Health Technology Assessment 2011; Vol. 15: No. 19
ISSN 1366-5278

Hysterectomy, endometrial ablation and Mirena® for heavy menstrual bleeding: a systematic review of clinical effectiveness and cost-effectiveness analysis

S Bhattacharya, LJ Middleton, A Tsourapas, AJ Lee, R Champaneria, JP Daniels, T Roberts, NH Hilken, P Barton, R Gray, KS Khan, P Chien, P O'Donovan, KG Cooper and the International Heavy Menstrual Bleeding Individual Patient Data Meta-analysis Collaborative Group

April 2011
10.3310/hta15190

Health Technology Assessment
NIHR HTA programme
www.hta.ac.uk

How to obtain copies of this and other HTA programme reports

An electronic version of this title, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (www.hta.ac.uk). A fully searchable DVD is also available (see below).

Printed copies of HTA journal series issues cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our despatch agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per issue and for the rest of the world £3 per issue.

How to order:

– fax (with **credit card details**)
– post (with **credit card details** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you to either print out your order or download a blank order form.

Contact details are as follows:

Synergie UK (HTA Department)
Digital House, The Loddon Centre
Wade Road
Basingstoke
Hants RG24 8QW

Email: orders@hta.ac.uk

Tel: 0845 812 4000 – ask for 'HTA Payment Services' (out-of-hours answer-phone service)

Fax: 0845 812 4001 – put 'HTA Order' on the fax header

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *University of Southampton* and drawn on a bank with a UK address.

Paying by credit card
You can order using your credit card by phone, fax or post.

Subscriptions

NHS libraries can subscribe free of charge. Public libraries can subscribe at a reduced cost of £100 for each volume (normally comprising 40–50 titles). The commercial subscription rate is £400 per volume (addresses within the UK) and £600 per volume (addresses outside the UK). Please see our website for details. Subscriptions can be purchased only for the current or forthcoming volume.

How do I get a copy of *HTA on DVD*?

Please use the form on the HTA website (www.hta.ac.uk/htacd/index.shtml). *HTA on DVD* is currently free of charge worldwide.

The website also provides information about the HTA programme and lists the membership of the various committees.

Hysterectomy, endometrial ablation and Mirena® for heavy menstrual bleeding: a systematic review of clinical effectiveness and cost-effectiveness analysis

S Bhattacharya,[1]* LJ Middleton,[2] A Tsourapas,[2] AJ Lee,[1] R Champaneria,[2] JP Daniels,[2] T Roberts,[2] NH Hilken,[2] P Barton,[2] R Gray,[2] KS Khan,[2] P Chien,[3] P O'Donovan,[4] KG Cooper[5] and the International Heavy Menstrual Bleeding Individual Patient Data Meta-analysis Collaborative Group†

[1]University of Aberdeen, Aberdeen, UK
[2]University of Birmingham, Birmingham, UK
[3]Ninewells Hospital, Dundee, UK
[4]Bradford Royal Infirmary, Bradford, UK
[5]Aberdeen Royal Infirmary, Aberdeen, UK

*Corresponding author

†See *Acknowledgements* for full list of individuals

Declared competing interests of authors: S Bhattacharya, K Cooper and P O'Donovan were authors of papers included in the individual patient data meta-analysis. K Cooper received support from Ethicon and Microsubs in the past (not since 2008). P O'Donovan is a member of the Medical Advisory Board of Microsubs since 2009 to the present. P Chien previously attended overseas meetings paid for by Hologic UK; he is also a member of the Clinical Reference Group for the National Audit of Patient Outcomes and Experience of Treatment for Women with Heavy Menstrual Bleeding. P Chien has published several manuscripts on menorrhagia.

Published April 2011
DOI: 10.3310/hta15190

This report should be referenced as follows:

Bhattacharya S, Middleton LJ, Tsourapas A, Lee AJ, Champaneria R, Daniels JP, *et al*. Hysterectomy, endometrial ablation and Mirena® for heavy menstrual bleeding: a systematic review of clinical effectiveness and cost-effectiveness analysis. *Health Technol Assess* 2011;**15**(19).

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

NIHR Health Technology Assessment programme

The Health Technology Assessment (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 research findings from the HTA programme directly influence decision-making bodies such as the National Institute for Health and Clinical Excellence (NICE) and the National Screening Committee (NSC). HTA findings also help to improve the quality of clinical practice in the NHS indirectly in that they form a key component of the 'National Knowledge Service'.

The HTA programme is needs led in that it fills gaps in the evidence needed by the NHS. There are three routes to the start of projects.

First is the commissioned route. Suggestions for research are actively sought from people working in the NHS, from the public and consumer groups and from professional bodies such as royal colleges and NHS trusts. These suggestions are carefully prioritised by panels of independent experts (including NHS service users). The HTA programme then commissions the research by competitive tender.

Second, the HTA programme provides grants for clinical trials for researchers who identify research questions. These are assessed for importance to patients and the NHS, and scientific rigour.

Third, through its Technology Assessment Report (TAR) call-off contract, the HTA programme commissions bespoke reports, principally for NICE, but also for other policy-makers. TARs bring together evidence on the value of specific technologies.

Some HTA research projects, including TARs, may take only months, others need several years. They can cost from as little as £40,000 to over £1 million, and may involve synthesising existing evidence, undertaking a trial, or other research collecting new data to answer a research problem.

The final reports from HTA projects are peer reviewed by a number of independent expert referees before publication in the widely read journal series *Health Technology Assessment*.

Criteria for inclusion in the HTA journal series

Reports are published in the HTA journal series 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 referees 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.

The research reported in this issue of the journal was commissioned by the HTA programme as project number 05/45/02. The contractual start date was in September 2007. The draft report began editorial review in February 2010 and was accepted for publication in September 2010. As the funder, by devising a commissioning brief, the HTA programme specified the research question and study design. 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 referees 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.

The views expressed in this publication are those of the authors and not necessarily those of the HTA programme or the Department of Health.

ISSN 1366-5278 (Print)
ISSN 2046-4924 (Online)
ISSN 2046-4932 (DVD)

Published by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk), on behalf of NETSCC, HTA.
Printed on acid-free paper in the UK by the Charlesworth Group.

Abstract

Hysterectomy, endometrial ablation and Mirena® for heavy menstrual bleeding: a systematic review of clinical effectiveness and cost-effectiveness analysis

S Bhattacharya,[1]* LJ Middleton,[2] A Tsourapas,[2] AJ Lee,[1] R Champaneria,[2] JP Daniels,[2] T Roberts,[2] NH Hilken,[2] P Barton,[2] R Gray,[2] KS Khan,[2] P Chien,[3] P O'Donovan,[4] KG Cooper[5] and the International Heavy Menstrual Bleeding Individual Patient Data Meta-analysis Collaborative Group†

[1]University of Aberdeen, Aberdeen, UK
[2]University of Birmingham, Birmingham, UK
[3]Ninewells Hospital, Dundee, UK
[4]Bradford Royal Infirmary, Bradford, UK
[5]Aberdeen Royal Infirmary, Aberdeen, UK

*Corresponding author

†See *Acknowledgements* for full list of individuals

Objective: The aim of this project was to determine the clinical effectiveness and cost-effectiveness of hysterectomy, first- and second-generation endometrial ablation (EA), and Mirena® (Bayer Healthcare Pharmaceuticals, Pittsburgh, PA, USA) for the treatment of heavy menstrual bleeding.

Design: Individual patient data (IPD) meta-analysis of existing randomised controlled trials to determine the short- to medium-term effects of hysterectomy, EA and Mirena. A population-based retrospective cohort study based on record linkage to investigate the long-term effects of ablative techniques and hysterectomy in terms of failure rates and complications. Cost-effectiveness analysis of hysterectomy versus first- and second-generation ablative techniques and Mirena.

Setting: Data from women treated for heavy menstrual bleeding were obtained from national and international trials. Scottish national data were obtained from the Scottish Information Services Division.

Participants: Women who were undergoing treatment for heavy menstrual bleeding were included.

Interventions: Hysterectomy, first- and second-generation EA, and Mirena.

Main outcome measures: Satisfaction, recurrence of symptoms, further surgery and costs.

Results: Data from randomised trials indicated that at 12 months more women were dissatisfied with first-generation EA than hysterectomy [odds ratio (OR): 2.46, 95% confidence interval (CI) 1.54 to 3.93; $p = 0.0002$), but hospital stay [WMD (weighted mean difference) 3.0 days, 95% CI 2.9 to 3.1 days; $p < 0.00001$] and time to resumption of normal activities (WMD 5.2 days, 95% CI 4.7 to 5.7 days; $p < 0.00001$) were longer for hysterectomy. Unsatisfactory outcomes associated with first- and second-generation

techniques were comparable [12.2% (123/1006) vs 10.6% (110/1034); OR 1.20, 95% CI 0.88 to 1.62; $p = 0.2$). Rates of dissatisfaction with Mirena and second-generation EA were similar [18.1% (17/94) vs 22.5% (23/102); OR 0.76, 95% CI 0.38 to 1.53; $p = 0.4$]. Indirect estimates suggested that hysterectomy was also preferable to second-generation EA (OR 2.32, 95% CI 1.27 to 4.24; $p = 0.006$) in terms of patient dissatisfaction. The evidence to suggest that hysterectomy is preferable to Mirena was weaker (OR 2.22, 95% CI 0.94 to 5.29; $p = 0.07$). In women treated by EA or hysterectomy and followed up for a median [interquartile range (IQR)] duration of 6.2 (2.7–10.8) and 11.6 (7.9–14.8) years, respectively, 962/11,299 (8.5%) women originally treated by EA underwent further gynaecological surgery. While the risk of adnexal surgery was similar in both groups [adjusted hazards ratio 0.80 (95% CI 0.56 to 1.15)], women who had undergone ablation were less likely to need pelvic floor repair [adjusted hazards ratio 0.62 (95% CI 0.50 to 0.77)] and tension-free vaginal tape surgery for stress urinary incontinence [adjusted hazards ratio 0.55 (95% CI 0.41 to 0.74)]. Abdominal hysterectomy led to a lower chance of pelvic floor repair surgery [hazards ratio 0.54 (95% CI 0.45 to 0.64)] than vaginal hysterectomy. The incidence of endometrial cancer following EA was 0.02%. Hysterectomy was the most cost-effective treatment. It dominated first-generation EA and, although more expensive, produced more quality-adjusted life-years (QALYs) than second-generation EA and Mirena. The incremental cost-effectiveness ratios for hysterectomy compared with Mirena and hysterectomy compared with second-generation ablation were £1440 per additional QALY and £970 per additional QALY, respectively.

Conclusions: Despite longer hospital stay and time to resumption of normal activities, more women were satisfied after hysterectomy than after EA. The few data available suggest that Mirena is potentially cheaper and more effective than first-generation ablation techniques, with rates of satisfaction that are similar to second-generation techniques. Owing to a paucity of trials, there is limited evidence to suggest that hysterectomy is preferable to Mirena. The risk of pelvic floor surgery is higher in women treated by hysterectomy than by ablation. Although the most cost-effective strategy, hysterectomy may not be considered an initial option owing to its invasive nature and higher risk of complications. Future research should focus on evaluation of the clinical effectivesness and cost-effectiveness of the best second-generation EA technique under local anaesthetic versus Mirena and types of hysterectomy such as laparoscopic supracervical hysterectomy versus conventional hysterectomy and second-generation EA.

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

Contents

Glossary

Adenomyosis The presence of endometrium in the myometrium. Can cause heavy menstrual bleeding and pain.

Amenorrhoea Absence of periods.

Cervix The lower, narrower end of the uterus.

Cornua The horn-shaped top of the uterus leading to the fallopian tubes.

Cystometry A method for measuring the pressure–volume relationship of the bladder.

Diathermy Use of a high-frequency electrical current to produce heat that destroys tissues through cutting or electrocoagulation. The patient's body forms part of the circuit.

Dysmenorrhoea Painful periods.

Electrocautery Cauterisation of tissue using an electric current to generate the heat. Cauterisation destroys the tissue and causes scarring.

Endometriosis A condition where tissue resembling the endometrium occurs outside the uterus. The tissue responds to the menstrual cycle causing internal bleeding and pain.

Endometrium The inner lining of the uterus that thickens and sloughs off during the menstrual cycle.

Fibroids Benign smooth muscle tumours of the uterus.

Fundus The higher, wider end of the uterus.

Haematometra A collection of blood and other menstrual fluids in the uterus, which causes it to distend.

Hyperplasia The abnormal increase in the number of normal cells in a tissue.

Hysterectomy The surgical removal of the uterus; may include removal of the cervix.

Hegar A German gynaecologist who gave his name to a series of graduated cylindrical instruments used to dilate the cervix.

Hysteroscope An instrument using fibre optic technology that allows direct visualisation of the uterine cavity. Channels in the instrument allow it to be inserted to perform ablations.

Iatrogenic An adverse effect inadvertently induced through treatment.

Laparoscope A device used in surgery that allows visualisation through the use of fibre optics.

Leiomyomas Fibroids.

Menopause Cessation of menstruation, usually around age 50 years.

Menometrorrhagia Frequent, excessive menstrual bleeding.

Menorrhagia Heavy menstrual bleeding, clinically defined as more than 80 ml of blood per cycle, but more usually defined subjectively by the woman.

Menstruation The cyclic physiological discharge of blood and mucosal tissues through the vagina from the non-pregnant uterus. It is under hormonal control and recurs at approximately 4-week intervals.

Metrorrhagia Irregular, sometimes prolonged, menstrual bleeding.

Myometrium The outer muscular layer of the uterus.

Necrosis Cell death.

Oligomenorrhoea Few or scanty periods.

Pelvic inflammatory disease An inflammatory process that may be caused by sexually transmitted infection, ovarian cystic disease or infections after childbirth.

Perimenopausal Around the time of the menopause.

Polyp A mass of tissue on the mucosal lining, in this case in the uterus.

Post-ablation sterilisation syndrome In previously sterilised women, accumulation of the blood in the fallopian tubes, which may cause severe pelvic pain.

Pre-menstrual syndrome A combination of emotional and physical features that occur cyclically in women. May include mood changes, bloating, breast tenderness, fatigue and other symptoms.

Pyrexia Fever.

Salpingo-oophorectomy Surgical removal of the fallopian tubes and the ovaries.

Uterus The womb. A hollow, muscular pear-shaped organ in which the embryo is nourished.

List of abbreviations

AD	aggregate data
CEAF	cost-effectiveness acceptability frontier
CI	confidence interval
CUA	cost–utility analysis
DUB	dysfunctional uterine bleeding
ELA	endometrial laser ablation
EA	endometrial ablation
ED	endometrial destruction
EQ-5D	European Quality of Life-5 Dimensions
EVPI	expected value of perfect information
GA	general anaesthetic
GI	gastrointestinal
GnRH	gonadotrophin-releasing hormone
GP	general practitioner
HMB	heavy menstrual bleeding
HA	hydrothermablator
ICD	*International Classification of Diseases*
ICD-9	*International Classification of Diseases*, Ninth Edition
ICD-10	*International Classification of Diseases*, Tenth Edition
ICER	incremental cost-effectiveness ratio
IPD	individual patient data
IQR	interquartile range
ISD	Information Services Division
ITT	intention to treat
IUD	intrauterine device
IUS	intrauterine system
LA	local anaesthetic
LNG	levonorgestrel
MEA	microwave endometrial ablation
NICE	National Institute for Health and Clinical Excellence
NNT	number needed to treat
NSAID	non-steroidal anti-inflammatory drug
OR	odds ratio
PBAC	pictorial blood loss assessment chart
PICOS	participants, interventions, comparisons, outcomes and study design
PSA	probabilistic sensitivity analysis
QALY	quality-adjusted life-year
QoL	quality of life
RB	rollerball
RBEA	rollerball endometrial ablation
RCOG	Royal College of Obstetricians and Gynaecologists
RCT	randomised controlled trial
RR	relative risk
SD	standard deviation
SE	standard error
SF-36	Short Form questionnaire-36 items
SMD	standardised mean difference
SMR	Scottish Morbidity Returns

TBEA	thermal balloon endometrial ablation
TCRE	transcervical resection of the endometrium
TVT	tension-free vaginal tape
UTI	urinary tract infection
VAS	visual analogue scale
WMD	weighted mean difference

All abbreviations that have been used in this report are listed here unless the abbreviation is well known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices, in which case the abbreviation is defined in the figure legend or at the end of the table.

Executive summary

Background

Heavy menstrual bleeding (HMB) is a common problem that affects approximately 1.5 million women in England and Wales and accounts for 20% of gynaecology outpatient referrals. Although objectively defined as cyclical loss of > 80 ml of blood during each menstrual period, HMB is diagnosed clinically in the presence of excessive menstrual blood loss that interferes with a woman's physical, emotional, social and material quality of life.

Medical treatments for HMB include oral drug regimens, such as tranexamic acid and mefenamic acid, and the combined oral contraceptive pill as well as the levonorgestrel intrauterine system (LNG IUS) (Mirena®, Bayer Healthcare Pharmaceuticals, Pittsburg, PA, USA), which can reduce menstrual loss by local release of progestogen. Surgical treatments include first- (hysteroscopic) and second- (non-hysteroscopic) generation endometrial ablation (EA), which destroys the lining of the cavity of the uterus (endometrium), and hysterectomy (surgical removal of the uterus). First-generation ablation techniques include endometrial laser ablation, transcervical resection of the endometrium and rollerball (RB) ablation. Examples of second-generation ablative techniques are fluid-filled thermal balloon endometrial ablation, radiofrequency (thermoregulated) balloon endometrial ablation, hydrothermal endometrial ablation, microwave EA (MEA) and impedance-controlled bipolar radiofrequency ablation (NovaSure®; Hologic Inc., Bedford, MA, USA).

In 1999–2000, half of the 51,858 hysterectomies performed in the public sector in England were for HMB. In contrast, 7179 hysterectomies were performed for HMB in 2004–5 while 9701 women underwent EA – over half of these (5457) by means of second-generation (non-hysteroscopic) techniques. The use of Mirena has increased concurrently, although its widespread use for contraception across a number of clinical settings in primary and secondary care means that it is difficult to gather accurate data on numbers prescribed for HMB.

Objective

The aim of this project was to determine the clinical effectiveness and cost-effectiveness of hysterectomy, first- and second-generation EA, and Mirena for the treatment of HMB. To address this question, the specific objectives were:

1. To determine, using individual patient data (IPD) meta-analysis of existing randomised controlled trials (RCTs), the short- to medium-term effects of each class of treatment in terms of patient dissatisfaction, time to resumption of normal activities and complication rate, and to explore these outcomes in clinical subgroups.
2. To report, using population-based data from record linkage, the long-term effects of ablative techniques and hysterectomy in terms of failure rates and complications.
3. To inform current treatment policy in this clinical area, while the value of information component serves to highlight future research needs and agendas, and inform possible future research funding decisions.

Design

Systematic review and individual patient data meta-analysis of available evidence

A detailed search was carried out to identify systematic reviews and RCTs involving hysterectomy, EA and Mirena. IPD were sought from RCTs of hysterectomy, EA techniques and Mirena to examine their relative effectiveness. A systematic review was conducted based on a protocol designed using widely recommended methods that complied with meta-analysis reporting guidelines.

