, with 202 out of 276 (73.2%) patients having at least 24 months of follow-up (van der Zee et al, 2008). The 3-year disease-specific survival for patients with unifocal vulval disease and negative SLN was 97.0%. In a subsequent paper, 5-year disease-specific survival was 77.3% in patients with positive SLN. However, survival varied depending on the histopathology technique used; 64.9% when malignant SLN were identified by routine pathology and 92.1% when identified by ultrastaging, and was higher with size of SLN metastases > 2 mm (Oonk et al, 2010).

**Quality of life.** One study (62 patients) investigating quality of life found few differences between SLN and IFL with EORTC QLQ-C30; only the financial difficulties scale was statistically significantly worse in the IFL group. For the FACT-V questionnaire, there were significantly worse results for the contentment functional scale, and oedema, complaints and stockings symptom scales (Oonk et al, 2009).

**Adverse events.** Information about adverse events was generally poorly reported. Eight studies (Table 3) provided data (Van den Eynden et al, 2003; Terada et al, 2006; Brunner et al, 2008; Johann et al, 2008; Moore et al, 2008; van der Zee et al, 2008; Achimas-Cadariu et al, 2009; Crosbie et al, 2010). Patients undergoing IFL had worse morbidity than those undergoing SLN alone. Definitions of morbidity are not standardised, and therefore statistical comparisons were not possible.

## DISCUSSION

This systematic review comprises 29 studies with information on test accuracy of 99mTc and/or blue dye identification of SLN biopsy with reference standard of either IFL for all (24 studies) or IFL for SLN-positive nodes (containing metastases) and clinical follow-up for SLN-negative nodes (5 studies). There were, in effect, three index tests (99mTc, blue dye and

99mTc with blue dye) and five reference standard groups (H&E only, insufficient details to determine histopathological techniques used, frozen section only, immunohistochemistry, ultrastaging and ultrastaging with immunohistochemistry). Therefore, calculating the sensitivity and specificity of finding metastases in a SLN biopsy compared with the reference standard was not straightforward and no meta-analysis of all 29 studies was appropriate. Unfortunately, it was not possible to establish the success rate of SLN detection (technical success) as distinct from sensitivity of SLN for detection of metastatic disease.

The largest group of studies (21 studies) using 99mTc, blue dye and immunohistochemistry demonstrated a pooled SLN detection rate of 97.7% (CI 96–98%) with blue dye and 99mTc, 94.0% (CI 90–96%) with 99mTc alone and 68.7% (CI 63–74%) with blue dye alone. We could not statistically compare accuracy of 99mTc with blue dye as there were insufficient studies of similar clinical characteristics to be able to conduct meta-regression. Nevertheless, these results provide evidence that a combination of blue dye/technetium and ultrastaging is the most accurate test. Using blue dye and technetium may also benefit the learning curve as blue dye enables direct visualisation of the SLN (Bass et al, 1999).

Recurrence occurred in two cohorts of patients; the first underwent IFL and then follow-up, and the second underwent follow-up after a negative SLN. The number of clinically apparent recurrences at follow-up was less than the number of SLN-positive patients. This is either because microscopic metastases detected with sophisticated histology techniques in SLN biopsy are less likely to be clinically significant than those detected with standard histopathology techniques, or because subsequent IFL and/or the subsequent adjuvant treatment is therapeutic. It is noteworthy that nodes retrieved as 'sentinel' were subject to intensive examination to detect micrometastases, whereas nodes retrieved at subsequent IFL were either subject to routine histopathology or details of histopathological examination were not presented. Thus, accuracy results are skewed to overscoring positives for SLN or underscoring positives for IFL. Unless both SLN and IFL are subjected to the same technique, the true value of micrometastases detected by ultrastaging of SLN will not be established. This review highlights the challenges in truly assessing the value of a test such as SLN when there is such variability in how the nodes are examined at pathology. There is an urgent need for consensus to define the standards of histopathology and the need for ultrastaging. Our review did not identify any papers that presented management of groin lymph nodes at recurrence.

One recent systematic review has reported high rates of detection of SLN nodes, but did not make any attempt to stratify studies by techniques used to examine SLN or differentiate studies based on follow-up for SLN negatives (Hassanzade et al, 2013). The strength of the test accuracy systematic review was the rigour of its conduct and the focus on comparing and contrasting the different types of index test and reference standard. The studies had considerable methodological limitations, including lack of an adequate description of inclusion criteria, population (especially stage of disease) and reference standard used. Histological methods varied considerably, particularly with regard to 'ultrastaging/additional sections', and the optimum methods of examining the SLN were unclear. The results from the two largest studies (van der Zee et al, 2008; Levenback et al, 2012) are extremely consistent showing FN rates of <3% in unifocal vulval tumours <4 cm in size and supporting the use of SLN biopsy in these patients. It is noteworthy that van der Zee et al (2008) implemented a robust training protocol before entry for SLN biopsy. Our results

show a higher FN rate of 9% with clinical follow-up for SLN biopsy negatives, reflecting pooled estimates from smaller studies and highlighting the importance of the learning curve effect. Gynaecological oncologists will value the clinical utility of knowing the FN rate of SLN in counselling patients.

