Cost-effectiveness of diagnostic strategies for the management of abnormal uterine bleeding (heavy menstrual bleeding and post-menopausal bleeding): a decision analysis

Natalie A M Cooper,[1,2] Pelham M Barton,[3] Maria Breijer,[4] Orla Caffrey,[3] Brent C Opmeer,[5] Anne Timmermans,[4] Ben W J Mol,[4] Khalid S Khan[6] and T Justin Clark[2]*

[1]School of Clinical and Experimental Medicine, University of Birmingham, Birmingham, UK
[2]Birmingham Women's NHS Foundation Trust, Birmingham, UK
[3]Health Economics Unit, School of Health and Population Sciences, University of Birmingham, Birmingham, UK
[4]Academic Department of Obstetrics and Gynaecology, Academic Medical Centre, University of Amsterdam, Amsterdam, the Netherlands
[5]Clinical Research Unit, Academic Medical Centre, University of Amsterdam, Amsterdam, the Netherlands
[6]Centre for Primary Care and Public Health, Barts and The London School of Medicine and Dentistry, London, UK

*Corresponding author

Declared competing interests of authors: none

Published April 2014
DOI: 10.3310/hta18240

Scientific summary

Diagnostic strategies for the management of abnormal uterine bleeding

Health Technology Assessment 2014; Vol. 18: No. 24
DOI: 10.3310/hta18240

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

Scientific summary

Background

Abnormal uterine bleeding (AUB) affects women of all ages and is the commonest gynaecological presentation in secondary care, imposing substantial demands on health service resources. The condition manifests predominantly as heavy menstrual bleeding (HMB) in women of reproductive age and causes significant morbidity, distress and restrictions on activities of daily living. It is invariably of benign origin, but optimal treatment depends upon accurate diagnosis of the underlying pathology. Post-menopausal bleeding (PMB) describes any vaginal bleeding occurring after the menopause has been reached. While PMB is not usually heavy, it is an alarming, anxiety-provoking symptom, as it is caused by endometrial cancer in between 5% and 10% of cases. Therefore, prompt diagnostic work-up is necessary to make a timely diagnosis so that potentially curative treatment can be instigated in the presence of malignant disease and reassurance given if this life-threatening aetiology is excluded. While the implications of HMB and PMB are different, it is clear that accurate diagnosis is of key importance to realising effective treatment. However, in stark contrast with the treatment of pathologies associated with AUB, diagnostic work-up is ill-defined and practice is eclectic as there are many tests [transvaginal scan (TVS), saline infusion sonography (SIS), focal and global endometrial biopsy (EBx), outpatient hysteroscopy (OPH), etc.] and test combinations. This diverse, irrational approach to investigation is likely to compromise patient outcomes and waste scarce health-care resources through inefficiency.

Objectives

1. To determine the most cost-effective diagnostic testing strategy for the diagnosis and treatment of HMB.
2. To determine the most cost-effective diagnostic testing strategy for the diagnosis and treatment of PMB.

Data sources

Parameter inputs to populate the relevant branches of the decision trees included estimates of disease prevalence, diagnostic test accuracy, treatment effectiveness and associated costs. These data were derived from systematic quantitative reviews, individual patient data (IPD) from existing data sets and focused searches for specific data contingent with rigorous methodological appraisal to select the highest quality available data. In the absence of data estimates, the consensus view of an expert clinical panel was obtained. MEDLINE was searched from 1950 and EMBASE was searched from 1980. Both were searched up until the date of data identification, which ranged from December 2010 to January 2012.

Methods

Two clinically informed cost-effectiveness models were built as comprehensive decision trees to reflect current service provision for the diagnostic work-up of women presenting with HMB and PMB respectively. The trees were constructed to examine the effectiveness of different diagnostic testing strategies for women referred to secondary care by their general practitioner. As there is no consensus regarding how best to investigate women with AUB, the currently available and relevant outpatient tests were evaluated either alone or in combination. The tests included in both models were TVS, EBx and OPH. The testing modality SIS was also included in the HMB model.

A series of decision trees evaluating various testing strategies for HMB were developed to represent alternative decision options and their possible consequences. The trees explicitly illustrate the patient pathway from suspected pathology underlying the clinical presentation through to the outcome of testing, distinguishing between correct and incorrect diagnosis. Then, conditional on the accuracy of the diagnostic testing strategy, the outcome of treatment for HMB (satisfaction) was analysed at 1 year post initial presentation. A similar approach was used for the PMB model, but because the prime reason for investigation is to diagnose endometrial cancer, the outcome selected was 5-year survival. In addition, because a previous cost-effectiveness analysis found initial testing with TVS to be the preferred option for investigating PMB, we developed additional diagnostic strategies, using previously developed multivariable models based on patient characteristics, to see if this improved upon the previous optimal TVS strategy.