Individual patient data on 2814 women were available from 17 of the 30 RCTs identified (14 trials including 2448 women for first- vs second-generation EA; seven trials including 1127 women for hysterectomy vs first-generation EA; five trials including 304 women for second-generation EA vs Mirena; three trials including 190 women for first-generation EA vs Mirena; one trial including 236 women for hysterectomy vs Mirena). Direct and indirect comparisons were made where appropriate to assess the effect of interventions on the primary outcome measure of patient dissatisfaction.

Follow-up of women following hysterectomy and endometrial ablation by record linkage

Patient-based data for inpatient and day case activity from the whole of Scotland which are routinely collected as Scottish Morbidity Returns (SMR) by the Scottish Information Services Division (ISD) were used for this study. Following linkage with the Scottish Cancer Registry, an anonymised data set containing follow-up hospital data on all women who had undergone either hysterectomy or EA for HMB between 1989 and 2006 was made available to the researchers. Socioeconomic status was assessed using the Carstairs index, which was divided into quintiles for analysis. Descriptive statistics were used to summarise each of the surgical outcomes and potential predictor variables (age, year of procedure and Carstairs quintile). Appropriate univariate analyses across the hysterectomy and EA groups were performed. Cox proportional hazards regression analysis was used to examine the survival experience for different surgical outcomes in the hysterectomy and EA groups and then between different types of hysterectomy following adjustment for age, year of primary operation and Carstairs quintile.

Cost-effectiveness evaluation

The authors developed a state transition (Markov) model using Microsoft EXCEL (Microsoft Corporation, Seattle, WA, USA). The structure was informed by the review of the clinical literature supplemented by clinical input. The model allows a comparison of four hypothetical cohorts of women with HMB who are treated separately by one of four alternative strategies: (1) Mirena coil; (2) first-generation EA techniques; (3) second-generation EA techniques; and (4) hysterectomy. Given the reliance on secondary data and the availability of data, the model-based economic evaluation takes the form of a cost–utility analysis and was carried out from the perspective of the UK NHS in a secondary care setting. The results are reported in terms of incremental cost per quality-adjusted life-year (QALY) gained based on quality of life data available from published sources. The presentation of results in QALYs allows comparison of the results with other available and recently published studies [Garside R, Stein K, Wyatt K, Round A, Price A. The effectiveness and cost-effectiveness of microwave and thermal balloon endometrial ablation for heavy menstrual bleeding: a systematic review and economic modelling. *Health Technol Assess* 2004;**8**(3)]. Resource use was estimated from the existing published evidence and additional cost data from other sources such as the annual review of unit health and social care costs (Personal Social Services Research Unit) and national schedule for reference costs.

Results

Clinical effectiveness from individual patient data meta-analysis

At around 12 months, 7.3% more women [12.6% (57/454) vs 5.3% (23/432)] were dissatisfied with the outcome of first-generation EA than with hysterectomy [OR (odds ratio) 2.46, 95% confidence interval (CI) 1.54 to 3.93; $p = 0.0002$], but hospital stay [WMD (weighted mean difference) 3.0 days, 95% CI 2.9 to 3.1 days; $p < 0.00001$] and time to resumption of normal activities (WMD 5.2 days, 95% CI 4.7 to 5.7 days; $p < 0.00001$) were longer for hysterectomy. Unsatisfactory outcomes were comparable with first- and second-generation EA techniques [12.2% (123/1006) vs 10.6% (110/1034); OR 1.20, 95% CI 0.88 to 1.62; $p = 0.2$], although second-generation techniques were quicker (WMD 14.5 minutes, 95% CI 13.7 to 15.3 minutes; $p < 0.00001$) and women recovered sooner (WMD 0.48 days, 95% CI 0.20 to 0.75 days; $p = 0.0008$) with fewer procedural complications. Indirect comparison suggested more unsatisfactory outcomes with second-generation EA techniques than with hysterectomy [10.6% (110/1034) vs 5.3% (23/432); OR 2.32, 95% CI 1.27 to 4.24; $p = 0.006$].

Rates of dissatisfaction with Mirena and second-generation EA were similar [18.1% (17/94) vs 22.5% (23/102); OR 0.76, 95% CI 0.38 to 1.53; $p = 0.4$]. Overall rates of dissatisfaction were 17.2% (22/128) for Mirena and 18.2% (25/137) for both first- and second-generation EA. Lack of IPD prohibited any further investigation of subgroups or repeated measures.

Indirect estimates suggest that hysterectomy is also preferable to second-generation EA (OR 2.32, 95% CI 1.27 to 4.24; $p = 0.006$) in terms of patient dissatisfaction. This is confirmed by the repeated measures analysis over all three time points, which only include IPD (OR 3.06, 95% CI 1.59 to 5.90; $p = 0.0008$). The evidence to suggest that hysterectomy is preferable to Mirena was weaker (OR 2.22, 95% CI 0.94 to 5.29; $p = 0.07$), but given the lack of precision from Mirena comparisons this was not a surprising result.

Medium- to long-term surgical outcomes following endometrial ablation and hysterectomy for heavy menstrual bleeding

Between 1989 and 2006, 37,120 Scottish women underwent hysterectomy and 11,299 had EA as a primary surgical procedure for HMB. The median [interquartile range (IQR)] duration of follow-up was 6.2 (2.7–10.8) and 11.6 (7.9–14.8) years, respectively, in the EA and hysterectomy cohorts.

A total of 2779 women in the original EA group went on to have a hysterectomy and were excluded from further analysis.

Of the remaining women originally treated by EA, 962 (8.5%) underwent further gynaecological surgery. While the risk of adnexal surgery was similar in both groups, women who had undergone hysterectomy were more likely to need further surgery for stress urinary incontinence. Vaginal hysterectomy was associated with a significantly higher chance of further surgery for urinary incontinence and pelvic floor repair than hysterectomy carried out through the abdominal route. The incidence of endometrial cancer following endometrial ablation was low at 0.02%.

Cost-effectiveness

The results of the cost-effectiveness model show that the strategy of hysterectomy is the most cost-effective. Hysterectomy dominates the first-generation EA strategy and, although more expensive, produces more QALYs than the other strategies of second-generation EA and

Mirena. The incremental cost-effectiveness ratios for hysterectomy compared with Mirena and hysterectomy compared with second-generation ablation are £1440 per additional QALY and £970 per additional QALY, respectively.

Discussion

Strengths and limitations of the analysis

For the systematic review, an extensive literature search was conducted, with no language restrictions, minimising the risk of missing information.

A limitation of the systematic review was the unavailability of IPD from at least 35% of randomised women, which could not be accessed as a number of triallists did not agree to collaborate or could not be contacted. Received data were sometimes incomplete and, on occasions, failed quality checks, and so were unusable. The review's inferences are also limited by the inconsistent outcome measure used across trials; studies involving endometrial destruction (ED) and Mirena focused on comparing reduction in bleeding, while hysterectomy trials focused on patient satisfaction and quality.

The follow-up study on women who had undergone hysterectomy or EA is, to our knowledge, the first large population-based study to use national data. Use of the *International Classification of Diseases* codes allowed us to define both the cause of HMB as well as the nature of surgery, but, as the diagnosis of dysfunctional uterine bleeding was performed by a process of exclusion, it is possible that the hysterectomy cohort could have included a few women with other causes of HMB. As a retrospective observational study, it is not free from problems of bias and confounding. The analysis was compromised by the limited availability of key socioeconomic as well as clinical variables. Although the numbers of women in the hysterectomy and ablation cohorts were large, a major drawback was our inability to discriminate between the individual types of first- and second-generation EA or adjust for the experience of the operator as has been done in previous national audits. We were also unable to analyse the long-term outcomes following laparoscopic hysterectomy as numbers were small and these were grouped with abdominal hysterectomy.

The major strength of the economic component of this study is that it was based on a state-of-the-art Markov model which was informed by data from an IPD meta-analysis of randomised trials. A multidisciplinary team including economists, expert clinicians and statisticians provided input into the model structure, primarily based on the evidence in the literature. All assumptions used in the model were made a priori, and were based on the best available evidence.

The quality of the health economic model was affected by the paucity of good-quality data such as those related to adverse outcomes following some types of EA and follow-up data on Mirena use. In addition, the complexity of the model meant a long running time, which inevitably affected the number and nature of additional sensitivity analyses undertaken.

Interpretation of available evidence and consensus regarding treatment

More women were dissatisfied following EA than hysterectomy. However, dissatisfaction rates were low after all treatments and hysterectomy was associated with an increased hospital stay and recovery period. The paucity of suitable trials means that definitive evidence on the effectiveness of Mirena compared with more invasive procedures is lacking. Hysterectomy would be considered the most cost-effective strategy in the light of the acceptable thresholds used by the National Institute for Health and Clinical Excellence (NICE). The results concur with those of other studies, but are sensitive to utility values used in the analysis.

A summary of the results on the clinical effectiveness and cost-effectiveness of Mirena, EA and hysterectomy was sent electronically to 15 national experts (gynaecological surgeons) along with a short questionnaire to encourage rapid response. After two mailings, responses were received from 10 clinicians, 9 of whom indicated that having considered effectiveness, cost-effectiveness and invasiveness/risks they would favour HMB LNG IUS (Mirena), second-generation EA techniques and hysterectomy as first-, second- and third-line approaches to HMB resistant to oral medication. This view was endorsed by three consumers who highlighted the need for a degree of flexibility in order to accommodate the preferences of individual women.

Conclusion

Although hospital stay and time to resumption of normal activities were longer, more women were satisfied after hysterectomy than after first-generation EA. In the absence of head-to-head trials, indirect estimates suggest that hysterectomy is also preferable to second-generation EA in terms of patient satisfaction. Dissatisfaction rates were comparable between first- and second-generation techniques, although second-generation techniques were cheaper, quicker and associated with faster recovery and fewer complications. There are few comparisons of Mirena with more invasive procedures.

The few data available suggest that Mirena is potentially cheaper and more effective than first-generation ablation techniques with rates of satisfaction that are similar to those of second-generation techniques. Owing to a paucity of trials, there is limited evidence to suggest that hysterectomy is preferable to Mirena. Hysterectomy is considered the most cost-effective strategy, but, owing to its invasive nature and higher risk of complications, is considered a final option by gynaecological experts and consumers.

Implications for service provision

Our review provides evidence that hysterectomy reduces dissatisfaction compared with EA, and this information could contribute to a consultation with women making a choice about treatment options when initial drug treatment fails to control HMB. EA is satisfactory for a very high proportion of women, but, if complete cessation of bleeding is sought, then hysterectomy may be offered. A decision to opt for hysterectomy needs also to take into account the invasive nature of the procedure and its potential for short- and long-term morbidity in some women.

Although conclusive evidence from randomised trials is still awaited, the evidence from our review is consistent with a recent NICE recommendation that women should be offered Mirena before more invasive procedures. This view reflects the minimally invasive nature of the intervention as well as the ability to offer it in primary care. This piece of research has highlighted the benefits and risks associated with the three broad strategies for the treatment of HMB and, while supportive of the existing NICE guideline on this subject, our results underline the need for a degree of flexibility in accommodating women's preferences.

Need for further research

This project has uncovered a number of areas for future research. These include:

- evaluation of the clinical effectiveness and cost-effectiveness of the best second-generation EA technique under local anaesthetic versus Mirena
- exploring the safety of second-generation EA and Mirena through a national audit
- longer term follow-up of randomised cohorts of women treated for HMB

- evaluation of the clinical effectiveness and cost-effectiveness of hydrothermablator (HA, the second-generation EA device which can be used under direct vision) against other second-generation techniques
- trials assessing conservative and less morbid types of hysterectomy such as laparoscopic supracervical hysterectomy versus conventional hysterectomy and second-generation EA.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Chapter 1

Background

Heavy menstrual bleeding (HMB) is a common problem[1] which affects approximately 1.5 million women in England and Wales.[2] The condition causes 1 in 20 women of reproductive age to consult their general practitioners (GPs) and accounts for 20% of gynaecology outpatient referrals. HMB can cause significant distress to women by affecting their performance at work as well as their social activities, and leads to a measurable reduction in quality of life (QoL).[3] Surgery has been traditionally used as the definitive treatment of HMB such that, in the past, by the age of 55 years one in five women in the UK had a hysterectomy,[4] over half of which were for HMB.[5]

Definition of heavy menstrual bleeding

Although objectively defined as the loss of > 80 ml of blood per cycle,[6] such measurement is impractical in most clinical settings. Between 35% and 60% of women who present with a subjective complaint of HMB have been shown to have normal levels of blood loss.[7,8] Conversely, many women with objectively demonstrable high blood loss do not seek help for associated symptoms.[9]

Of various methods used to measure menstrual blood loss, the alkaline haematin technique has been considered to be the gold standard.[10] Despite the introduction of modifications in an effort to simplify it,[11] this method remains laborious and involves extraction of haemoglobin from used sanitary wear. As such, it is unsuitable for regular clinical use.

A more practical method of assessing menstrual blood loss is the pictorial blood loss assessment chart (PBAC).[12] This takes into account the number of items of sanitary wear used and the degree of staining, which are in turn converted into a score. This technique is now more widely used than the alkaline haematin method although the correlation between actual measured blood loss and the PBAC score has been questioned.[13] Another indirect method for estimating menstrual blood loss is the 'menstrual pictogram',[14] which is similar to the PBAC but additionally requires women to comment on the absorbency of the towel or tampon and any extraneous blood loss.

From a clinical perspective, HMB is defined as excessive menstrual blood loss which interferes with a woman's physical, emotional, social and material QoL, and which can occur alone or in combination with other symptoms.[15]

Causes of heavy menstrual bleeding

Possible causes of HMB are shown in *Table 1*. In most cases a definite cause is not found and the condition is labelled as dysfunctional uterine bleeding (DUB).[17]

TABLE 1 Causes of HMB (from Garside *et al*.[16])

Anatomical	Biochemical	Endocrine	Haematological	Iatrogenic	Associated factors
Fibroids	Prostaglandins	Hypothalamic–pituitary–gonadal–adrenal axis dysfunction	Von Willebrand's disease	IUDs	Obesity
Polyps		Oestrogen-producing tumours		Anticoagulants	Heavy smoking
Adenomyosis		Thyroid dysfunction	Leukaemia	Exogenous hormones	Excessive alcohol
Infection			Increased endometrial fibrinolytic activity		Depression
Malignancies					Endometriosis

IUD, intrauterine device.

Estimating the severity of heavy menstrual bleeding

Subjective estimates of menstrual blood loss do not correlate well with objective measures,[13,18] and over half of women who have surgery for HMB do not experience a blood loss of 80 ml or more in each cycle.[8] Women's expectations of normal menstrual loss can shape their perception of the gravity of their condition, inform their demand for treatment and influence their judgement about treatment success.

The presence of other menstrual symptoms may also have an impact on perceptions of bleeding and account for some of the differences between objective and subjective estimates of menorrhagia. Thus, many women presenting with HMB describe other additional symptoms such as painful periods while associated symptoms are more likely to encourage a diagnosis of HMB by clinicians. The impact of HMB is conventionally measured by means of a number of QoL measures. A systematic review of QoL measures in HMB described 15 generic and two condition-specific scales[19] and suggested that there was scope for better ways of assessing the severity of the condition and its impact on women's lives.

Generic scales allow comparison between different clinical conditions in terms of their impact on QoL and may provide a single score or scores across dimensions of QoL, but are relatively insensitive to specifics of a particular condition. Generic measures of QoL used in HMB include the Short Form questionnaire-36 items (SF-36), Nottingham Health Profile, health-status structured history and single global item.[16] SF-36 is generally a well-validated measure used to assess health-related QoL[19] and includes items on global health perception, physical function, social function, role (physical and mental), pain, mental health and energy/vitality.[20] The SF-36 has been considered to be a feasible way of assessing QoL in women with HMB, but it has some limitations in this setting[21] as some questions can be inappropriate for these women. In addition, internal reliability, as assessed by Cronbach's statistic, has been shown to be lower in women with HMB, especially for general health perception and mental health scales.

Clark *et al.*[19] have also reported on the use of generic measures that address particular aspects of QoL such as physical (Modified Townsend Score), mental (General Health Questionnaire) and sexual health (Revised Sabbatsberg Sexual Rating Scale) and social function (Lifestyle Index) in studies of women with HMB bleeding.

Condition-specific scales have the advantage of incorporating attributes of QoL that are specifically affected by the condition of interest. They may therefore be more sensitive to small but important

changes and may be considered to have greater face validity (that is, they include items that are of importance to sufferers and reflect their experience and concerns). Two condition-specific outcome measures have been developed for women with HMB. These include the Menorrhagia Outcomes Questionnaire[22] and the Multi-attribute Questionnaire.[23] The Menorrhagia Outcomes Questionnaire includes items on symptoms and satisfaction with care, physical function, psychological and social well-being, global judgement of health and QoL, and personal constructs. The Multi-attribute Questionnaire includes items on practical difficulties, social function, psychological function, physical health, interruption to work and family life.

Preference-based measures elicit preferences for a given health state and, if appropriately scaled, provide weights that can be used in cost–utility analyses.

The EQ-5D™ (European Quality of Life-5 Dimensions; Euro Qol Group, Rotterdam, the Netherlands), which has been used in studies as a measure of QoL in HMB, includes a multi-attribute scale, with dimensions of mobility, self-care, usual activities, pain/discomfort and anxiety/depression, and a global rating scale for QoL (visual analogue scale).

Measuring patient satisfaction

Patient satisfaction is widely used as a primary outcome measure in studies of treatments for HMB.[24] Satisfaction is a subjective and relative concept and represents the extent to which a service meets users' expectations. It is not clear whether satisfaction can be measured on a continuum, from dissatisfied through to satisfied, or whether factors resulting in satisfaction are different from those leading to dissatisfaction.[16] Satisfaction is influenced by patient characteristics[24] such as age and health status.

The extent to which these potential biases are addressed in the patient satisfaction measures used in studies of HMB is difficult to assess in the absence of detailed accounts of the development and validation of the measures used. While the use of a similar tool to measure subjective satisfaction for women in both arms of an RCT may provide a comparative measure between these groups, it may remain unclear exactly what is being measured for the reasons outlined above.[16] In addition, the range of techniques and scales used to elicit a measure of satisfaction across studies can limit any attempts to aggregate data by means of meta-analysis.