It is also uncertain whether patients would rather risk groin metastases by forgoing IFL if they are SLN negative. One small study surveyed 106 patients who had undergone IFL as part of treatment; 66% would recommend IFL if the risk of missing metastasis from SLN biopsy was 1 in 80 and 84% would recommend IFL if the risk of missing metastasis from SLN biopsy was 1 in 8. Age and the presence or degree of side effects experienced by the patients surveyed, including 39% with severe lymphoedema and 28% with severe pain, did not affect preferences for each procedure (de Hullu et al, 2001). Further research on factors that influence lymphatic spread, for example, age, stage of disease and grade of tumour and exploration of patient's preferences, may aid decision making in the individual patient. Sophisticated quality-of-life studies need to investigate the impact of SLN vs IFL in patients.

At this stage, given the relatively small numbers of studies evaluating SLN with clinical follow-up, SLN should only be implemented within a research protocol, for unifocal tumours <4 cm, at selected centres with sufficient numbers and expertise to establish quality control. Careful patient counselling is essential, referring to the trade-off between morbidity from lymphadenectomy and the slightly higher rate of recurrence with SLN biopsy. Consensus standards for histopathological examination and reporting for SLN and IFL nodes are urgently needed, particularly with regard to ultrastaging and immunohistochemistry protocols. Given the higher recurrence rate in patients receiving SLN biopsy, we recommend that where patients have opted for SLN only and have not undergone full IFL, and the SLN is negative, they should be followed-up at close intervals (e.g., 2-monthly for 2 years) to detect any missed groin node metastases early, and facilitate an attempt at salvage therapy. In the absence of data to guide optimal method of follow-up (clinical vs imaging), careful clinical monitoring would seem pragmatic.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## REFERENCES

Achimas-Cadariu P, Harter P, Fisseler-Eckhoff A, Beutel B, Traut A, Du BA (2009) Assessment of the sentinel lymph node in patients with invasive squamous carcinoma of the vulva. Acta Obstet Gynecol Scand 88: 1209–1214.

Bass SS, Cox CE, Ku NN, Berman C, Reintgen DS (1999) The role of sentinel lymph node biopsy in breast cancer. *J Am Coll Surg* 189: 183–194.

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

Beesley V, Janda M, Eakin E, Obermair A, Battistutta D (2007) Lymphedema after gynecological cancer treatment: prevalence, correlates, and supportive care needs. *Cancer* 109(12): 2607–2614.

Bell JG, Lea JS, Reid GC (2000) Complete groin lymphadenectomy with preservation of the fascia lata in the treatment of vulvar carcinoma. *Gynecol Oncol* 77: 314–318.

Berek JS, Hacker NF (2005) (eds) *Practical Gynecologic Oncology*. 4th edn. pp 585–602. Williams & Wilkins: Philadelphia.

Brunner AH, Polterauer S, Tempfer C, Joura E, Reinthaller A, Horvat R, Hefler L (2008) The accuracy of intraoperative frozen section of the inguinal sentinel lymph node in vulvar cancer. *Anticancer Res* 28: 4091–4094.

Burger MP, Hollema H, Emanuels AG, Krans M, Pras E, Bouma J (1995) The importance of the groin node status for the survival of T1 and T2 vulval carcinoma patients. *Gynecol Oncol* 57: 327–334.

Camara O, Gonnert H, Herrmann J, Egbe A, Diebolder H, Gajda M, Michels W, Runnebaum IB (2009) Sentinel lymph node biopsy in vulvar cancer: a pilot study. *Eur J Gynaecol Oncol* 30: 622–624.