The model-based economic evaluation took the form of a cost-effectiveness analysis from the perspective of the NHS in a contemporary, 'one-stop' secondary care clinical setting, where all indicated testing modalities would be available during a single visit.

Deterministic results were obtained using point estimates of the parameters to estimate the expected cost, outcome and incremental cost-effectiveness. For the HMB model, the use of the levonorgestrel-releasing intrauterine system (LNG-IUS), recommended by the National Institute for Health and Care Excellence (NICE) as first-line treatment in HMB, was used as the reference-case scenario to compare all other testing strategies against. In the PMB model, no initial diagnostic work-up (investigation being restricted to representation) was used as the reference case. The stability of the results for both models was then tested through probabilistic sensitivity analysis and acceptability curves and frontiers drawn.

Results

Two potentially cost-effective testing strategies for the investigation of women with HMB were identified. These were initial testing with OPH alone or in combination with EBx. To adopt a strategy of OPH, £360 needs to be invested to gain one more woman satisfied at 1 year compared with a strategy of empirical treatment with a LNG-IUS. Although a testing strategy of OPH + EBx is marginally more effective, the ICER is approximately £21,000 to gain one more satisfied patient compared with OPH. We estimate that this equates to around £26,500 per quality-adjusted life-year (QALY), which falls within the £20,000–30,000 per QALY threshold. These findings were stable during sensitivity analyses, varying model inputs including disease prevalence, test feasibility and accuracy, with OPH remaining more cost-effective than the LNG-IUS reference strategy even at relatively low willingness-to-pay (WTP) thresholds. In women wishing to preserve their fertility, therapy with endometrial ablation and hysterectomy is contraindicated. SIS was cost-effective in this situation, with an incremental cost-effectiveness ratio (ICER) of approximately £2000, but for an additional financial outlay of £2720, testing instead with OPH produces a further satisfied patient, which is likely to be considered affordable and worthwhile by the NHS. At WTP thresholds of around £5000, there was a > 90% certainty of OPH being the most cost-effective option. In the scenario that only women who had HMB refractory to treatment with the LNG-IUS were referred to secondary care, OPH continued to be the most cost-effective option with an ICER of £5480 for each additional woman with HMB satisfied and increasing certainty of cost-effectiveness with increasing, but viable, WTP levels. Although the combination of TVS and EBx was a more effective approach in this latter situation, its ICER was over £500,000.

For the investigation of PMB, our analysis identified three potentially cost-effective testing approaches, all utilising TVS: (i) 'selective TVS with history', where TVS was restricted to women with a > 4% chance of endometrial cancer based upon risk factors identified from the patient history; (ii) 'history + TVS', where historical risk factors are taken into account along with the result of a TVS; and (iii) a combination of TVS and OPH. Selective TVS based upon historical risk prediction for the diagnostic work-up of women presenting with PMB generated an ICER compared with our reference strategy of 'no initial work-up' of £129,000 per extra woman surviving 5 years. Across the NICE threshold range of £20,000–30,000 per QALY gained, this option of selective TVS combined with risk factors acquired from the clinical history

would appear to be cost-effective if each additional 5-year survivor gains 4 to 6 QALYs. This seems plausible and suggests that this is the preferred option for investigating PMB at a cost acceptable to the NHS. The ICER of combining history and TVS was £2.4M, and the combination of OPH and TVS required an investment of £2.7M to acquire one additional woman with PMB living 5 years. Probabilistic sensitivity analyses showed that 'history + TVS' was cost-effective compared with 'selective TVS with history' in a small proportion of model runs. However, this proportion never exceeded 20% across a plausible range of WTP values from £0 to £600,000.

Limitations

Test accuracy and performance data were obtained from high-quality systematic quantitative reviews of the literature. Furthermore, in the PMB model, IPD integrated patient characteristics into the testing strategies to reflect the real clinical situation, although these data were available only for strategies utilising TVS. Other clinical parameter inputs, including treatment effectiveness and disease prevalence, were obtained following systematic searches and selected based upon a rigorous evaluation of data quality. We had hoped to use IPD derived from published systematic reviews of test accuracy to provide estimates of accuracy of tests used in combination. However, we adapted our approach when it became clear, following literature searches and scrutiny of the primary studies included in the available aggregated reviews, that reporting accuracy data in HMB or PMB for more than one test was rare. Tests in agreement presented no uncertainty, but where test results were modelled as being discordant, decision-making, as regards the assumed diagnosis or need for further testing, was determined by the consensus view of an expert clinical panel. The accord of this representative panel was key to informing the model inputs in other areas of practice where the evidence base was lacking or clinical decision-making was contentious: notably, the effectiveness of treatments for HMB in the presence of undetected pathologies (HMB model), and the effect of delayed diagnosis of endometrial cancer on disease progression and prognosis (PMB model).