Current service provision

Treatment for HMB aims to improve women's quality of life through reducing menstrual loss. The current National Institute for Health and Clinical Excellence (NICE) guideline advocates full gynaecological examination followed by appropriate tests such as a full blood count and recognises the need for endometrial biopsy, ultrasound scan and hysteroscopy in specific cases.[15]

Medical therapy

According to the recent NICE guideline on HMB,[15] medical treatment should be considered where structural and histological abnormalities of the uterus have been excluded or for fibroids < 3 cm in diameter which do not appear to distort the cavity of the uterus. In addition, the contraceptive needs of the woman should be taken into consideration. In addition to being licensed as a contraceptive device, the levonorgestrel-releasing intrauterine system (LNG IUS or Mirena®, Bayer Healthcare Pharmaceuticals, Pittsburg, PA, USA) is an effective non-surgical

treatment for HMB which is reversible and fertility sparing. The device, which has to be fitted by a qualified practitioner, has a T-shaped plastic frame and a rate-limiting membrane on the vertical stem which releases a daily dose of 20 µg of LNG. The effects of the LNG IUS are local and hormonal, including prevention of endometrial proliferation, thickening of cervical mucous and suppression of ovulation in a minority of women. It reduces estimated menstrual blood loss by up to 96% by 12 months, with up to 44% of users reporting amenorrhoea,[25,26] at a cost which is a third of that for hysterectomy.[27] It has been recommended that LNG IUS should be considered before oral medication such as tranexamic acid, non-steroidal anti-inflammatory drugs (NSAIDs) or combined oral contraceptives.[15] Mirena can lead to troublesome spotting in some women, causing early discontinuation of the device. There are relatively few randomised trials comparing the relative effectiveness of LNG IUS with that of hysterectomy, as well as endometrial ablation (EA), or long-term follow-up data on Mirena use.

Surgical treatment

Despite the availability of a number of medical options, long-term medical treatment is unsuccessful or unacceptable in many cases and surgical alternatives such as EA techniques and hysterectomy may be required.[28] In a randomised controlled trial (RCT) of medical management versus transcervical resection of the endometrium (TCRE) in secondary care, total satisfaction with treatment was higher in women who were treated surgically (39% vs 61%; $p = 0.01$).[28] The current NICE guideline on HMB suggests that EA may be offered to women who do not desire future fertility and in whom bleeding is considered to have a major impact on QoL. The guideline development group felt that ablative surgery could be offered as the initial surgical treatment for HMB after full discussion about the risks and benefits of other options.[15]

Incidence of surgical operations for heavy menstrual bleeding

Of 51,858 hysterectomies in the public sector in England in 1999–2000, it is likely that half were for HMB.[29] Between 1999–2000 and 2004–5, 6500 fewer hysterectomies were performed.[30] In contrast there were 826 hysteroscopic EAs in England in 1989, rising to 7173 in 1992–3, before falling to 3847 in 1997–8. In 2004–5, 9701 EAs were performed, of which over half (5457) used second-generation (non-hysteroscopic directed) methods.[30] With just 7179 hysterectomies performed for HMB over this period, the predominant operation for HMB was now ablation. The use of LNG IUS has increased concurrently, although the widespread use of this device for contraception as well as for the control of HMB across a number of clinical settings (primary care, sexual reproductive health as well as gynaecology clinics in secondary care) makes it difficult to gather accurate data on uptake rates.

Hysterectomy

Hysterectomy is defined as the surgical removal of the uterus. It offers a definitive treatment for menorrhagia and guarantees amenorrhoea, but is particularly invasive and carries risk of significant morbidity.[31] Hysterectomy can be performed through a number of routes. In abdominal hysterectomy the uterus is approached through the anterior abdominal wall, while vaginal hysterectomy involves surgical removal of the uterus through the vagina. In laparoscopic hysterectomy surgery is accomplished without the need for a laparotomy. Laparoscopic hysterectomy includes three subtypes: (1) laparoscopically assisted vaginal hysterectomy; (2) laparoscopic hysterectomy; and (3) total laparoscopic hysterectomy. In addition, laparoscopically assisted supracervical hysterectomy involves removal of the body of the uterus while the cervix is retained.

Hysterectomy can also be categorised on the basis of the extent of the operation and organs removed. Removal of the uterus and cervix constitutes total hysterectomy, while excision of the

body of the uterus while conserving the cervix is defined as subtotal hysterectomy. Removal of the uterus alone is conventionally known as simple hysterectomy, while additional removal of the fallopian tubes and ovaries or ovaries alone is referred to as salpingo-oophorectomy or oophorectomy, respectively. Oophorectomy is usually performed in the presence of ovarian pathology but can also be carried out prophylactically to avoid the risk of cancer. Removal of the ovaries in cases of HMB is incidental.[15] Of 37,000 hysterectomies performed in the UK in 1994–5, two-thirds were abdominal (4% of these were subtotal hysterectomies) and 57% were accompanied by removal of tubes and ovaries.[5]

Hysterectomy is generally performed as an inpatient procedure. The need for general anaesthesia, prolonged hospital stay and delayed recovery makes it a potentially expensive treatment.[32] Overall, 1 in 30 women suffers a major adverse event after hysterectomy, and the mortality rate is 0.4–1.1 per 1000 operations. Around 3% of women suffer from perioperative problems such as haemorrhage, trauma to other pelvic organs and anaesthetic problems. Immediate postoperative complications include sepsis, bleeding and venous thromboembolism. Adverse events following hysterectomy are summarised in *Table 2*. Although delayed complications including urinary incontinence, fatigue, pelvic pain, hot flushes and sexual problems have been reported,[5,33–35] satisfaction rates following hysterectomy are very high.[31]

Endometrial ablation

Endometrial ablative techniques, which aim to destroy functionally active endometrium along with some underlying myometrium,[36,37] offer a conservative surgical alternative to hysterectomy. The first-generation ablative techniques including endometrial laser ablation (ELA),[38,39] TCRE[40] and rollerball endometrial ablation (RBEA) were all endoscopic procedures. Although none guarantees amenorrhoea, their effectiveness (in comparison with hysterectomy) has been demonstrated in a number of RCTs.[41–46]

National audits[47,48] revealed that, although first-generation ablative techniques were less morbid than hysterectomy, they were associated with a number of complications, including uterine perforation, cervical laceration, false passage creation, haemorrhage, sepsis and bowel injury. In addition, fluid overload associated with the use of 1.5% urological glycine (non-ionic) irrigation fluid in TCRE and RBEA resulted in serious and occasionally fatal consequences due to hyponatraemia.[49,50] Mortality from these techniques has been estimated at 0.26 per 1000.[47,48]

Second-generation ablative techniques represent simpler, quicker and potentially more efficient means of treating menorrhagia, which require less skill on the part of the operator. Examples of second-generation ablative techniques are fluid-filled thermal balloon EA (TBEA), radiofrequency (thermoregulated) balloon endometrial ablation, hydrothermal EA, 3D bipolar radiofrequency EA, microwave endometrial ablation (MEA), diode laser hyperthermy, cryoablation and photodynamic therapy. The most common techniques in the UK are TBEA

TABLE 2 Complications following hysterectomy (from Garside *et al.*[16])

Very common (>1 in 10)	Common (>1 in 100, <1 in 10)	Uncommon (>1 in 1000, <1 in 100)
Sepsis	Haemorrhage	Death
Pyrexia	Blood transfusion	Fluid overload
Wound haematoma	Anaemia	Visceral damage
Hypergranulation	Vault haematoma	Respiratory/heart complications
Urinary tract infection (UTI)	Anaesthetic	Deep-vein thrombosis
	Diarrhoea	
	Ileus	

(ThermaChoice®, Ethicon, Livingston, UK and Cavaterm™, Pnn Medical SA, Morges, Switzerland)[51–53] and MEA,[54,55] while the NovaSure® (Hologic Inc., Bedford, MA, USA) device[56] is becoming more widely used.

First-generation endometrial ablation techniques

Early methods of EA which require direct hysteroscopic visualisation of the endometrial cavity such as TCRE, RBEA and ELA are known as 'first-generation' ablation techniques.[57] A national survey demonstrated that 99% of first-generation ablative procedures were performed under general anaesthetic.[47] Endometrial thinning agents are conventionally used prior to ablation in order to ensure an adequate depth of destruction. Drugs such as danazol and gonadotrophin-releasing hormone (GnRH) analogues have been shown to improve operating conditions for the surgeon and increase postsurgical amenorrhoea rates.[58] GnRHs were found to produce slightly more consistent endometrial thinning than danazol, although both produced satisfactory results.[58]

Transcervical resection of the endometrium requires a hysteroscope with a fibre optic cable to transmit light from an external power source. The cervix is dilated prior to insertion of the resectoscope, which provides a 12° angle of view. A continuous-flow outer sheath circulates liquid (usually glycine) to provide a clear view of the uterine cavity. A cutting loop is used to remove the endometrial lining. TCRE provides good samples of endometrium for biopsy.[16] TCRE may also be used for the excision of small fibroids, and the operation[17] is usually done as a day case.

Rollerball EA also requires visualisation and irrigation using a resectoscope. EA is achieved by means of a rollerball (RB) electrode rather than a cutting loop. A current is passed through the ball which is moved across the surface of the endometrium.[59] As the RB fits better within the cornua and decreases the chance of perforating this relatively thin-walled part of the uterus,[47] some surgeons prefer to use the RB to treat this area. In the UK, it is usual for TCRE to be used to treat the uterine walls while RBEA is used for the fundus and cornua.[60]

Potential perioperative adverse effects associated with TCRE and RBEA include electrosurgical vaginal and vulval burns, uterine perforation, haemorrhage, gas embolism, infection and fluid overload (which may cause congestive cardiac failure, hypertension, haemolysis, coma and death). Strategies for avoiding fluid absorption include maintaining the minimum intrauterine pressure compatible with safe surgery, using an efficient suction system to retrieve irrigation fluid and maintaining a strict fluid balance.[39]

Possible adverse effects of first-generation ablation techniques are shown in *Table 3*.

TABLE 3 Complications of first-generation EA techniques (from Garside *et al.*)[16]

Common (>1 in 100, <1 in 10)	Uncommon (>1 in 1000, <1 in 100)
Haemorrhage	Death
Uterine perforation	Pregnancy
Sepsis	Cardiovascular/respiratory complications
Pyrexia	Visceral burn
Fluid overload	Haematoma
	GI obstruction/ileus
	Laparotomy

GI, gastrointestinal.

Second-generation endometrial ablation techniques

Since the 1990s, several new methods of EA have been developed. These are often referred to as second-generation techniques. They do not require direct visualisation of the uterine cavity and employ a variety of means to destroy the endometrium – circulation of heated saline within the uterine cavity, use of a diode laser [endometrial laser intrauterine thermo-therapy (ELITT)], punctual vaporising methods, photodynamic methods, radiofrequency, microwaves, a balloon catheter filled with heated fluid and cryotherapy. The treatments are much less dependent on the skill of the surgeon than first-generation techniques, and much more dependent on the reliability of the machines used to ensure safety and efficacy. Complications associated with second-generation techniques include equipment failure, uterine infection, perforation, visceral burn, bleeding and cyclical pain. A limited number of randomised trials indicate that these procedures appear to be as effective as first-generation ablative techniques.[61] In addition, some have the added benefit of being performed under local anaesthetic.

Microwave endometrial ablation

Microwave EA uses microwave energy (at a frequency of 9.2 GHz) to destroy the endometrium with a tissue penetration depth of ⩾6 mm. An 8-mm applicator inserted through the cervix delivers the microwaves using a dielectrically loaded waveguide.[62] Power is controlled by the surgeon using a footswitch and the temperature inside the uterus is monitored by thermocouples on the surface of the waveguide. Prior to microwave ablation treatment, oral and vaginal thinning agents may be given. Immediately prior to MEA, hysteroscopy is performed to exclude false passages, wall damage and perforation.

Following measurement of uterine cavity length, the cervix is dilated to Hegar 8 or 9 under general or local anaesthetic and the uterine cavity length is measured again. Next, the microwave probe is inserted until the tip reaches the fundus. Graduated centimetre markings on the applicator shaft confirm the length and if these three measurements of uterine length are the same the device is activated.[63] When, after a few seconds, the temperature reaches 80 °C, the probe is moved laterally so that the tip is placed in one of the uterine cornua. The temperature briefly falls and rises again and when 80 °C is reached again the probe is moved to the other cornual region and the procedure is repeated.

Maintaining a temperature of 70–90 °C, the probe is withdrawn with side-to-side movements. The temperature measured by the thermocouple is actually the heat transmitted back from the tissue through the plastic sheath to the applicator shaft. Tissue temperature is higher than these measured levels during active treatment. As a marker on the probe appears at the external os, the applicator is switched off to avoid treating the endocervix. The procedure takes 2–3 minutes.[62] Postoperative analgesia is provided as required.

Thermal balloon endometrial ablation

Thermal balloon EA aims to destroy the endometrium by means of heated liquid within a balloon inserted into the uterine cavity, which should be of normal size and regular shape. Available devices such as ThermaChoice and Cavaterm have an electronic controller, a single-use latex or silicone balloon catheter with a heating element, thermocouples and an umbilical cable. As the balloon must be in direct contact with the uterine wall, the device is unsuitable for women with large or irregular uterine cavities.

In the case of the ThermaChoice device, following dilatation of the cervix to about 5 mm, the balloon is introduced within the uterus and filled with sterile fluid (5% dextrose in water) which causes it to expand to fit the cavity. Once intrauterine pressure is stabilised to 160–180 mmHg, the temperature of the fluid is raised to 87°C and maintained for 8 minutes. Pressure, temperature and time are continuously monitored and the device is switched off if safety parameters

are breached. The heat produced by the device causes destruction of the endometrium. Postoperatively, oral analgesia is prescribed and the treated endometrium sloughs off over the following week to 10 days.

The Cavaterm device acts in a similar fashion. The cervix is dilated to about 6 mm and a silicone balloon is inserted and filled with sterile 5% glucose solution to a pressure of 230–240 mmHg. The liquid is heated at a target temperature of 78°C for 10 minutes.

Endometrial thinning agents are not recommended although curettage immediately prior to the procedure can be used. NSAIDs are given to reduce perioperative cramping.

Impedance-controlled bipolar radiofrequency (NovaSure)

The NovaSure system consists of a single-use bipolar ablation device which is inserted into the uterine cavity transcervically after dilatation to 8 mm. This is connected to a generator which functions at 500 kHz and has a power cut-off limit set at a tissue impedance of 50 Ω. The cavity and cervical length are measured and the difference in centimetres is determined; this setting is selected on the shaft of the device. The device is inserted and the trigger is deployed which delivers the bipolar triangular array into the cavity. With gentle tapping and slight rotation in both directions the array fully deploys with the tips sited in each cornua. The distance between the cornuae is displayed and then entered into the generator, and this determines the energy required. A cavity integrity test is then automatically performed, which must be passed before the energy is delivered. Active treatment times are under 2 minutes, during which time suction pulls the walls onto the device. After completion the array is retracted into the device sheath and withdrawn. While the device is versatile, it cannot effectively treat larger cavities (> 11 cm) or distorted cavities. Pre-treatment endometrial thinning is not required and the procedure can be performed under local anaesthetic. A number of randomised trials has been undertaken, one comparing with RBEA[56] and the others with thermal balloon devices.[64,65] One of the trials has published follow-up to 5 years.[66] A randomised trial comparing NovaSure with microwave ablation has been completed and awaits publication. It has approval from NICE.[15] Results have been consistent through the trials, with amenorrhoea rates varying between 42% and 56%, high satisfaction rates of over 90% and low hysterectomy rates. Active treatment times vary between 90 and 120 seconds.

Adverse effects associated with second-generation EA devices include the following:[16] uterine infection, perforation, visceral burn, bleeding, haematometra, laceration, intra-abdominal injury and cyclical pain.

Use of local anaesthetic

Use of local anaesthetic (LA) is a potential advantage of second-generation EA techniques, although this may not be suitable for all women. In a partially randomised trial of general anaesthetic (GA) and LA,[67] the procedure was considered acceptable under GA in both preferred (100%) and randomised (97%) groups. However, under LA, 97% of those who chose this method and 85% of those allocated to LA found the procedure acceptable.

Selecting an appropriate treatment for heavy menstrual bleeding

The introduction of new EA techniques over the last two decades has been accompanied by a series of randomised clinical trials aimed at evaluating their clinical effectiveness and cost-effectiveness. Initially, first-generation EA techniques such as TCRE and laser EA were compared with hysterectomy.[31] Subsequent trials, which compared alternative first-generation techniques such as TCRE, laser EA and RBEA, established TCRE as the gold standard for this group of treatments. As less invasive and more user-friendly second-generation techniques such as MEA became available, these were compared with earlier methods of ablation like TCRE and RBEA.

Although not all techniques have been subjected to head-to-head comparisons in the context of randomised trials, an overview of the literature demonstrates that MEA (second generation) has been shown to be comparable with TCRE (first generation) – which in turn has been shown to be an effective alternative to hysterectomy (gold standard). However, questions about the long-term clinical effectiveness and cost implications of alternative forms of surgical treatment remain unanswered. Published data report no more than 5 years of follow-up.[46,68] Inevitably, some women treated by EA will eventually require repeat ablation or hysterectomy. Following hysterectomy, a proportion of women will also develop further complications such as postsurgical adhesions and pelvic floor dysfunction, which may lead to further surgery. The necessity for a head-to-head comparison between the two most common second-generation methods – MEA and TBEA – has been identified.[15] Given the widespread use of ablative techniques as first-line surgical treatment for menorrhagia at the present time, it is uncertain whether it is either necessary or feasible to compare second-generation techniques directly with hysterectomy. At the same time, the need to obtain comparative information on long-term outcomes is clearly accepted, as is the need to identify the best technique for individual women.

From a clinical perspective, the most relevant research questions at the present time are:

1. How do the currently used ablative techniques compare with hysterectomy in the medium to long term?
2. Which among the commonly used second-generation ablation techniques is the most effective and cost-effective?
3. Are there subgroups of women who are most likely to benefit from either hysterectomy or specific types of ablation?

In this project we have performed a series of studies in order to address these questions by analysis of data from national data sets and randomised trials. Long-term outcomes following EA and hysterectomy in a national cohort have been explored by means of record linkage, while the effectiveness of Mirena, hysterectomy and EA have been determined by individual patient data (IPD) meta-analysis of existing trials. The output has been used (along with other data from the literature) to create a model for the utilisation and costs of the different treatments.

Project objectives

1. To determine, using data from record linkage and follow-up of randomised and non-randomised cohorts of British women, the long-term effects of various second-generation ablative techniques and hysterectomy in terms of failure rates, complications, QoL and sexual function.
2. To determine, using IPD meta-analysis of existing RCTs, the short- to medium-term effects of various second-generation ablative techniques and hysterectomy, including the exploration of outcomes in clinical subgroups.
3. To undertake a model-based clinical effectiveness and cost-effectiveness analysis comparing various second-generation ablative techniques with hysterectomy using output from the above analyses and to conduct extensive sensitivity analyses to explore the robustness of the results to the assumptions made.
4. To devise an algorithm for clinical decision making regarding the choice of surgery for women with HMB in whom medical treatment has failed.