Crosbie EJ, Winter-Roach B, Sengupta P, Sikand KA, Carrington B, Murby B, Slade RJ (2010) The accuracy of the sentinel node procedure after excision biopsy in squamous cell carcinoma of the vulva. *Surg Oncol* 19: e150–e154.

de Cicco C, Sideri M, Bartolomei M, Grana C, Cremonesi M, Fiorenza M, Maggioni A, Bocciolone L, Mangioni C, Colombo N, Paganelli G (2000) Sentinel node biopsy in early vulvar cancer. *Br J Cancer* 82: 295–299.

de Hullu JA, Ansink AC, Tymstra T, van der Zee AG (2001) What doctors and patients think about false-negative sentinel lymph nodes in vulvar cancer. *J Psychosom Obstet Gynecol* 22: 199–203.

de Hullu JA, Hollema H, Piers DA, Verheijen RH, van Diest PJ, Mourits MJ, Aalders JG, van Der Zee AG (2000) Sentinel lymph node procedure is highly accurate in squamous cell carcinoma of the vulva. *J Clin Onc* 18: 2811–2816.

de Hullu JA, Van Der Avoort IA, Oonk MH, van der Zee AG (2006) Management of vulvar cancers. *Eur J Surg Oncol* 32: 825–831.

Diagnostic Test Accuracy Working Group (2012) Handbook for DTA reviews.

Gaarenstroom KN, Kenter GG, Trimbos JB, Agous I, Amant F, Peters AA, Vergote I (2003) Postoperative complications after vulvectomy and inguinofemoral lymphadenectomy using separate groin incisions. *Int J Gynecol Cancer* 13: 522–527.

Gould N, Kamelle S, Tillmanns T, Scribner D, Gold M, Walker J, Mannel R (2001) Predictors of complications after inguinal lymphadenectomy. *Gynecol Oncol* 82(2): 329–332.

Hampl M, Hantschmann P, Michels W, Hillemanns P (2008) Validation of the accuracy of the sentinel lymph node procedure in patients with vulvar cancer: results of a multicenter study in Germany. *Gynecol Oncol* 111: 282–288.

Hassanzade M, Attaran M, Treglia G, Yousefi Z, Sadeghi R (2013) Lymphatic mapping and sentinel node biopsy in squamous cell carcinoma of the vulva: systematic review and meta-analysis of the literature. *Gynecol Oncol* 130: 237–245.

Hauspy J, Beiner M, Harley I, Ehrlich L, Rasty G, Covens A (2007) Sentinel lymph node in vulvar cancer. *Cancer* 110: 1015–1023.

Higgins JPT, Green S. (eds) (2011) *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. Available from www.cochrane-handbook.org.

Johann S, Klaeser B, Krause T, Mueller MD (2008) Comparison of outcome and recurrence-free survival after sentinel lymph node biopsy and lymphadenectomy in vulvar cancer. *Gynecol Oncol* 110: 324–328.

Klar M, Bossart M, Stickeler E, Brink I, Orlowska-Volk M, Denschlag D (2011) Sentinel lymph node detection in patients with vulvar carcinoma; Feasibility of intra-operative mapping with technetium-99m-labeled nanocolloid. *Eur J Surg Oncol* 37: 818–823.

Klat J, Sevcik L, Simetka O, Graf P, Waloschek T, Kraft O, Jaluvkova Z, Prochazka M (2009) Characteristics of sentinel lymph nodes' metastatic involvement in early stage of vulvar cancer. *Aust N Z J Obstet Gynaecol* 49: 672–676.

Levenback C, Coleman RL, Burke TW, Bodurka-Bevers D, Wolf JK, Gershenson DM (2001) Intraoperative lymphatic mapping and sentinel node identification with blue dye in patients with vulvar cancer. *Gynecol Oncol* 83: 276–281.

Levenback CF, Ali S, Coleman RL, Gold MA, Fowler JM, Judson PL, Bell MC, de Geest K, Spirtos NM, Potkul RK, Leitao Jr MM, Bakkum-Gamez JN, Rossi EC, Lentz SS, Burke 2nd JJ, Van Le L, Trimble CL (2012) Lymphatic mapping and sentinel lymph node biopsy in women with squamous cell carcinoma of the vulva: a gynecologic oncology group study. *J Clin Oncol* 30: 3786–3791.

(2008) *Cochrane Handbook for Systematic Reviews of Interventions*. Wiley-Blackwell: Chichester.

Lindell G, Jonsson C, Ehrsson RJ, Jacobsson H, Danielsson KG, Källström BN, Larson B (2010) Evaluation of preoperative lymphoscintigraphy and sentinel node procedure in vulvar cancer. *Eur J Obstet Gynecol Reprod Biol* 152(1): 91–95.

Martinez-Palonez JM, Perez-Benavente MA, Gil-Moreno A, Diaz-Feijoo B, Roca I, Garcia-Jimenez A, Aguilar-Martínez I, Xercavins J (2006) Comparison of recurrence after vulvectomy and lymphadenectomy with and without sentinel node biopsy in early stage vulvar cancer. *Gynecol Oncol* 103: 865–870.

