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

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

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

*Corresponding author

Declared competing interests of authors: none

Published December 2013
DOI: 10.3310/hta17600

Scientific summary

Sentinel lymph node status in vulval cancer
Health Technology Assessment 2013; Vol. 17: No. 60
DOI: 10.3310/hta17600

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

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

Objectives

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

Methods

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

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

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

Data sources

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

Results

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

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

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

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

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

Limitations

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

Conclusions

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

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

Implications for practice

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

Recommendations for further research

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

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

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

Funding

The National Institute for Health Research Health Technology Assessment programme.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Five-year impact factor: 5.804

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

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

Editorial contact: nihredit@southampton.ac.uk

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

Criteria for inclusion in the *Health Technology Assessment* journal

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

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

HTA programme

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

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

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

This report

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

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

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