Chapter 2

Hysterectomy, endometrial ablation and Mirena for heavy menstrual bleeding: a systematic review and individual patient data meta-analysis

Introduction

Many women with HMB are not satisfied with medical treatment and end up undergoing surgery.[69] Hysterectomy was once the only surgical option for HMB; indeed, almost half of the hysterectomies currently performed worldwide are for the treatment of HMB.[5] EA techniques, which aim to destroy or remove the endometrial tissue,[70] have become increasingly popular alternatives, and, as a result, the number of hysterectomies in the UK has declined by 64% between 1995 and 2002.[71] EA techniques were introduced in the 1980s, with RB ablation and transcervical resection emerging as the predominant approaches under direct hysteroscopic vision.[30] Subsequently, a second generation of non-hysteroscopic techniques, which are easier to perform, have become available. Here, devices are sited and activated to treat the whole endometrial cavity simultaneously without visual control. Destruction is achieved through a variety of modalities, including high temperature fluids and bipolar electrical or microwave energy. Intrauterine coil devices were initially introduced as contraceptives, but the addition of progestogen resulted in reduced menstrual bleeding. Mirena, an LNG-releasing IUS, provides a non-surgical alternative, which is reversible and fertility sparing.[72]

Women and clinicians now have greater choice of treatment, although evidence to support decision making is inadequate. In the UK, guidelines from NICE[15] recommend the use of Mirena in the first instance for women with benign HMB, followed by EA if pharmaceutical treatments fail to resolve symptoms. Syntheses of evidence from RCTs comparing these treatments have been limited,[31,61,73] partly because of the scarcity of head-to-head comparisons and variation in outcome measurements used to evaluate effectiveness. We undertook a meta-analysis of IPD from all relevant trials to address previous deficiencies in evidence synthesis. The aim of the study was to compare the relative efficacy of hysterectomy, first- and second-generation EA techniques, and Mirena in women with HMB using a primary outcome measure of patient dissatisfaction. IPD meta-analysis has a number of advantages over traditional published data reviews,[74] including the ability to carry out data checks, standardise analytical methods and undertake subgroup analyses.

Methods

We sought IPD from RCTs of hysterectomy, EA techniques and Mirena to examine their relative efficacy as a second-line treatment of HMB. The systematic review was conducted based on a protocol designed using widely recommended methods[75,76] that complied with meta-analysis reporting guidelines[77] (see *Appendix 10*) (www.bctu.bham.ac.uk/systematicreview/hmb/protocol.shtml).

Literature search and study selection

The Cochrane Library, MEDLINE (1966–2010), EMBASE (1980 to May 2010) and CINAHL databases (1982 to May 2010) were searched using relevant terms and word variants for population [e.g. menorrhagia, hypermenorrhea, (excessive) menstrual blood loss, HMB, dysfunctional uterine bleeding] and interventions (e.g. hysterectomy, vaginal hysterectomy, total abdominal hysterectomy, subtotal abdominal hysterectomy, laparoscopic hysterectomy, LNG IUS, Mirena coil and all types and variants of first- and second-generation ablative techniques). Variant names (e.g. EA, resection) and different makes for EA (e.g. Microsulis Medical Ltd, Denmead, UK; Cryogen – now American Medical Systems, Minnetonka, MN, USA) were also searched (see *Appendix 1*). We also hand-searched the bibliographies of all relevant primary articles and reviews to identify any articles missed by the electronic searches. Experts were contacted to identify further studies. To identify any ongoing RCTs, the Meta-Register of Controlled Trials and the International Standard Randomised Control Trial Number register were searched. No language restriction was applied.

Studies were selected in a two-step process. Firstly, we scrutinised the citations identified by the electronic searches and obtained full manuscripts of all the citations that met, or were thought likely to meet, the pre-determined inclusion criteria based on patient entry criteria (women with HMB or abnormal/excessive/prolonged uterine bleeding that was unresponsive to other medical treatment) and study design, the latter limited to RCTs only. We then considered four categories of intervention with the intention of comparing them against each other: *hysterectomy* (performed abdominally, vaginally or laparoscopically); *'first-generation' EA techniques* (using operative hysteroscopy, including endometrial laser ablation, TCRE and RBEA); *'second-generation' EA techniques* [those that use a 'blind' device to simultaneously treat the whole cavity, including thermal balloon (Cavaterm, ThermaChoice and Vesta), microwave (Microsulis), laser (ELITT), bipolar radio frequency (NovaSure), cryoablation and hydrothermal ablation]; and *LNG IUS* (Mirena). Studies making a comparison within these categories could not contribute to the meta-analysis, but these data were also requested to allow further exploration of possible predictors of the primary outcome measure.

Data collection and study quality assessment

Repeated attempts were made to contact corresponding authors via post, email or telephone to access data. Where initial attempts failed, we attempted personal contact via our links through the British and European Societies for Gynaecological Endoscopy. Authors were asked to supply anonymised data for each of the pre-specified outcome measures (both published and unpublished to reduce the chance of selective reporting bias) and were invited to become part of the collaborative group with joint ownership of the final publication. Where the investigators declined to take part in the study or could not be contacted, published data were extracted from manuscripts using pre-designed proformas by two independent reviewers (RC and LJM). Any disagreements were resolved by consensus or arbitration by a third reviewer (JPD). Received data were merged into a master database, specifically constructed for the review. The data were cleaned and results were cross-checked against published reports of the trials. Where discrepancies existed authors were contacted for clarification.

Authors of the protocol reviewed all relevant outcome measures to be used in the meta-analysis from articles identified in the literature search. Level of satisfaction with treatment was the most frequently measured outcome across all identified studies, with 21 out of 30 (70%) using this measure, and was used as the primary outcome measure. Dissatisfaction rates are presented to simplify interpretation of statistical output. Responses of 'very satisfied' or 'satisfied' were taken as a positive response, likewise 'very dissatisfied' or 'dissatisfied' as a negative response. Where a 'not sure' or 'uncertain' response was given these were conservatively taken to be a negative rating of treatment, although sensitivity analysis was undertaken to test the robustness of this assumption.

For a small number of studies,[78–81] surrogate outcomes for satisfaction were used (major problem resolved/improvement of health state/menstrual symptoms successfully treated/degree of recommendation). This assumption was also tested by sensitivity analysis without these studies (indicated in the results section where important) (see *Appendix 2*). A more disease-specific QoL tool[19] would have been the ideal choice for primary measure, but these data were not available from the studies identified. We have shown from the data in this review, though, that a strong relationship between dissatisfaction and patient QoL is apparent (see *Results*).

Other outcome measures were bleeding scores (ranging from a minimum of zero to no upper limit),[12] amenorrhoea rate (converted from a bleeding score of zero where data existed, otherwise as reported), heavy bleeding rate (converted from bleeding scores of > 100[12] where data existed, otherwise as reported), EQ-5D utility score,[82] SF-36 scores,[20] duration of surgery/hospital stay, general anaesthesia rates, postoperative pain score (standardised from visual analogue and ordinal scale scores onto a 0–10 scale), time to return to work/normal activities/sexual activity, dysmenorrhoea/dyspareunia rate and proportion undergoing subsequent ablation/hysterectomy or discontinuing use of Mirena. Pre-defined subgroups were age at randomisation (≤ 40 vs > 40 years), parity (nulliparous vs parous), uterine cavity length (≤ 8 vs > 8 cm), presence or absence of fibroids/polyps and, where available, severity of bleeding at baseline (bleeding score ≤ 350 or > 350).

All selected trials were assessed for their methodological quality, using received data sets where available in addition to the reported information. Quality was scrutinised by checking the adequacy of randomisation, group comparability at baseline (examining baseline characteristics for any substantive differences), blinding (where appropriate), use of intention-to-treat (ITT) analysis, completeness of follow-up, compliance, reliability using a priori sample size estimation and generalisability using description of the sample recruited. Adequacy of randomisation was assessed with subquestions examining information on sequence generation, the process of allocation and allocation concealment.

Statistical analysis

To minimise the possibility of bias IPD and aggregate data (AD) were combined in a two-stage approach.[83] IPD were reduced to AD to allow studies with AD only to be combined with those where IPD were obtained. Unless specifically stated in the text of the results section, all estimates shown are from all available data (both IPD and AD). Point estimates and 95% confidence intervals (CIs) were calculated for individual studies at each time point. For the primary outcome measure, differences in effect estimates between trials and the pre-defined subgroups of patients are displayed using odds ratio (OR) plots; results from other outcome measures are summarised in tables in the appendices for ease of reference. Heterogeneity was investigated using Cochran's Q,[84] I^2 statistics[24] and Higgins *et al.*[85] Subgroup analyses to explore the causes of heterogeneity were undertaken if the *p*-values of these tests were < 0.1. Differences between studies contributing IPD and those with AD only were examined in the same fashion to check that the latter results were consistent with those we received IPD for. Further details are given in the *Results* section if any inconsistency exists. Likewise, further details are given on any obvious publication bias if noted from the assessment of funnel plots. Only a limited amount of data were available for studies comparing Mirena with ED, so Mirena was compared with first- and second-generation studies combined as well as separately. Assumption-free 'fixed-effect' methods were used to combine dichotomous outcome measures and estimate pooled ORs using the method of Peto *et al.*,[86] and, for continuous variables, weighted mean differences (WMD) were calculated[87] at each time point. Data at less than 12 months were combined and are described as results at 6 months. Results from the limited number of studies with follow-up longer than 2 years are not referred to in the text but are given in the appendices.

The primary outcome measure of dissatisfaction was investigated comprehensively using received data. Results at 12 months, where the majority of studies had collected data, were used as the focus for analysis. Where responses were not available at this time point, data were substituted in the first instance from 2 years and, failing that, from 6 months. If it was not possible to make a direct comparison between treatments (e.g. hysterectomy vs second-generation EA), indirect estimates were made[88] using a logistic regression model[89] allowing for trial and treatment.[90] Estimates using dissatisfaction at any time were also examined, along with an analysis allowing for the correlation of the repeated measurements using generalised estimating equations (IPD only).[91]

Access to IPD also allowed the inclusion of patient-level covariates to examine possible predictors of dissatisfaction. First, covariates were considered individually, while allowing for differences between trial estimates. If considered statistically important ($p < 0.1$), covariate parameters were included together in a multivariate analysis to examine adjusted estimates. In addition to the analysis of the primary outcome measure described above, as a sensitivity analysis IPD were also used to explore the effect observed in compliance rates for comparisons between first- and second-generation EA (unfortunately there were insufficient data to extend this analysis to Mirena comparisons). For example, for those women 'satisfied' with treatment but subsequently undergoing a hysterectomy, positive responses were substituted with negative ones. The relationship between dissatisfaction and responses from the SF-36 QoL questionnaire was examined at the patient level using a regression model, allowing for trial. Given the number of analyses performed, any interpretation of *p*-values greater than the conservative threshold of 0.01 has been cautious owing to the likelihood of increased type 1 error rates. REVMAN v5.0 (Cochrane Collaboration, Denmark) and SAS v9.2 (SAS Institute, Cary, NC, USA) software were used for analysis.

Results

Trials and patients

A total of 556 potentially relevant citations were identified by electronic searches. After detailed evaluation, 30 trials were eligible for inclusion in the review (*Figure 1*). Of these trials, seven compared hysterectomy with ED techniques. Six of these studies involved first-generation techniques.[41–45,78] The seventh study used a combination of first and second generation in equal proportions[92] and was included here as a first-generation comparison, with a sensitivity analysis performed without the trial. One study compared hysterectomy[93] with Mirena. Fourteen studies compared first-generation EA techniques with those of second generation[53,54,56,79,94–103] and eight studies compared Mirena with EA, three of which were first generation[80,104,105] and five second generation.[81,106–109] Characteristics of these studies are shown in *Appendix 2*. Data from a further five studies,[64,65,110–112] which involved comparisons within first- and second-generation EA, were also received.

Trials comparing hysterectomy with EA and those comparing first- and second-generation ED involved women of a similar age, with average ages of 40.6 years [standard deviation (SD) 5.1 years] and 41.0 years (SD 4.9 years), respectively. Women involved in trials comparing Mirena with ED were slightly older, with an average age of 43.6 years (SD 3.5 years). Eligibility criteria for women with uterine pathology varied between trials; inclusion of women with fibroids was generally limited by size or number. Where included, they amounted to a maximum of 30% of the women in each individual study.

A high proportion of data was received from trials involving hysterectomy (seven of eight studies; 1278 of 1363 women), with less for trials of EA techniques (7 of 14 studies; 1359 of 2448 women)

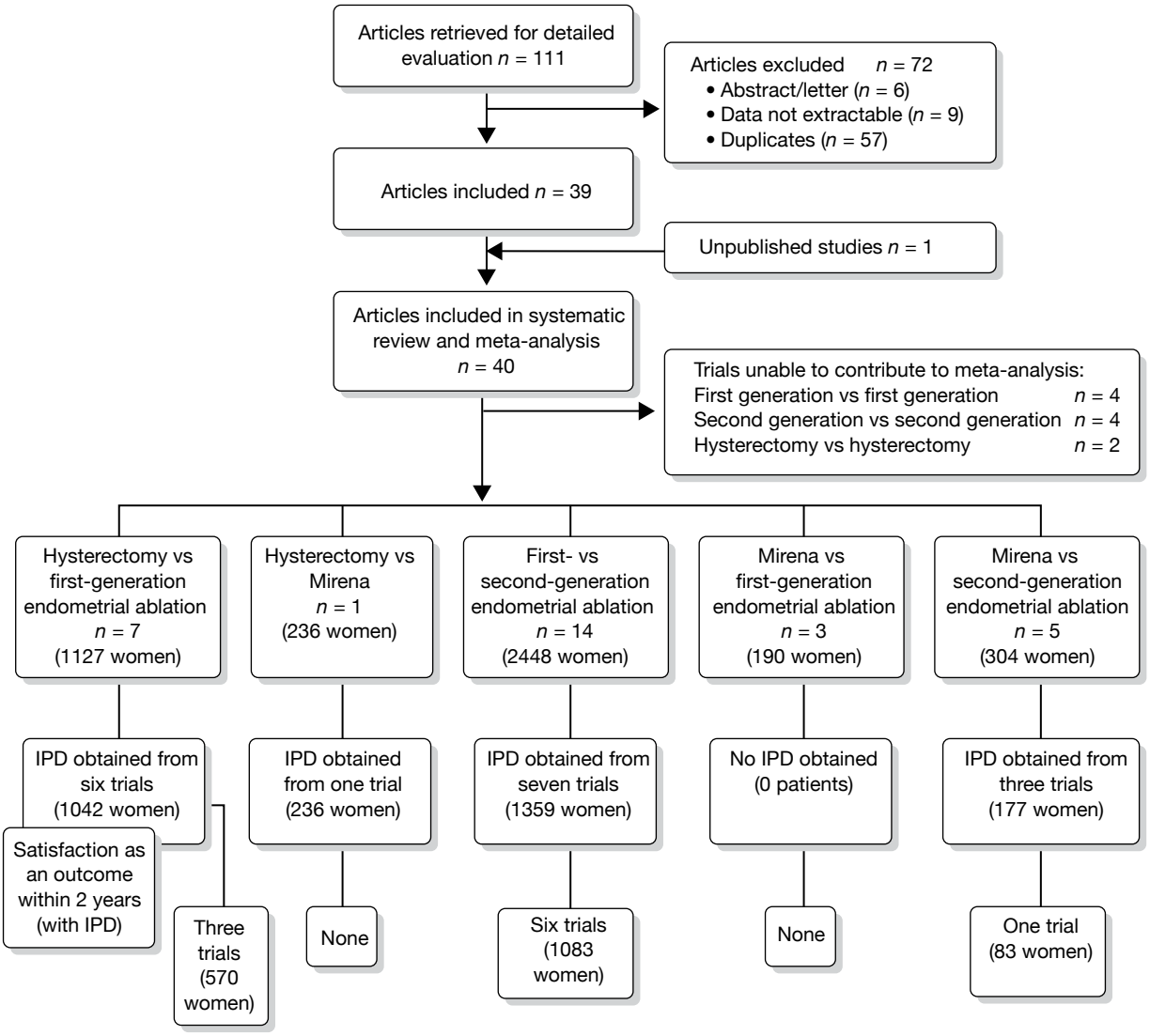


FIGURE 1 Study selection process for systematic review and IPD meta-analysis of randomised trials comparing hysterectomy, EA techniques and Mirena for HMB (details of selected trials in *Appendix 2*).

and those involving Mirena (three of eight studies; 177 of 494 women) (see *Appendix 2*). Overall, we received some IPD from 65% (2814/4305) of women involved in the trials, although only eight studies were able to provide all requested variables.[41,42,53,94,95,99,102,109] The remaining studies had some missing information, with limited details on patient follow-up covering subsequent operations (e.g. hysterectomy following Mirena). See the section on statistical analysis for details on how data from studies providing IPD were utilised.

Study quality

The methodological quality of the published data from the studies was variable (*Figure 2* and *Appendix 3*).

Over half the studies failed to give adequate information about their randomisation procedure and details on allocation concealment. There was a general lack of true ITT analysis, with some studies stating that an ITT analysis had been performed but only analysing those women who

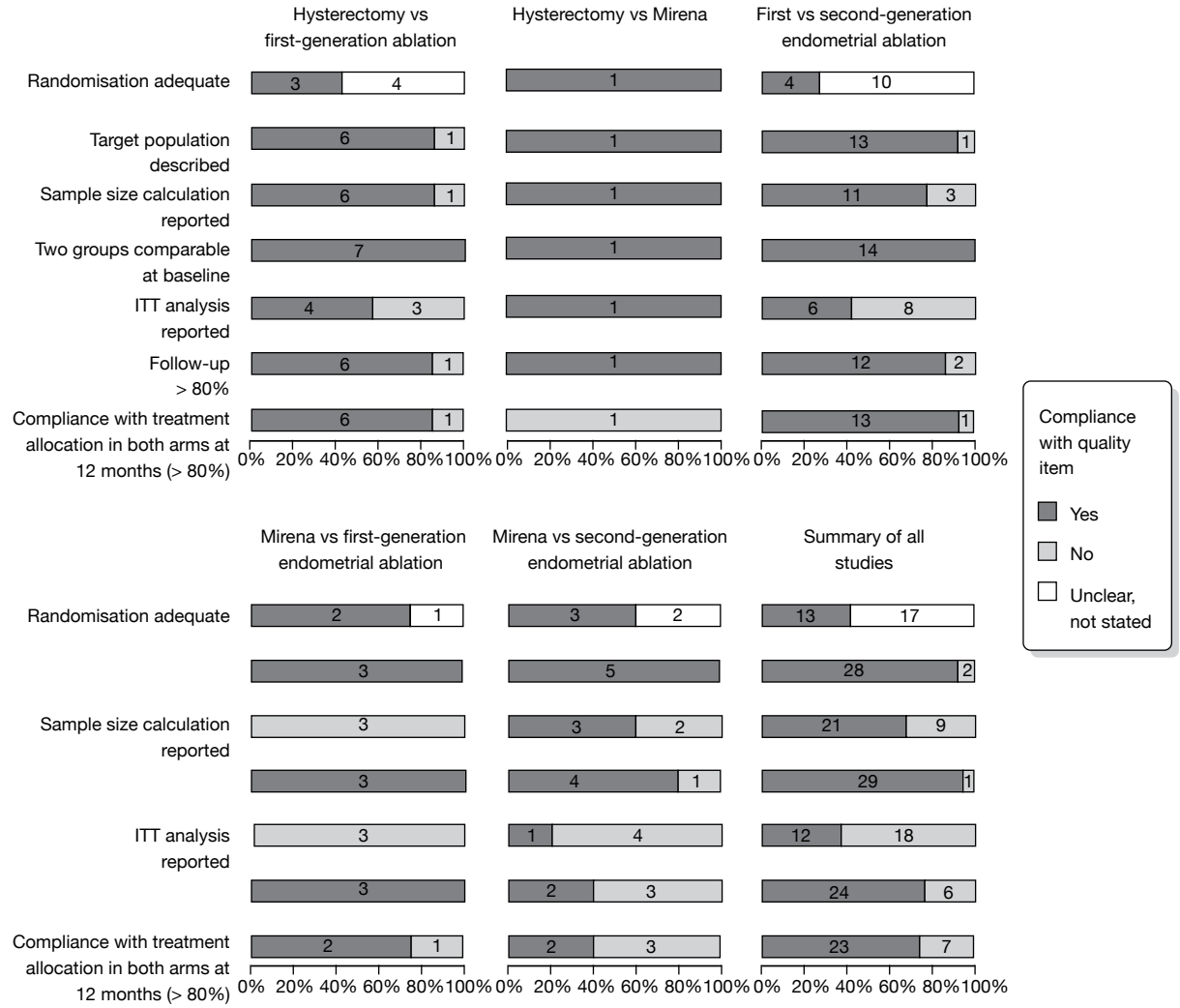


FIGURE 2 Quality of studies included in systematic review and IPD meta-analysis of randomised trials comparing hysterectomy, EA techniques and Mirena for HMB. Numbers inside bars are numbers of studies (details given in *Appendix 3*).

had received treatment. For four studies that reported non-ITT analysis,[53,95,99,107] ITT analyses were undertaken using the available IPD, although it was not always clear if protocol-deviant patients were followed up correctly in these cases. Small sample sizes often lacked a sensible justification, especially in studies involving Mirena. In the nine trials involving Mirena, only four had greater than 80% of women with Mirena in situ at 12 months post randomisation.