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

Merisio C, Berretta R, Gualdi M, Pultrone DC, Anfuso S, Agnese G, Aprile C, Mereu L, Salamano S, Tateo S, Melpignano M (2005) Radioguided sentinel lymph node detection in vulvar cancer. *Int J Gynecol Cancer* 15: 493–497.

Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol* 62: 1006–1012.

Moore RG, Robison K, Brown AK, Disilvestro P, Steinhoff M, Noto R, Brard L, Granai CO (2008) Isolated sentinel lymph node dissection with conservative management in patients with squamous cell carcinoma of the vulva: a prospective trial. *Gynecol Oncol* 109: 65–70.

Nyberg RH, Iivonen M, Parkkinen J, Kuoppala T, Maenpaa JU (2007) Sentinel node and vulvar cancer: a series of 47 patients. *Acta Obstet Gynecol Scand* 86: 615–619.

Oonk MH, van Hemel BM, Hollema H, de Hullu JA, Ansink AC, Vergote I, Verheijen RH, Maggioni A, Gaarenstroom KN, Baldwin PJ, van Dorst EB, Van DV, Hermans RH, van der Putten HW, Drouin P, Runnebaum IB, Sluiter WJ, van der Zee AG (2010) Size of sentinel-node metastasis and chances of non-sentinel-node involvement and survival in early stage vulvar cancer: results from GROINSS-V, a multicentre observational study. *Lancet Oncol* 11: 646–652.

Oonk MH, Van Os MA, de Bock GH, de Hullu JA, Ansink AC, van der Zee AG (2009) A comparison of quality of life between vulvar cancer patients after sentinel lymph node procedure only and inguinofemoral lymphadenectomy. *Gynecol Oncol* 113: 301–305.

Pereira de Godoy BN, Braile DM, de Fatima Godoy M, Longo J (2002) Quality of life and peripheral lymphoedema. *Lymphology* 35: 72–75.

Pitynski K, Basta A, Oplawski M, Przeszlakowski D, Hubalewska-Hola A, Krysztopowicz W (2003) [Lymph node mapping and sentinel node detection in carcinoma of the cervix, endometrium and vulva]. *Ginekol Pol* 74: 830–835.

Radziszewski J, Kowalewska M, Jedrzejczak T, Kozlowicz-Gudzinska I, Nasierowska-Guttmejer A, Bidzinski M, Siedlecki JA (2010) The accuracy of the sentinel lymph node concept in early stage squamous cell vulvar carcinoma. *Gynecol Oncol* 116: 473–477.

Rob L, Robova H, Pluta M, Strnad P, Kacirek J, Skapa P, Taborska K (2007) Further data on sentinel lymph node mapping in vulvar cancer by blue dye and radiocolloid Tc99. *Int J Gynecol Cancer* 17: 147–153.

Royal College of Obstetricians and Gynaecologists (1999) Management of vulval cancer.

Terada KY, Shimizu DM, Jiang CS, Wong JH (2006) Outcomes for patients with T1 squamous cell cancer of the vulva undergoing sentinel node biopsy. *Gynecol Oncol* 102: 200–203.

Vakselj A, Bebar S (2007) The role of sentinel lymph node detection in vulvar carcinoma and the experiences at the Institute of Oncology Ljubljana. *Radiol Oncol* 41: 167–173.

Van den Eynden J, Lannoo L, Amant F, Stroobants S, Vergote I (2003) [Sentinel node biopsy in the treatment of vulvar cancer]. *Tijdschr voor Geneeskunde* 59: 1169–1178.

van der Zee AG, Oonk MH, de Hullu JA, Ansink AC, Vergote I, Verheijen RH, Maggioni A, Gaarenstroom KN, Baldwin PJ, Van Dorst EB, Van DV, Hermans RH, Van Der Putten H, Drouin P, Schneider A, Sluiter WJ (2008) Sentinel node dissection is safe in the treatment of early-stage vulvar cancer. *J Clin Oncol* **26**: 884–889.

Vidal-Sicart S, Puig-Tintore LM, Lejarcegui JA, Paredes P, Ortega ML, Munoz A, Ordi J, Fuste P, Ortin J, Duch J, Martin F, Pons F (2007) Validation and application of the sentinel lymph node concept in malignant vulvar tumours. *Eur J Nucl Med Mol Imaging* **34**: 384–391.

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

Zekan J, Mutvar A, Huic D, Petrovic D, Karelovic D, Mitrovic L (2012) Reliability of sentinel node assay in vulvar cancer: the first Croatian validation trial. *Gynecol Oncol* **126**: 99–102.

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)