While we aimed to develop economic models that accurately and explicitly mirrored clinical practice, some simplification was necessary, driven by a desire to keep the extensive and comprehensive decision trees workable; thus, subtle differences between strategies may have been overlooked. Moreover, the economic evaluation took the UK NHS as its perspective. This meant that only costs incurred by the NHS were included. Inevitably, this perspective will have excluded other potentially important costs and benefits. For example, the scarcity of health-related quality of life (HRQL) data in AUB precluded a cost–utility analysis. While we took account of the feasibility of testing, we did not consider the morbidity (anxiety, discomfort, complications) and psychological implications for women and their families of undergoing investigation. Similarly, we did not incorporate patient preferences for testing into the models or indeed the added value of individual tests outside the focus of uterine assessment, for example simultaneous assessment of the ovaries.

Costs were derived from up-to-date Healthcare Resource Group data, and these costs were added to the cost of a standard new consultation. This method allowed us to compare the different strategies fairly by breaking down the different aspects of each appointment. However, while this approach allowed us to apply costs in a contemporary, 'one stop' clinical setting, the costs may not accurately reflect the true cost to the NHS.

Conclusions

Implications for service provision

For the initial investigation of women presenting to secondary care with HMB who do not require preservation of their fertility, our research suggests a choice between OPH alone or a combination of OPH and EBx. From our investigation, OPH appears to be the optimal first-line diagnostic test used for the

investigation of women presenting to secondary care with HMB wishing to preserve their fertility or refractory to previous medical treatment with the LNG-IUS. We would suggest that the current recommendation of basing the initial investigation of women with PMB on the universal TVS measurement of endometrial thickness at a 5 mm threshold may need to be replaced by a strategy of restricting TVS to women with risk factors (e.g. increasing age-raised BMI, diabetes or nulliparity), obtained from the preceding clinical assessment.

Suggested research priorities

Future research should be aimed at generating estimates of diagnostic test accuracy of test combinations from IPD so that the added value of tests used in combination from the outset or in sequence, conditional on preceding test results, can be more rigorously estimated. In addition, systematic documentation of known and unknown but potential risk factors would allow further validation of the model and optimise the diagnostic algorithm. This would allow predictive models to be built and the formulation of rational, tailored testing strategies based upon individual risk assessment. These data can then inform future economic evaluations.

The completion of this extensive and comprehensive economic modelling evaluation incorporating all contemporary testing alternatives for HMB and PMB has delineated potential cost-effectives options. Thus, it would now be feasible to test the rigour of these findings in focused randomised trials should further confirmation be deemed necessary.

Validated, condition-specific, patient-reported outcome measures are available for HMB, but these have not as yet been widely applied as outcomes in clinical research. Future clinical trials should employ such measures to better capture the impact of HMB symptoms, and its investigation and management, on HRQL. In addition, the underlying causes of HMB should be defined so that response to investigation and treatment, in terms of HRQL, can be delineated. The management of PMB is aimed at identifying and treating a potentially life-threatening condition. However, while the detection of endometrial cancer is of over-riding importance, the majority of women with PMB have benign underlying causes. An evaluation of optimal testing approaches for all possible causes of PMB, benign and malignant, would require a relevant clinical outcome applicable uniformly. To facilitate this more comprehensive assessment requires further studies to capture HRQL data for the treatment of specific conditions underlying PMB.

Diagnostic assessment should incorporate patient preferences and take into account associated morbidity, including psychological implications. To understand these preferences, studies should be designed in which participants acquire a detailed knowledge of the accuracy and understand the added value of particular testing approaches as well as become familiar with the experience and risks of adverse effects. Awareness of patient's views will better inform the development of testing strategies, facilitate more individualised testing and maybe enhance compliance. This qualitative knowledge of the whole patient experience will be especially important where choice between potential cost-effective alternatives is contentious.

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 09/63/01. The contractual start date was in November 2009. The draft report began editorial review in February 2012 and was accepted for publication in March 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).