Dissatisfaction as an outcome measure

Data from four studies that provided IPD on both outcomes[42,54,99,107] showed that satisfied patients had significantly increased scores in seven of eight domains of the SF-36 QoL questionnaire when compared with dissatisfied patients in the analysis of change from baseline scores, including the general health perception (7.4 points, 95% CI 3.1 to 11.8 points; $p = 0.0008$) and mental health (10.5 points, 95% CI 5.4 to 15.6 points; $p < 0.0001$) domains (*Table 4*). Differences from absolute differences (not adjusted for baseline score) were highly significant ($p < 0.0001$) in all eight domains in favour of satisfied patients.

TABLE 4 Results from regression analysis comparing satisfaction response with results from the SF-36 quality of life questionnaire at 12 months

		Change from baseline			Absolute		
SF-36 domain		Mean (SD; *n*)	Difference[a] (95% CI)	*p*-value	Mean (SD; *n*)	Difference[a] (95% CI)	*p*-value
General health	Dissatisfied	–4.7 (14.2; 71)	7.4 (3.1 to 11.8)	0.0008	60.3 (20.5; 91)	12.5 (8.5 to 16.6)	<0.0001
	Satisfied	5.7 (17.0; 507)			77.7 (17.8; 642)		
Physical function	Dissatisfied	0.4 (19.4; 70)	2.8 (–2.6 to 8.3)	0.3	78.1 (27.2; 89)	10.9 (6.7 to 15.1)	<0.0001
	Satisfied	6.0 (20.7; 497)			91.0 (16.6; 637)		
Role physical	Dissatisfied	5.3 (51.6; 71)	17.4 (5.3 to 29.4)	0.005	60.8 (45.1; 90)	24.0 (17.0 to 31.0)	<0.0001
	Satisfied	25.2 (44; 504)			88.4 (27.9; 641)		
Role emotional	Dissatisfied	4.2 (54.6; 71)	15.0 (2.9 to 27.0)	0.02	61.1 (44.2; 90)	23.4 (16.3 to 30.4)	<0.0001
	Satisfied	18.2 (44.4; 505)			87.4 (28.2; 641)		
Mental health	Dissatisfied	–2.1 (22.7; 71)	10.5 (5.4 to 15.6)	<0.0001	58.5 (21.6; 90)	16.9 (12.8 to 21.0)	<0.0001
	Satisfied	7.6 (18.7; 504)			76.9 (17.1; 638)		
Social function	Dissatisfied	4.9 (26.2; 70)	6.7 (0.9 to 12.5)	0.02	61.0 (24.2; 90)	17.6 (13.5 to 21.7)	<0.0001
	Satisfied	11.6 (21.2; 471)			85.5 (18.6; 629)		
Vitality	Dissatisfied	6.5 (23.7; 70)	8.5 (2.4 to 14.6)	0.006	43.6 (23.1; 91)	18.9 (14.1 to 23.8)	<0.0001
	Satisfied	15.7 (22.9; 503)			65.2 (21.0; 637)		
Pain	Dissatisfied	6.4 (34.3; 71)	9.1 (0.8 to 17.4)	0.03	57.6 (27.2; 91)	20.2 (14.7 to 25.6)	<0.0001
	Satisfied	20.1 (31.4; 504)			81.0 (23.4; 642)		

a Adjusted for study.

Effectiveness in reducing dissatisfaction with treatment

Hysterectomy versus first-generation endometrial ablation

More women were dissatisfied at 12 months following first-generation EA than hysterectomy [12.6% vs 5.3%; (57/454 vs 23/432); OR 2.46; 95% CI 1.54 to 3.93; $p = 0.0002$] (*Figure 3*), with no significant heterogeneity between study estimates ($p = 0.9$; $I^2 = 0\%$). This estimate of effect size was consistent with, although slightly less than, the estimate from the repeated measures analysis (IPD only) over all time points (OR 3.75; 95% CI 2.18 to 6.46; $p < 0.0001$) and an analysis using dissatisfaction at any time point (OR 3.37; 95% CI 2.14 to 5.31; $p < 0.0001$). There was no evidence of any differences between subgroups (see the *Data collection and study quality assessment* section), including between studies providing IPD or AD (test for heterogeneity: $p = 0.9$).

First- versus second-generation endometrial ablation techniques

Similar dissatisfaction rates were seen with first- and second-generation EA [12.2% vs 10.6% (123/1006 vs 110/1034); OR 1.20; 95% CI 0.88 to 1.62; $p = 0.2$; test for heterogeneity: $p = 0.7$)] (*Figure 4*). Comparison estimates were obtained from the repeated measures analysis of IPD (OR 1.21; 95% CI 0.84 to 1.74; $p = 0.3$), the analysis using dissatisfaction at any time (OR 1.22; 95% CI 0.91 to 1.62; $p = 0.2$), and also an analysis adjusting for patients who went on to receive a hysterectomy (OR 1.25; 95% CI 0.93 to 1.67; $p = 0.1$). Results were consistent over all subgroups, including those studies providing IPD or AD only (test for heterogeneity: $p = 0.8$).

Mirena versus endometrial ablation techniques

Rates of dissatisfaction with Mirena and second-generation EA were similar [18.1% vs 22.5% (17/94 vs 23/102); OR 0.76; 95% CI 0.38 to 1.53; $p = 0.4$] (*Figure 5*), although the latter rate was

Study or subgroup	First generation Events	First generation Total	Hysterectomy Events	Hysterectomy Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
Dickersin 2007[92]	13	107	7	103	26.1%	1.86 (0.55 to 6.21)
Crosignani 1997[45]	5	38	2	39	9.3%	2.61 (0.34 to 19.85)
O'Connor 1997[44]	7	106	2	50	10.7%	1.61 (0.24 to 10.65)
Pinion 1994[43]	13	104	4	93	22.4%	2.81 (0.76 to 10.39)
Dwyer 1993[42]	19	99	6	97	31.5%	3.20 (1.06 to 9.61)
Total (95% CI)		**454**		**382**	**100.0%**	**2.46 (1.54 to 3.93)**
Total events	57		21			

Heterogeneity: $\chi^2 = 1.15$, df = 4 ($p = 0.89$); $I^2 = 0\%$
Test for overall effect: $z = 3.75$ ($p = 0.0002$)

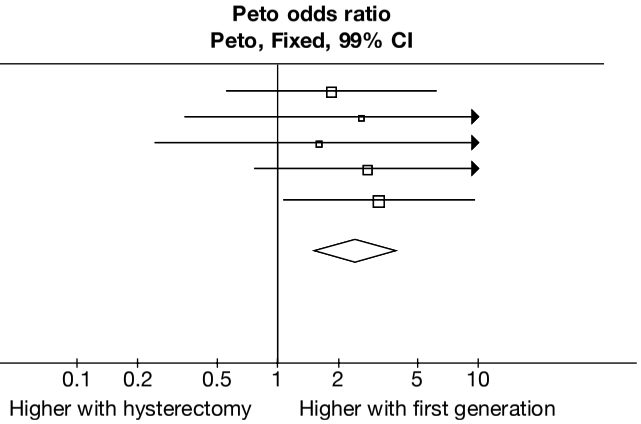


FIGURE 3 Dissatisfaction at 12 months – hysterectomy versus first-generation EA.

Study	First generation Events	First generation Total	Second generation Events	Second generation Total	Weight	Peto odds ratio Fixed, 99% CI
Brun 2006[103]	3	16	2	27	2.5%	2.94 (0.24 to 36.00)
Cooper 2004[99]	1	101	3	201	2.1%	0.68 (0.04 to 10.59)
Perino 2004[100]	5	55	3	56	4.5%	1.74 (0.26 to 11.42)
Duleba 2003[98]	10	72	16	156	12.0%	1.43 (0.45 to 4.53)
Hawe 2003[94]	3	33	4	37	3.8%	0.83 (0.11 to 6.38)
van Zon-Rabelink 2004[95]	13	58	15	75	13.1%	1.16 (0.38 to 3.47)
Cooper 2002[56]	5	82	11	154	8.1%	0.85 (0.21 to 3.43)
Pellicano 2002[102]	14	38	7	37	9.2%	2.40 (0.64 to 8.96)
Soysal 2001[96]	19	48	15	45	13.1%	1.31 (0.43 to 3.94)
Cooper 1999[54]	33	128	29	121	28.0%	1.10 (0.52 to 2.34)
Meyer 1998[53]	1	120	5	125	3.5%	0.27 (0.03 to 2.24)
Total (95% CI)		**751**		**1034**	**100.0%**	**1.20 (0.88 to 1.62)**
Total events	107		110			

Heterogeneity: $\chi^2 = 7.47$, df = 10 ($p = 0.68$); $I^2 = 0\%$
Test for overall effect: $z = 1.17$ ($p = 0.24$)

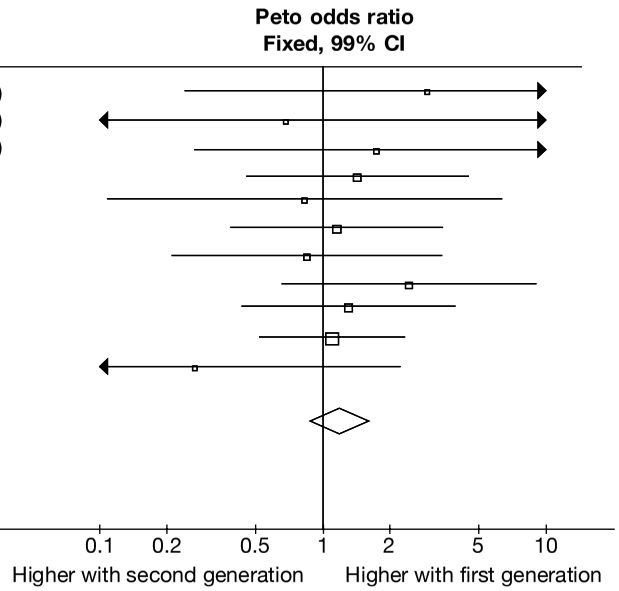


FIGURE 4 Dissatisfaction at 12 months – first- versus second-generation EA.

twice as high as that seen for second-generation EA when it was compared with first-generation EA (see *Figure 4*). The slightly older age of women in these studies (see *Trials and patients*) could be a possible explanation for this increase, although given the small number of women studied in these trials this difference in rate could easily have arisen by chance. The combined estimate of this and the one study that compared Mirena with first-generation EA[80] (test for differences between subgroups: $p = 0.2$) also showed no evidence of a difference (OR 0.94; 95% CI 0.50 to 1.77; $p = 0.9$). Heterogeneity of estimates overall was of borderline statistical significance ($p = 0.09$; $I^2 = 54\%$). Overall rates of dissatisfaction were 17.2% (22/128) for Mirena and 18.2% (25/137) for both first- and second-generation EA. Lack of IPD prohibited any further investigation of subgroups or repeated measures. Sensitivity analysis performed without two studies where surrogates for dissatisfaction were used significantly reduced the data available for analysis but did not change the findings.

Indirect comparisons of hysterectomy with second-generation endometrial ablation techniques and Mirena

Indirect estimates (*Figure 6*) suggest that hysterectomy is also preferable to second-generation EA [5.3% vs 10.6% (23/432 vs 110/1034); OR 2.32; 95% CI 1.27 to 4.24; $p = 0.006$) in terms of patient dissatisfaction. This is confirmed by the repeated measures analysis over all three time points, which per force only include IPD (OR 3.06; 95% CI 1.59 to 5.90; $p = 0.0008$). The evidence to suggest hysterectomy is preferable to Mirena was weaker [5.3% vs 17.2% (23/432 vs 22/128); OR 2.22; 95% CI 0.94 to 5.29; $p = 0.07$], but given the lack of precision from Mirena comparisons this was not a surprising result and should be cautiously interpreted.

Predictors of dissatisfaction

For second-generation EA, IPD showed that uterine cavity length was the strongest predictor of dissatisfaction ($p = 0.02$), with shorter uterine cavity length (≤ 8 cm vs > 8 cm) associated with reduced rates (OR 0.59; 95% CI 0.38 to 0.93; $p = 0.02$) (*Table 5*). Absence of fibroids/polyps also showed a trend towards reduced dissatisfaction ($p = 0.07$), although no further adjusted estimates including both parameters were attempted as only three studies had data on fibroids/polyps. There were no convincing associations with any of the variables for hysterectomy or first-generation EA.

Effectiveness in improving other outcomes

Hysterectomy versus endometrial ablation and Mirena

These comparisons focused on recovery times and QoL, as estimates of postoperative menstrual blood loss are redundant after hysterectomy (see *Appendix 4*). EA offered quicker surgery (WMD 32 minutes; 95% CI 30 to 34 minutes; $p < 0.0001$), shorter hospital stay (WMD 3.0 days; 95% CI 2.9 to 3.1 days; $p < 0.00001$), faster recovery periods (time to return to normal activities: WMD 5.2 days; 95% CI 4.7 to 5.7 days; $p < 0.00001$) and less pain postoperatively (WMD 2.5 points; 95% CI 2.2 to 2.9 points; $p < 0.0001$), although estimates of differences for some of these parameters should be used with caution given the high variability between studies (see *Appendix 4*). One study[92] suggested no obvious difference in EQ-5D utility scores, while another[78] suggested differences in favour of hysterectomy in the general health (WMD 9.6 points; 95% CI 5.7 to 13.5 points; $p < 0.0001$), social function (WMD 24 points; 95% CI 21 to 27 points; $p < 0.0001$) and vitality (WMD 13 points; 95% CI 9.3 to 16 points; $p < 0.0001$) domains of the SF-36 questionnaire. Perioperative adverse events associated with hysterectomy were relatively few (0.5%–2.0% each), but UTIs were more common with hysterectomy (43/530; 8.1%) than with EA (9/585; 1.5%). Of the women who were initially treated with EA, 15% had undergone a hysterectomy within 2 years.

Study	Mirena Events	Mirena Total	First generation/ second generation Events	First generation/ second generation Total	Weight	Peto odds ratio Fixed, 99% CI
First generation						
Crosignani 1997[45]	5	34	2	35	16.8%	2.64 (0.34 to 20.32)
Subtotal (95% CI)		**34**		**35**	**16.8%**	**2.64 (0.56 to 12.48)**
Total events	5		2			
Heterogeneity: not applicable						
Test for overall effect: $z = 1.23$ ($p = 0.22$)						
Second generation						
Shaw 2007[108]	5	25	5	28	21.7%	1.15 (0.19 to 6.90)
Busfield 2006[107]	2	37	9	39	25.1%	0.24 (0.05 to 1.30)
Soysal 2002[106]	10	32	9	35	36.3%	1.31 (0.33 to 5.24)
Subtotal (95% CI)		**94**		**102**	**83.2%**	**0.76 (0.38 to 1.53)**
Total events	17		23			
Heterogeneity: $\chi^2 = 4.43$, df = 2 ($p = 0.11$); $I^2 = 55\%$						
Test for overall effect: $z = 0.76$ ($p = 0.44$)						
Total (95% CI)		**128**		**137**	**100.0%**	**0.94 (0.50 to 1.77)**
Total events	22		25			
Heterogeneity: $\chi^2 = 6.48$, df = 3 ($p = 0.09$); $I^2 = 54\%$						
Test for overall effect: $z = 0.19$ ($p = 0.85$)						
Test for subgroup differences: $\chi^2 = 2.05$, df = 1 ($p = 0.15$), $I^2 = 51.3\%$						

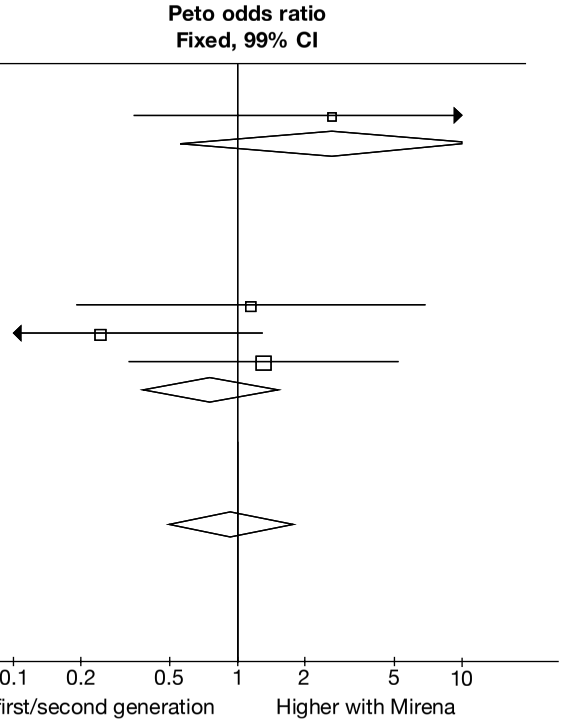


FIGURE 5 Dissatisfaction at 12 months – first- and second-generation EA versus Mirena.

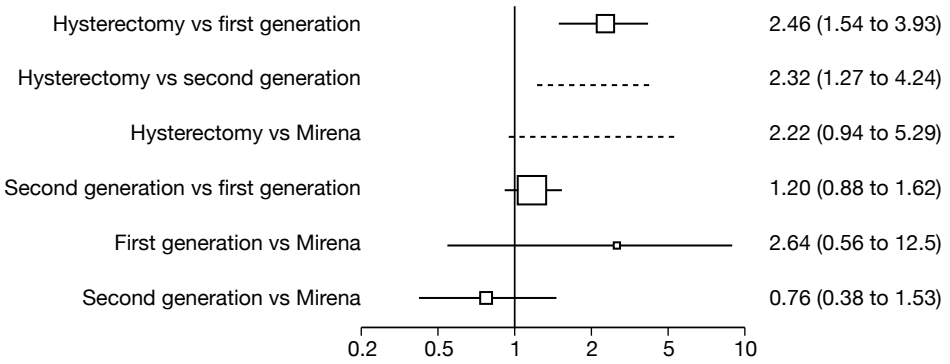


FIGURE 6 Dissatisfaction at 12 months summary. Odds ratios (95% CIs) are shown. Estimates >1 indicate increased dissatisfaction for the second treatment listed. Indirect estimates are represented by a dotted line.

TABLE 5 Results from logistic regression analysis with dissatisfaction at 12 months as the outcome

	Hysterectomy			First-generation EA			Second-generation EA		
	OR (95% CI)[a]	*n*	*p*-value	OR (95% CI)[a]	*n*	*p*-value	OR (95% CI)[a]	*n*	*p*-value
Individual estimates[b]									
Uterine cavity length, cm (≤8 vs >8)	–	–	–	0.97 (0.38 to 2.44)	418	0.9	0.59 (0.38 to 0.93)	817	0.02
Age, years (≤40 vs >40)	2.28 (0.66 to 7.89)	239	0.2	1.21 (0.81 to 1.81)	971	0.4	1.30 (0.87 to 1.93)	942	0.2
Fibroids/polyps (absence vs presence)	0.51 (0.14 to 1.93)	233	0.3	1.15 (0.55 to 2.38)	476	0.7	0.36 (0.12 to 1.07)	302	0.07
Parity (nullparous vs parous)	–	–	–	1.27 (0.36 to 4.43)	778	0.7	0.84 (0.33 to 2.16)	734	0.7
Baseline bleeding score (≤350 vs 350)	–	–	–	0.73 (0.27 to 1.97)	328	0.5	0.96 (0.48 to 1.91)	551	0.9

[a]Estimates <1 favour the first subgroup listed, i.e. have reduced dissatisfaction.
[b]After allowing for study.

No differences in EQ-5D scores were seen at 6 or 12 months in the single study comparing hysterectomy with Mirena (see *Appendix 5*), while the only statistically significant effect observed in the SF-36 questionnaire was in the pain domain, which favoured hysterectomy (WMD 9.6 points; 95% CI 2.7 to 16.6 points; $p = 0.007$). All results were consistent over subgroups.

First- versus second-generation endometrial ablation techniques

The proportion of women with amenorrhoea or still experiencing heavy bleeding was similar in both groups at all time points apart from at 2 years, where there was a borderline significant difference in favour of second-generation techniques (amenorrhoea: OR 0.64, 95% CI 0.41 to 0.99, $p = 0.04$; HMB: OR 1.85, 95% CI 1.04 to 3.32, $p = 0.04$) (see *Appendix 6*). High heterogeneity for the estimate of amenorrhoea rate at 12 months (OR 1.12, 95% CI 0.93 to 1.35, $p = 0.3$) appeared to be due to two outlying studies (Duleba[98] and Perino[100]), the results of which could not be verified as IPD were not available. However, analysis without these studies gave very similar results (OR 1.10, 95% CI 0.90 to 1.35, $p = 0.4$) with lowered heterogeneity ($p = 0.1$; $I^2 = 36\%$). Change from baseline analysis of bleeding scores showed no evidence of a difference at any of the time points. Inconsistency of estimates for this outcome at 12 months was due to a single study;[103] sensitivity analysis without this study also showed no change to the overall result (WMD –0.3; 95% CI –27.5 to 27.0; $p = 1.0$; heterogeneity: $p = 0.4$; $I^2 = 10\%$). Two studies[54,99]

using the SF-36 questionnaire and one small study[94] using the EQ-5D questionnaire showed no consistent difference between first- and second-generation techniques, in terms of change from baseline results.

Second-generation EA was quicker (WMD 15 minutes; 95% CI 14 to 15 minutes; $p < 0.0001$) and less likely to need general anaesthesia than first generation (OR 0.16; 95% CI 0.12 to 0.20; $p < 0.0001$), although highly significant heterogeneity makes estimates difficult to interpret. Less frequent use of general anaesthesia with second-generation EA translated to a slightly quicker time to return to work (WMD 1.36 days; 95% CI 0.69 to 2.03 days; $p < 0.0001$) and time to return to normal activities (WMD 0.48 days; 95% CI 0.20 to 0.75 days; $p = 0.0008$), although the overall estimate for the latter was somewhat inconsistent (heterogeneity: $p = 0.04$; $I^2 = 59\%$). Postoperative pain was similar following either method of EA, although estimates from different studies varied widely (heterogeneity: $p < 0.0001$; $I^2 = 89\%$) without any obvious explanation. Adverse events were relatively low in both groups (each < 2%), but perioperative complications such as uterine perforation (OR 0.20; 95% CI 0.07 to 0.57; $p = 0.003$), excessive bleeding (OR 0.14; 95% CI 0.07 to 0.55; $p = 0.005$), fluid overload (OR 0.12; 95% CI 0.04 to 0.36; $p = 0.0001$) and cervical laceration (OR 0.12; 95% CI 0.05 to 0.33; $p < 0.0001$) were lower with second-generation EA. The number of women requiring a subsequent hysterectomy was lower for second-generation EA, but these differences were not large enough to be statistically significant within the first 2 years (12 months: OR 0.77, 95% CI 0.47 to 1.24, $p = 0.3$; 2 years OR 0.68, 95% CI 0.41 to 1.13, $p = 0.1$). Overall rates were 3.3% (74/2265) and 7.6% (71/939) at these time points. Any differences amongst subgroups were confined to single time points only. Results from studies providing IPD were consistent with those with AD only.

Mirena versus endometrial ablation techniques

Fewer women experienced HMB after Mirena at 6 months (OR 0.23; 95% CI 0.09 to 0.57; $p = 0.001$) and at 2 years (OR 0.08; 95% CI 0.01 to 0.50; $p = 0.007$), although total numbers here were small compared with the estimate at 12 months, where there was no evidence of any difference (OR 0.74; 95% CI 0.34 to 1.61; $p = 0.5$) (see *Appendices 7* and *8*). Amenorrhoea rates were similar at all time points, although the overall estimate at 12 months, along with the estimate for HMB, was somewhat inconsistent (heterogeneity: $p = 0.05$ and 0.06, respectively; $I^2 = 59\%$ and 55%, respectively). Changes in bleeding scores favoured EA at 12 months only (WMD 38 points; 95% CI 15 to 60 points; $p = 0.0009$), a result consistent over all studies (heterogeneity: $p = 0.5$; $I^2 = 0\%$). Other outcome measures could not separate the two treatments. Two studies[107,109] provided SF-36 changes from baseline scores, and no differences were found in any of the domains. The number of women subsequently undergoing a hysterectomy was slightly higher for Mirena, although total numbers in this comparison were very small; rates at 12 months were 2.3% (2/86) for EA and 6.7% (6/89) for Mirena. A high proportion of women originally prescribed Mirena discontinued use of this treatment – 15.7% (30/191) at 12 months, rising to 27.6% (29/105) by 2 years. Reported adverse events were low with Mirena; around only 3% reported an expelled/migrated coil within the first month. These results were from studies of first- and second-generation studies combined where first-generation data existed, and were consistent over both types of EA.

Discussion

Principal findings

In this review, access to IPD enabled a more rigorous analysis than is possible from published data from trials comparing second-line treatments for HMB. The primary outcome measure of dissatisfaction was shown to be strongly related to increased QoL. Based on direct and indirect

comparisons using all available data, the review found that both first- and second-generation EA techniques were associated with greater dissatisfaction than hysterectomy, although rates were low for all treatments and absolute differences were small. Recovery times and length of hospital stay were longer for hysterectomy. Dissatisfaction levels with second-generation techniques were slightly lower than those associated with first-generation techniques. In addition, second-generation methods of EA were quicker, had faster recovery times, were associated with fewer adverse procedural events and could be offered under local anaesthetic. Fewer women subsequently underwent hysterectomy after second-generation EA than with first-generation EA, but this difference was not statistically significant. Shorter uterine cavity length was associated with lower levels of dissatisfaction for second-generation EA. Comparisons of EA with Mirena suggest comparable efficacy, although studies involving the latter treatment were generally small and consequently imprecise. Substantial discontinuation of Mirena use was noted and makes interpretation of findings for this treatment difficult.

Strengths and limitations of the review

We used optimal methodology, complying with guidelines on reporting of systematic reviews and meta-analyses.[113] An extensive literature search was conducted, with no language restrictions, minimising the risk of missing information. The collection of IPD allowed us to use previously unreported data, improve the assessment of study quality, standardise outcome measures, undertake ITT analysis and use optimal analytical methods. Subgroup, repeated measures and multivariable analyses would not have been possible without the collection of IPD, and, along with the indirect measures analysis, previously have not been reported.

The review was hampered by the unavailability of IPD from at least 35% of randomised women, which could not be accessed as a number of triallists did not agree to collaborate or could not be contacted. Received data were sometimes incomplete and on occasion failed quality checks and so were unusable. The review's inferences are also limited by the inconsistent outcome measure used across trials; studies involving ED and Mirena focused on comparing reduction in bleeding, while hysterectomy trials focused on patient satisfaction, QoL and resource usage.

Interpretation

In this review we found that more women were dissatisfied following EA than following hysterectomy, although this should be placed in context of longer operating time, total hospital stay and recovery period for hysterectomy. Rates of dissatisfaction are relatively low for EA and it is an effective alternative for women with abnormal uterine bleeding who do not seek amenorrhoea. While this review has shown that hysterectomy is a relatively safe operation, other studies with a more comprehensive follow-up of large populations have shown higher levels of morbidity following hysterectomy.[6] In contrast, EA has low rates of complication.[47] All these factors need to be taken into consideration when considering any potential benefit of hysterectomy.

We found that second-generation techniques, such as thermal balloon ablation (ThermaChoice and Cavaterm),[51–53] the NovaSure device[56] or microwave (Microsulis),[54,68] were not significantly different to first-generation techniques in terms of patient dissatisfaction. Moreover, they are simpler and quicker, require less skill on the part of the operator and can be attempted under local anaesthetic. Importantly, fewer operative complications have also been recorded. Thus, they are clearly preferable to first-generation techniques. The association of shorter uterine

cavity length and lower dissatisfaction with second-generation EA could be because endoscopic treatment is technically more difficult, although given the borderline statistical significance it could also have arisen by chance.

The comparisons involving Mirena were encouraging and, given that it is a relatively cheap and minimally invasive procedure, it could be considered first if drug treatment for heavy bleeding fails.[114] It may even be an alternative to oral drug treatment as a first-line agent, but we did not address this question in our review. However, the current body of evidence comparing Mirena with more invasive techniques is limited and prohibits us from making any strong conclusions about the current findings of this treatment. Furthermore, research on Mirena presents some specific difficulties in interpretation owing to the high proportion of women discontinuing treatment. This can be seen in the trial by Hurskainen *et al.*,[93] which compared Mirena with hysterectomy. While the study was well conducted and reported, the lack of further investigation into the analysis of the primary outcome measure (EQ-5D QoL measure) made the interpretation that there was no evidence of a difference questionable. Of the women allocated Mirena, 20% had received hysterectomy before the main analysis time point at 12 months, with a further 12% no longer using the Mirena. Unfortunately, missing IPD from this trial meant we could not examine further.

Implications for practice

Our review provides evidence that hysterectomy reduces dissatisfaction compared with EA, and this information should be used as part of a consultation with women making a choice about treatment options when initial drug treatment fails to control HMB. EA is satisfactory for a very high proportion of women, but, if complete cessation of bleeding is sought, then hysterectomy may be offered.

Despite the relative paucity of trials evaluating Mirena (particularly in comparison with hysterectomy), it is available in primary care and is less invasive than surgical options. In view of this we can concur with a recent NICE recommendation that women should be offered Mirena before more invasive procedures.[15]

Implications for research

This review has shown that further investment in an RCT comparing hysterectomy with second-generation EA would be of limited value given the similar efficacy of first- and second-generation techniques. Questions remain about the long-term clinical effectiveness of all the treatments; evidence from trials with longer term follow-up (≥ 4 years) is limited to a handful of studies involving differing comparisons.[92,93,115,116] Mirena in particular versus alternative forms of surgical treatment requires further research. While the small studies included in this review have indicated promising results for this treatment, the substantial levels of non-compliance makes interpretation of outcomes difficult and casts some doubt on the equivalent efficacy conclusions.

Individual patient data meta-analysis is an extremely powerful tool if used correctly[117] and provides the most definitive synthesis possible of the available evidence. Such collaborative meta-analyses are well established in cancer and have greatly influenced clinical practice, resulting in striking improvements in, for example, breast cancer survival.[84] Clinicians in speciality groups, such as gynaecology, need to be aware that contributing study results to an IPD is certainly as important as conducting the original research, if not more so. Consensus on optimal outcome measures would also be helpful for meta-analysis.

Chapter 3

Long-term sequelae following hysterectomy or endometrial ablation in Scotland

Introduction

The last two decades have seen the emergence of EA as a conservative surgical alternative to hysterectomy. While the results of hysterectomy are good and amenorrhoea is guaranteed, hysterectomy is invasive and can carry significant short-term morbidity.[31] Overall, 1 in 30 women suffers a major adverse event, and the mortality rate is 0.4–1.1 per 1000 operations. The need for general anaesthesia, prolonged hospital stay and delayed recovery also makes hysterectomy potentially a more expensive treatment.[32] Recent national data from England suggest that EA is now more common than hysterectomy for HMB and second-generation methods are now more commonly performed than hysteroscopic EA.[30]

Endometrial ablative techniques do not guarantee amenorrhoea, but their effectiveness (in comparison with hysterectomy) has been demonstrated in a number of RCTs[41–44] (Aberdeen Endometrial Ablation Trials Group, 1999[46]). National audits of first-generation EA (Scottish Hysterectoscopy Audit Group, 1995[48]) have revealed a number of short-term complications including uterine perforation, fluid overload, cervical laceration, false passage creation, haemorrhage, sepsis and bowel injury.[49,50] Mortality from these techniques has been estimated at 0.26 per 1000.[47,48] Second-generation ablative techniques represent simpler, quicker and potentially more efficient means of treating menorrhagia, which require less skill on the part of the operator but are associated with complications such as equipment failure, uterine infection, perforation, visceral burn, bleeding and cyclical pain. A number of randomised trials indicate that these procedures appear to be as effective as first-generation ablative techniques.[61]

Studies on outcomes after hysterectomy have mainly concentrated on short-term outcomes (in unselected groups of women undergoing the procedure for different underlying reasons[118]). Similarly, most evaluative studies on first- and second-generation EA have reported short-term complications, although some have included medium- and long-term outcomes (there have been few long-term controlled comparisons of hysterectomy with ablation techniques in women with HMB[116]).

Objective

To determine, using population-based data from record linkage, long-term effects of ablative techniques and hysterectomy in terms of failure rates and complications.

Research questions

1. What is the risk of further gynaecological surgery following EA compared with that following hysterectomy in women with HMB?

2. What is the risk of further gynaecological surgery following different types of hysterectomy in women with HMB?
3. What is the risk of gynaecological cancer following EA compared with that following hysterectomy in women with HMB?
4. What is the association of age with risk of further gynaecological surgery following EA compared with that following hysterectomy in women with HMB?

Methods

Anonymised patient-based data for inpatient and day case activity from the whole of Scotland which are routinely collected as Scottish Morbidity Returns (SMR) by the Information Services Division (ISD) were used for this study. More information on the ISD is available at www.isdscotland.org. The SMR register is subjected to regular quality assurance checks and has been shown to be more than 99% complete since the late 1970s.[119] Approval to perform the study was granted by the Privacy Advisory Committee of the ISD. As researchers had no access to any patient identifiers, the North of Scotland Research Ethics Service were of the opinion that formal ethical approval was not necessary.

The original database supplied by ISD contained a total of 549,223 records. From this, records with an *International Classification of Diseases*, Tenth Edition (ICD-10) diagnostic code beginning with either N92 (excessive, frequent and irregular menstruation) or N93 (other abnormal uterine and vaginal bleeding excluding neonatal vaginal haemorrhage and pseudo menses) plus any record with ICD, Ninth Edition (ICD-9) codes of –6226, –6270 and –6268 were selected. This identified 61,880 records (29,100 records with relevant ICD-10 codes and 32,780 with relevant ICD-9 codes). A total of 791 subjects who were aged < 25 or > 55 years were then excluded to avoid clinically implausible diagnoses and those unlikely to have dysfunctional bleeding, leaving 61,089 records. Thirty-seven records were subsequently excluded from women recorded as implausibly having EA following a hysterectomy, leaving 61,052 records. We also excluded 9891 women who had a hysterectomy before 1989 (after making sure that no EA was lost) to ensure a comparable time frame (1989–2006) of initial operation in both the EA and hysterectomy groups. This left 51,198 records (14,078 in the EA group and 37,120 in the hysterectomy group). A total of 2779 women (19.7%) in the EA group went on to have a hysterectomy. We excluded these women since it would be difficult to determine whether subsequent sequelae should be attributed to the initial EA or to the subsequent hysterectomy. The median (interquartile range, IQR) duration between the date of EA and subsequent hysterectomy for these 2779 women was 1.25 (0.66–2.67) years. Similarly, of the original 14,078 women undergoing EA, 379 (2.7%) went on to have a repeat EA procedure within a median (IQR) of 1.17 (0.66–2.83) years. These 379 women have been retained in the EA group.

Following the exclusions detailed above, we were left with a data file containing 48,419 analysable records from women aged 25–55 years who had either an EA ($n = 11,299$) or a hysterectomy ($n = 37,120$) as a primary surgical procedure for dysfunctional uterine bleeding between 1989 and 2006. Note that it was not possible to discriminate between different types of EA owing to inconsistencies with coding in the early years following the introduction of the new technology.

The ISD then linked these 48,419 women to the cancer registry, and ICD-9 and ICD-10 codes corresponding to gynaecological cancers (breast, vaginal, cervical, uterine and ovarian) diagnosed between 1989 and 2006 were made available for analysis. All cancers with a date of diagnosis subsequent to the date of EA or hysterectomy were included in the analysis.

Statistical analysis

Socioeconomic status was assessed using the Carstairs index,[120] which was divided into quintiles for analysis. Descriptive statistics (percentage, *n*, mean and SD, median and IQR as appropriate) were used to summarise each of the surgical outcomes and potential predictor variables (age, duration of follow-up and Carstairs quintile) in the EA and hysterectomy groups. Appropriate univariate analyses (chi-squared test for comparing two categorical variables, *t*-test to compare means and the Mann–Whitney *U*-test to compare medians) across the hysterectomy and EA groups were performed.

Cox proportional hazards regression analysis was used to examine the survival experience for different surgical outcomes in the hysterectomy and EA groups. Hazard ratios and their 95% CI for different outcomes were calculated both before and after adjustment for age, year of primary operation and Carstairs quintile. Kaplan–Meier survival curves for outcomes that were significantly different between groups following adjustment were plotted and the assumption of a constant hazard ratio over time was checked. Similar survival analysis was performed comparing cancer outcomes between the EA and hysterectomy groups and then comparing surgical outcomes between different types of hysterectomy. The association of age with risk of further surgical procedures in the EA and hysterectomy groups was examined by inclusion of an age by group interaction term in the regression model along with the main effects.

International Classification of Diseases codes used in the analysis

Hysterectomy codes used

Hysterectomy	Codes
Any hysterectomy	q07.4, q07.8, q07.9, q07.5, q08.8, q08.9
Any hysterectomy + oophorectomy (bi + removal of only ovary)	(q07.4, q07.8, q07.9, q07.5, q08.8, q08.9) + (q 22.1, q22.3, q23.2, q23.6)
Any total hysterectomy	q 07.4, q07.8, q07.9
Any subtotal hysterectomy	q07.5
Any vaginal hysterectomy	q08.8, q08.9

Operation codes used

Operations	Codes
Oophorectomy	q232, q236, q432
Ovarian surgery	q438, q439, q473, q474, q478, q479, q491, q493, q498, q499
Hysterectomy	q078, q079, q088, q089, q072, q074, q075, q082, q083
Uterine operations	q093, q098, q099, q103, q108, q109, q161, q168, q169
Repeat ablations	q171
Adnexal surgery	q228, q229, q238, q239, q248, q249
Vaginal repair	m531, p228, p229, p238, p239
Tension free vaginal tape (TVT)	m538
Vault repair	p241, p243, p244, p248, p249
Fistula repair	p251, p252, p253, p254
Colporrhaphy (anterior or posterior vaginal repair)	p221, p222, p223, p231, p232, p233

Cancer codes used

Cancer type	ICD-10	ICD-9
Breast	C50	174
Vagina	C52	184
Cervix	C53	180
Uterine	C54	182
Ovary	C56	183

Results

Between 1989 and 2006, 37,120 Scottish women underwent hysterectomy and 11,299 underwent EA as a primary surgical procedure for dysfunctional uterine bleeding (*Table 6*). Women who underwent EA were significantly older and belonged to a higher socioeconomic group than women who underwent hysterectomy. The median duration of follow-up in women post ablation was shorter, reflecting the increased numbers of ablations performed in more recent years.

Table 7 shows the different types of hysterectomy performed for HMB. Women who underwent bilateral oophorectomy were significantly older than those whose ovaries were conserved ($p < 0.001$).

Table 8 lists the frequencies and hazard ratios for further surgical outcomes following either EA or hysterectomy. Women were significantly more likely to require further gynaecological surgery after EA (adjusted hazard ratio 3.56; 95% CI 3.26 to 3.89) than after hysterectomy. Most

TABLE 6 Baseline characteristics of women in the two groups

		Endometrial ablation ($n = 11,299$)	Hysterectomy ($n = 37,120$)	p-value
Age at treatment (years), mean (SD)		42.5 (5.6)	41.0 (6.0)	< 0.001
Duration of follow-up (years), median (IQR) minimum, maximum		6.2 (2.7–10.8), (0.002, 17.91)	11.6 (7.9–14.8), (0.002, 17.91)	< 0.001
Carstairs quintile, *n* (%)	1	2941 (26)	5617 (15)	< 0.001
	2	2266 (20)	6870 (19)	
	3	1905 (17)	7682 (21)	
	4	1957 (19)	7814 (21)	
	5	2160 (20)	8669 (24)	

TABLE 7 Types of hysterectomy[a]

	n (%)	Age in years, mean (SD)
Hysterectomy with conservation of ovaries	20,864 (56)	39.1 (5.5)
Any hysterectomy + oophorectomy (bi + removal of only remaining ovary)	13,036 (35)	44.2 (5.6)
Total hysterectomy	28,961 (78)	41.1 (6.1)
Subtotal hysterectomy	2948 (8)	41.5 (5.8)
Vaginal hysterectomy	5211 (14)	40.4 (5.9)

a The ICD codes for each of these five types of hysterectomy correspond with those listed in the statistical analysis section. Note that these subcategories of hysterectomy are not mutually exclusive, hence the total is > 37,120.

TABLE 8 Surgical outcomes following endometrial ablation and hysterectomy

	EA, *N* = 11,299 (*n* %)	Hysterectomy *N* = 37,120 (*n* %)	Unadjusted hazard ratio (95% CI)	Adjusted[a] hazard ratio (95% CI)
All gynaecological surgery	962 (8.5)	1446 (3.9)	3.60 (3.31 to 3.91)	3.56 (3.26 to 3.89)
Adnexal surgery	37 (0.3)	277 (0.8)	0.64 (0.45 to 0.90)	0.80 (0.56 to 1.15)
Pelvic floor repair	102 (0.9)	817 (2.2)	0.68 (0.55 to 0.84)	0.62 (0.50 to 0.77)
Intrauterine procedures	577 (5.1)	–	–	–
Repeat EA	278 (2.5)	–	–	–
TVT	52 (0.5)	388 (1.1)	0.82 (0.62 to 1.11)	0.55 (0.41 to 0.74)
Genital fistula repair	3 (0.03)	61 (0.2)	0.18 (0.05 to 0.56)	0.18 (0.06 to 0.58)

a Adjusted for age, year of primary operation and Carstairs quintile with hysterectomy as the base group.

of these further procedures were intrauterine procedures (such as dilatation, curettage and hysteroscopy) and repeat EA. Women who underwent EA were less likely to undergo pelvic floor repair (adjusted hazard ratio 0.62; 95% CI 0.50 to 0.77), TVT (adjusted hazard ratio 0.55; 95% CI 0.41 to 0.74) or genital fistula repair (adjusted hazard ratio 0.18; 95% CI 0.06 to 0.58) than the hysterectomy group. *Figures 7–9* show Kaplan–Meier survival curves of all gynaecological surgery, pelvic floor repair and TVT, respectively. The survival curves for genital fistula repair are not shown owing to the small number of women who have this outcome, in the EA group in particular.

Cancer outcomes

Table 9 shows both the incidence of gynaecological cancers in women following hysterectomy or EA and the results of the survival analyses. Overall, the number of women diagnosed with cancer was small. The adjusted hazard ratio of breast cancer was higher and that of ovarian cancer was lower for women in the EA group compared with the hysterectomy group, but neither reached statistical significance. For illustration purposes, we calculated the number of EAs needed to treat to give one extra cancer compared with the hysterectomy group. For breast cancer, the number needed to treat (NNT) was 237 and for ovarian cancer it was 1073.

Outcomes following endometrial ablation versus different types of hysterectomy

Table 10 shows the number and percentage of women having further surgery in the EA group compared with those in the group undergoing different types of hysterectomy. The unadjusted and adjusted hazard ratios are then presented in *Table 11*. The risks of future gynaecological surgery in these subgroups are broadly similar to those seen in the previous comparison between ablation and hysterectomy (all types) combined. Women in the EA group had a significantly lower adjusted risk of having adnexal surgery than women who had undergone hysterectomy with conservation of ovaries (adjusted hazard ratio 0.65; 95% CI 0.45 to 0.95). The adjusted hazard ratios of pelvic floor repair, TVT and genital fistula repair were all significantly lower in the EA group than in women who had undergone vaginal hysterectomy. Similarly, the adjusted hazard ratios of undergoing either a TVT procedure or a genital fistula repair were all lower in women who had undergone EA than in those having either hysterectomy with ovarian conservation, hysterectomy with oophorectomy, total hysterectomy or vaginal hysterectomy.

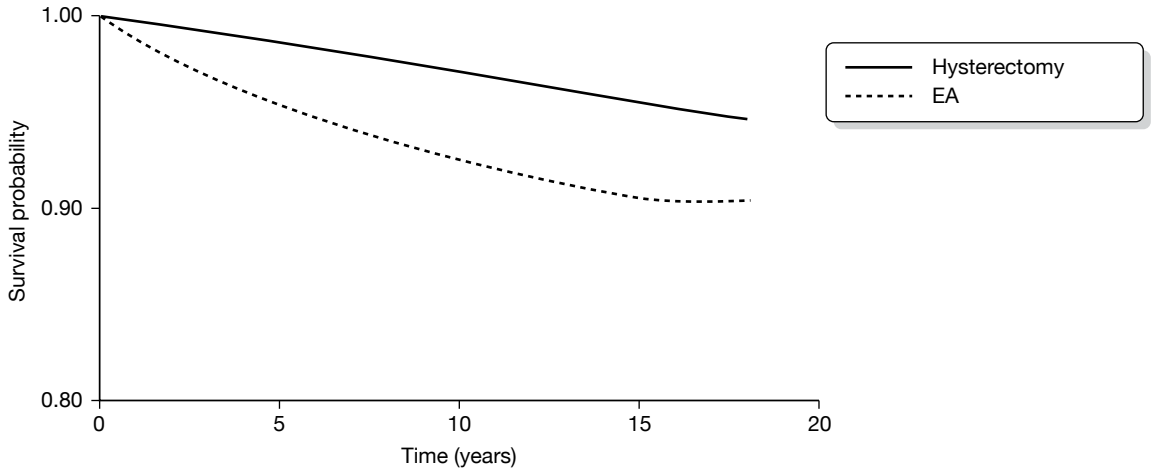


FIGURE 7 Kaplan–Meier survival curve of all gynaecological surgery among the EA and hysterectomy groups.

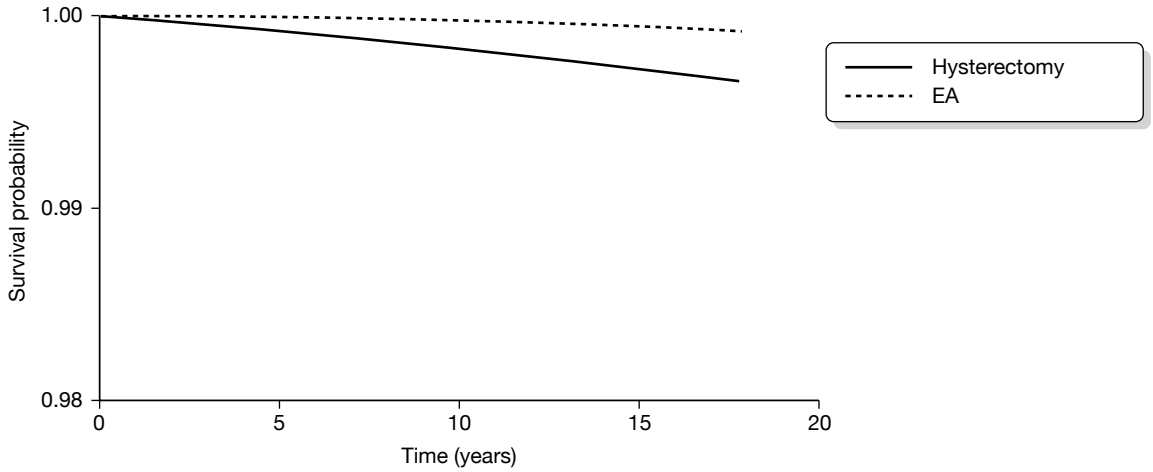


FIGURE 8 Kaplan–Meier survival curve of pelvic floor repair among the EA and hysterectomy groups.

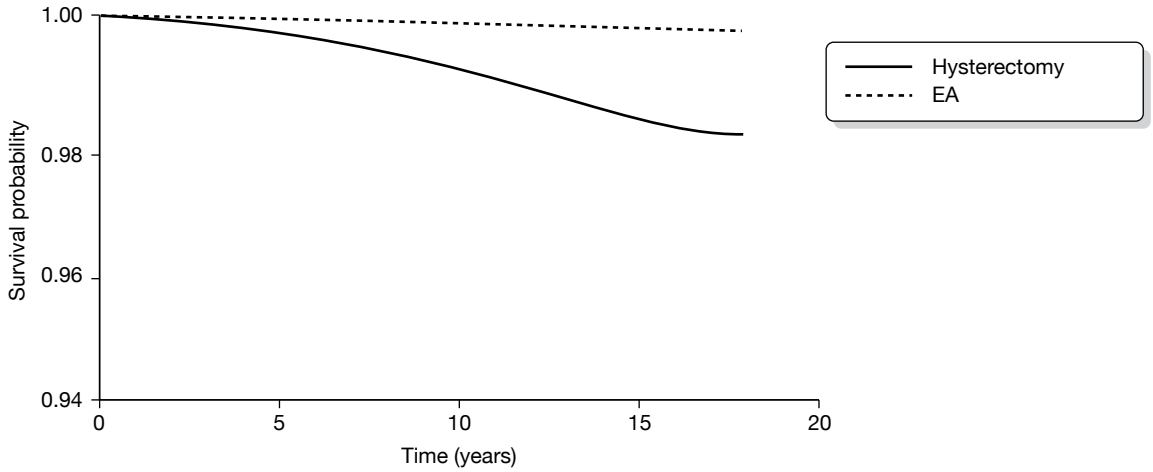


FIGURE 9 Kaplan–Meier survival curve of TVT among the EA and hysterectomy groups.

TABLE 9 Cancer outcomes following EA and hysterectomy

	EA, *N* = 11,299 (*n* %)	Hysterectomy, *N* = 37,120 (*n* %)	Unadjusted hazard ratio (95% CI)	Adjusted[a] hazard ratio (95% CI)
Breast cancer	130 (1.15)	584 (1.57)	1.22 (1.01 to 1.49)	1.14 (0.93 to 1.39)
Cervical cancer	4 (0.04)	–	–	–
Endometrial cancer	2 (0.02)	–	–	–
Ovarian cancer	5 (0.04)	51 (0.14)	0.70 (0.28 to 1.78)	0.91 (0.35 to 2.39)
Vaginal cancer	–	4 (0.02)	–	–

a Adjusted for age, year of primary operation and Carstairs quintile with hysterectomy as the base group.

TABLE 10 Number of women (%) having further surgery following EA or different types of hysterectomy

	EA, *n* = 11,299	Hysterectomy with conservation of ovaries, *n* = 20,864	Hysterectomy with oophorectomy, *n* = 13,036	Total hysterectomy, *n* = 28,961	Subtotal hysterectomy, *n* = 2948	Vaginal hysterectomy, *n* = 5211
All gynaecological surgery	962 (8.5)	901 (4.3)	402 (3.1)	1113 (3.8)	70 (2.4)	263 (5.1)
Adnexal surgery	37 (0.3)	217 (1.0)	–	230 (0.8)	16 (0.5)	31 (0.6)
Pelvic floor repair	102 (0.9)	480 (2.3)	281 (2.2)	618 (2.1)	31 (1.1)	168 (3.2)
TVT	52 (0.5)	228 (1.1)	127 (1.0)	294 (1.0)	25 (0.9)	69 (1.3)
Genital fistula repair	3 (0.03)	39 (0.2)	19 (0.2)	49 (0.2)	–	12 (0.2)

TABLE 11 Risk of further surgery following EA versus different types of hysterectomy

	Adjusted hazard ratio (95% CI)[a]				
	Ablation vs hysterectomy with conservation of ovaries	Ablation vs hysterectomy with oophorectomy	Ablation vs total hysterectomy	Ablation vs subtotal hysterectomy	Ablation vs vaginal hysterectomy
All gynaecological surgery	3.27 (2.95 to 3.63)	4.49 (3.96 to 5.07)	3.85 (3.49 to 4.24)	4.30 (3.37 to 5.48)	2.53 (2.19 to 2.92)
Adnexal surgery	0.65 (0.45 to 0.95)	–	0.81 (0.57 to 1.17)	0.75 (0.42 to 1.36)	0.85 (0.51 to 1.41)
Pelvic floor repair	0.51 (0.40 to 0.64)	0.74 (0.58 to 0.94)	0.70 (0.56 to 0.88)	0.94 (0.62 to 1.40)	0.36 (0.28 to 0.46)
TVT	0.52 (0.38 to 0.72)	0.59 (0.43 to 0.84)	0.57 (0.42 to 0.78)	0.62 (0.38 to 1.01)	0.41 (0.28 to 0.61)
Genital fistula repair	0.15 (0.04 to 0.50)	0.23 (0.06 to 0.80)	0.16 (0.05 to 0.52)	–	0.16 (0.04 to 0.59)

a Adjusted for age, year of primary operation and Carstairs quintile with type of hysterectomy as base group.

The hazard ratios of further surgery following different types of hysterectomy are shown in *Table 12*. Women whose ovaries were conserved were significantly more likely to undergo further gynaecological surgery (adjusted hazard ratio 1.39; 95% CI 1.22 to 1.59) as well as pelvic floor repair (adjusted hazard ratio 1.31; 95% CI 1.11 to 1.55) than women who had undergone a hysterectomy and bilateral oophorectomy. In contrast, women having an abdominal total hysterectomy or a subtotal hysterectomy were significantly less likely to undergo either further gynaecological surgery or pelvic floor repair than women undergoing a vaginal hysterectomy. *Figures 10–15* show Kaplan–Meier survival curves for those end points in which there was a significant difference in survival between the different hysterectomy types.

TABLE 12 Adjusted hazard ratio of further surgery following different types of hysterectomy

	Adjusted hazard ratio (95% CI)[a]			
	Hysterectomy with conservation of ovaries vs hysterectomy + bilateral oophorectomy	Total hysterectomy vs subtotal hysterectomy	Abdominal total vs vaginal hysterectomy	Subtotal hysterectomy vs vaginal hysterectomy
All gynaecological surgery	1.39 (1.22 to 1.59)	1.19 (0.93 to 1.53)	0.68 (0.60 to 0.78)	0.54 (0.41 to 0.71)
Adnexal surgery	–	0.99 (0.59 to 1.66)	1.36 (0.93 to 1.99)	1.07 (0.57 to 2.00)
Pelvic floor repair	1.31 (1.11 to 1.55)	1.31 (0.91 to 1.90)	0.54 (0.45 to 0.64)	0.39 (0.26 to 0.58)
TVT	1.09 (0.85 to 1.39)	1.09 (0.71 to 1.66)	0.77 (0.59 to 1.01)	0.69 (0.43 to 1.11)
Genital fistula repair	1.39 (0.76 to 2.55)	–	0.72 (0.38–to 1.36)	–

a Adjusted for age, year of primary operation and Carstairs quintile with the second listed group as base group.
Women who underwent both or neither types of hysterectomy being compared were excluded from the model on a pairwise basis.

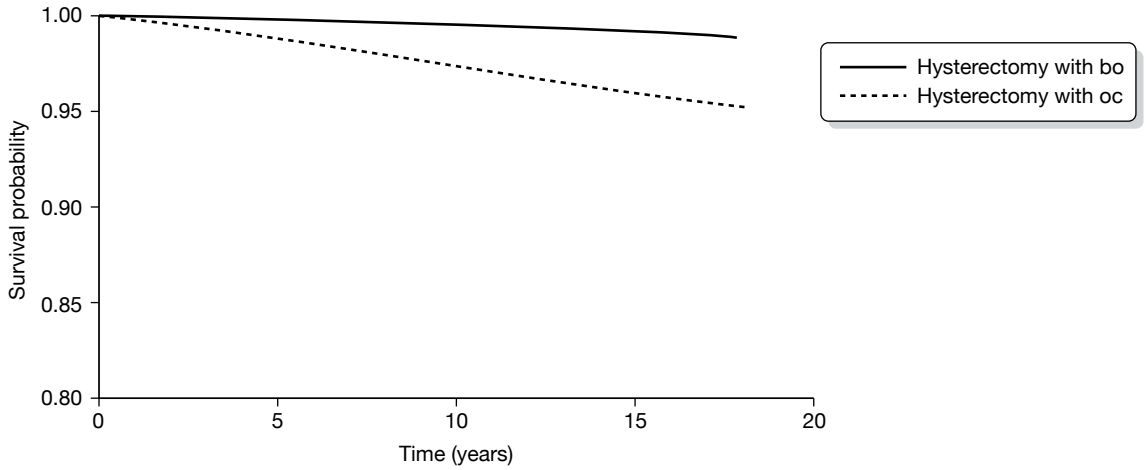


FIGURE 10 Kaplan–Meier survival curve of all gynaecological surgery among the hysterectomy with ovarian conservation (oc) and hysterectomy with bilateral oophorectomy (bo) groups.

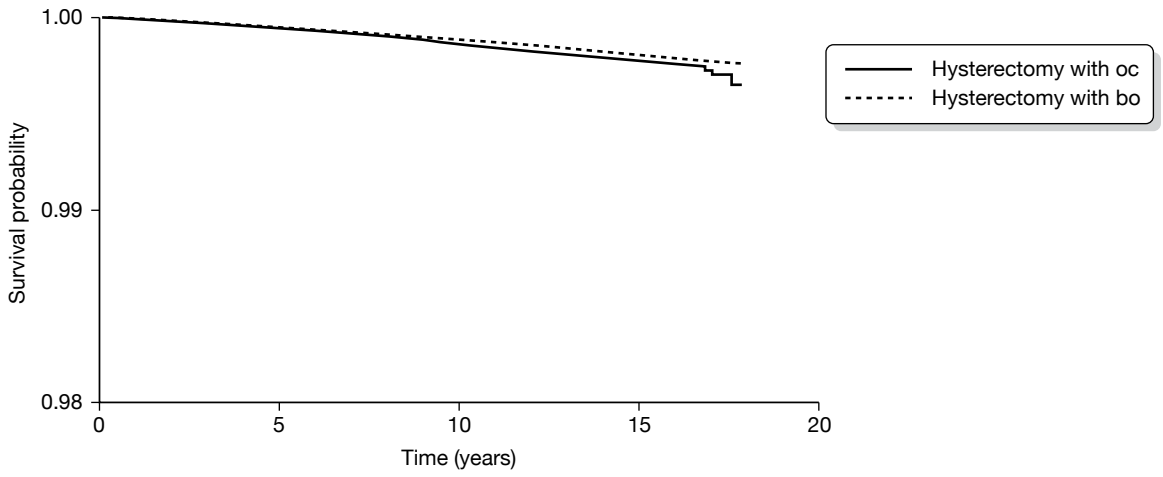


FIGURE 11 Kaplan–Meier survival curve of pelvic floor repair among the hysterectomy with ovarian conservation (oc) and hysterectomy with bilateral oophorectomy (bo) groups.

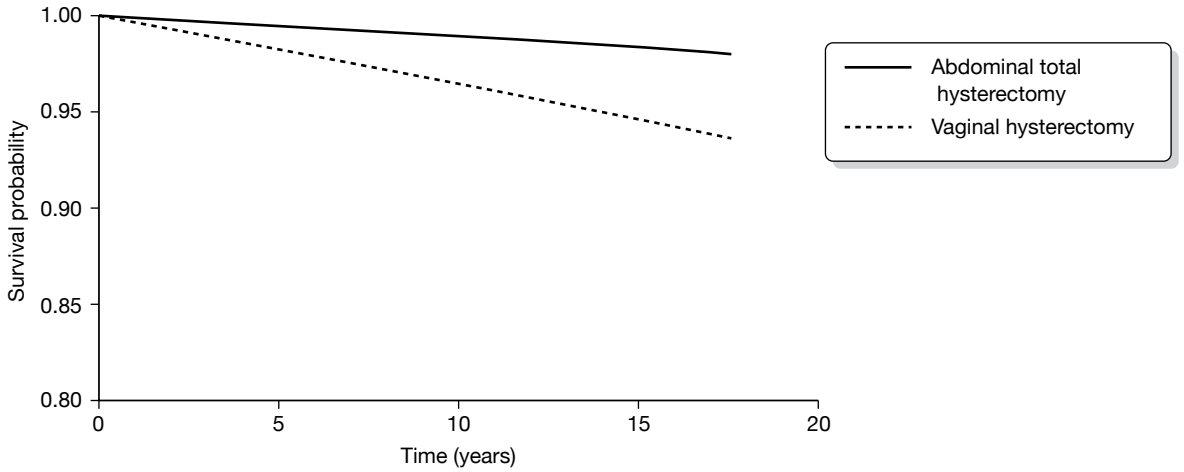


FIGURE 12 Kaplan–Meier survival curve of all gynaecological surgery among the abdominal total and vaginal hysterectomy groups.

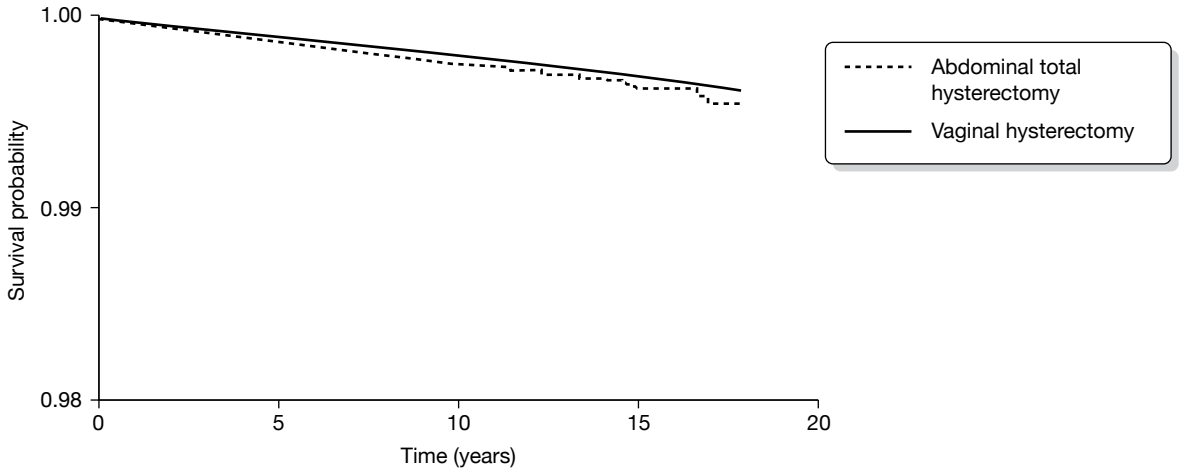


FIGURE 13 Kaplan–Meier survival curve of pelvic floor repair among the abdominal total and vaginal hysterectomy groups.

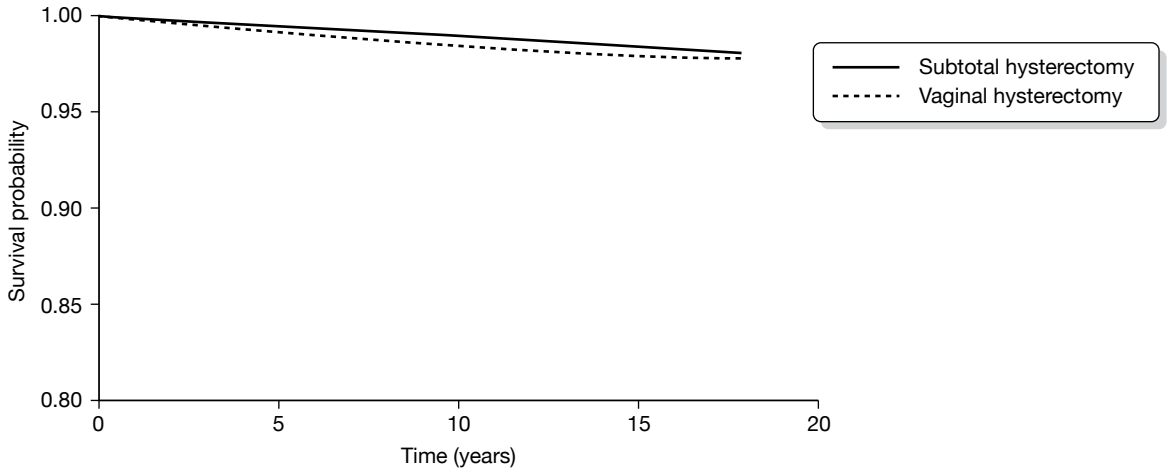


FIGURE 14 Kaplan–Meier survival curve of all gynaecological surgery among the subtotal hysterectomy and vaginal hysterectomy groups.

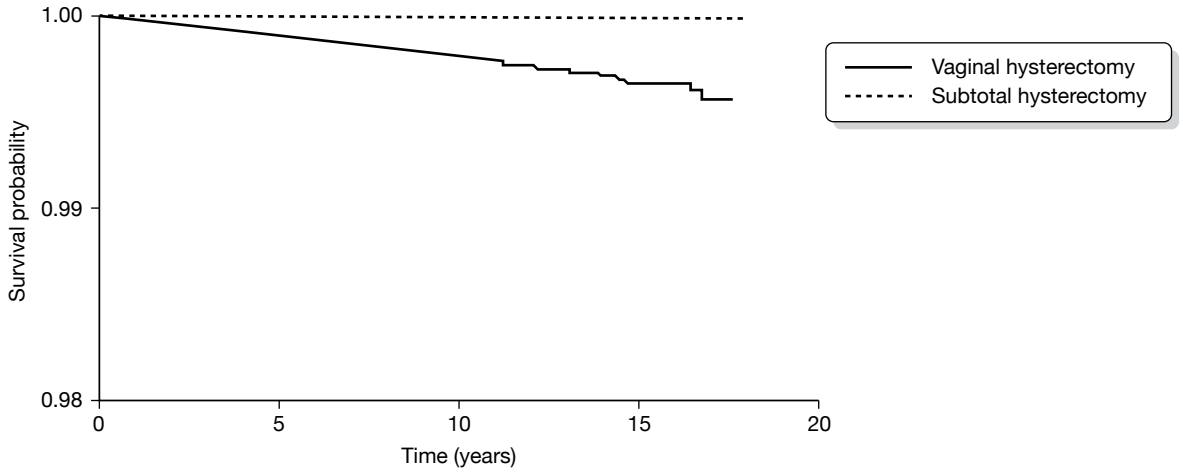


FIGURE 15 Kaplan–Meier survival curve of pelvic floor repair among the subtotal hysterectomy and vaginal hysterectomy groups.

Association of age with risk of further surgical procedures

The association of subsequent surgery with age in the EA and hysterectomy groups was examined by inclusion of an age by group interaction term in the regression model along with the main effects of age and group status. No significant interaction of age with risk of subsequent surgery was found (the *p*-value for the interaction term for all gynaecological surgery was 0.0569; adnexal surgery $p = 0.120$; pelvic floor repair $p = 0.416$; TVT $p = 0.151$ and genital fistula repair $p = 0.515$).

Discussion

Principal findings

The sociodemographic profile of women who underwent either first- or second-generation EA was different to that of those who received hysterectomy for HMB. Hysterectomy was more likely to lead to surgery for pelvic floor repair and for stress urinary incontinence. Vaginal hysterectomy was associated with a higher chance of further surgery and surgery for pelvic floor prolapse compared with hysterectomy carried out through the abdominal route. The incidence of cancers was generally low in both groups (< 1.6%), with endometrial cancer following EA having an incidence of 0.02%.

Strengths and limitations

To our knowledge this is the first large population-based study using national data on long-term outcomes in women who had received alternative surgical treatments for HMB. Use of ICD codes allowed us to define both the cause of HMB as well as the nature of surgery, but, as the diagnosis of dysfunctional uterine bleeding was performed by a process of exclusion, it is possible that the hysterectomy cohort could have included a few women with other causes of HMB.

This was a retrospective cohort study based on routinely collected national data. Like any observational study, it is not free from the usual problems of bias and confounding. Additionally, the analysis was compromised by the limited availability of key socioeconomic as well as clinical variables. Although the numbers of women in the hysterectomy and ablation cohorts were large, a major drawback was our inability to discriminate among the individual types of first- and second-generation EA or adjust for the experience of the operator as has been done in previous

national audits.[47] We were also unable to analyse the long-term outcomes following laparoscopic hysterectomy as numbers were small; these were therefore grouped with abdominal hysterectomy.

Interpretation of findings

Our findings suggest that women from higher socioeconomic groups are more likely to have EA than hysterectomy. This inverse correlation between hysterectomy rates and social class has been noted by some authors[121] but not by others.[122] None of these studies had a control group of women with HMB who went on to have EA.

Our results confirm previous work[46] that suggests that women treated initially with EA for HMB are more likely to need subsequent surgery for the same condition than those treated by hysterectomy, which is a more definitive operation. Although they were excluded from the current analysis, we noted that around one in five women had a subsequent hysterectomy following initial EA, while repeat ablations and exploration of the uterus accounted for a substantial minority of cases. As has been noted previously,[46,112,123] the survival curve indicates that most repeat surgery for persistent HMB occurred within the first 2 years of initial surgery.

Women were more likely to undergo a TVT procedure for stress urinary incontinence after hysterectomy than after EA – corroborating the results of some previous studies that suggested a link between hysterectomy and urinary incontinence.[34,124] The biological justification of this has been debated in the past and could be due to compromise to the pelvic floor caused by surgical damage to muscular, connective or neurological tissue. The lower risk of genital fistula repair after ablation is explained by the higher probability of pelvic organ damage during hysterectomy.

The risk of gynaecological cancer following EA has been identified as a key clinical and research question in the past. Our results are in agreement with Krogh *et al.*,[125] who found no significant increase in the incidence of endometrial cancer after ablation in a Danish population. We found no significant difference in the risk of ovarian cancer between women undergoing EA and those who had a hysterectomy. In contrast, Loft *et al.*[126] reported that the risk of ovarian cancer was lower in women who had hysterectomy (with conservation of at least one ovary) than in those who had not [relative risk (RR) 0.78; 95% CI 0.60 to 0.96]. Owing to an average age of 41 years at initial surgery and the relatively short follow-up (median duration 6.2 and 11.6 years for EA and hysterectomy groups, respectively), many of the women in our study have yet to reach the peak age of incidence for many of the cancers in question and hence the incidence of a common malignancy like breast cancer is low in both groups.

We are able to explore the impact of different types of hysterectomy versus ablation and show that the overall risk of further surgery was significantly higher in the ablation group irrespective of type of hysterectomy. Compared with the ablation group, the odds of TVT were significantly higher in women after a total hysterectomy, but not after subtotal hysterectomy, thus fuelling the ongoing debate on the potential association between total (as opposed to subtotal) hysterectomy and urinary stress incontinence.[123,127–129] Subtotal hysterectomy has other potential disadvantages, as shown in a small series of women: laparoscopic subtotal hysterectomy led to a 22.9% (16/70) chance of potentially difficult further surgery to remove the cervical stump.[130] The chances of undergoing surgery for pelvic floor prolapse and urinary stress incontinence were significantly higher in women who had undergone vaginal hysterectomy than in the ablation group. It is impossible to rule out the possibility that the decision to opt for a vaginal route for a hysterectomy may have been informed by prior knowledge of a degree of descent or pelvic laxity. Comparison of different types of hysterectomy revealed that the vaginal route was more likely to be associated with future surgery for prolapse and incontinence. This is supported by data from Blandon *et al.*,[131] who reported that, in comparison with women without prolapse, women who had a hysterectomy for prolapse were at increased risk for subsequent pelvic floor repair.

In the absence of follow-up data from randomised trials, the existing literature on this subject is conflicting as described by Thakar and Sultan.[123]

There was no significant association between age and risk of further gynaecological operations.

Clinical implications

The results of this study provide clinicians and women with useful data on expected outcomes after EA which can be used to counsel women regarding options of treating HMB. The lower perioperative complications of the less invasive ablation procedures need to be balanced against the 8% chance of repeat surgery for the same symptoms, although the chance of long-term pelvic floor problems may be less. It is also useful to confirm data from other follow-up studies on smaller cohorts that indicate that most repeat procedures occur within 2 years of the initial operation. Our data are broadly reassuring in terms of identifying the risk of endometrial cancer after ablation.

Research implications

This study underlines the limitations of the available literature in this field, which include retrospective studies that have been prone to selection and reporting bias as well as lack of data on key confounders. In the absence of adequately powered large randomised trials with high rates of follow-up, well planned prospective cohort studies with pre-determined end points are needed. It is also important to consider the need for large national audits of EA especially now that new second-generation ablation technologies are being adopted.

Chapter 4

Health economics

Background

The objectives of the economic evaluation were to undertake a cost-effectiveness analysis of hysterectomy versus first- and second-generation ablative techniques and Mirena as additional comparisons were agreed at the request of the National Co-ordinating Centre for Health Technology Assessment prior to funding.

The purpose of the economic evaluations is to inform current treatment policy in this clinical area, while the value of information component will serve to highlight future research needs and agendas, and inform possible future research funding decisions.

The model development process, planned to use as a starting point, was the recently published menorrhagia clinical pathway Markov model.[16] That model, from researchers at the University of Exeter, formed the basis of the national coverage decision by NICE on microwave and TBEA for HMB.[16] As part of model development, the requirements for structural model adjustment were determined through consultation within the research team, drawing on the requisite clinical and modelling expertise.

The model developed by Garside *et al.*[16] was a state transition (Markov) model using Microsoft EXCEL (Microsoft Corporation, Seattle, WA, USA). Their structure was informed by clinical input and the model examined the progress of five hypothetical cohorts of women with HMB who were treated separately with either TBEA, MEA, TCRE, RBEA or hysterectomy. Their evaluation took the perspective of the NHS and the outcomes were presented in quality-adjusted life-years (QALYs). The basic structure and many of the main assumptions from the model-based evaluation of Garside *et al.*[16] formed the basis of our model and assumptions in our model and these will be referred to throughout the report. Refinements that led to the current model structure are described throughout the *Methods* section but are relatively modest and represent a pragmatic adjustment to available data. The key difference in the data used to inform the current model was additional data from the IPD meta-analysis and the addition of the Mirena strategy.

The principal clinical data used in populating the model were drawn from other aspects of our research work on this project, namely the individual patient meta-analyses, data from national registers and existing RCTs.

Methods

The cost-effectiveness component of the work reports the results in terms of incremental cost per QALY gained based on QoL data available from published sources.[132] The presentation of results in QALYs allows comparison of the results with other available and recently published studies.[16] Resource use was estimated from the existing published evidence and additional cost data from other sources such as the annual review of unit health and social care costs (Personal Social Services Research Unit) and national schedule for reference costs.

The results of the analysis are presented using cost-effectiveness acceptability curves to reflect sampling variation and uncertainties in the appropriate threshold cost-effectiveness value. In addition to probabilistic sensitivity analyses on the base case model, we have included a range of alternative analyses to explore the robustness of these results to plausible variations in key assumptions and variations in the analytical methods used. We also carried out a value of information analysis to explore the extent of uncertainty in the model-based economic evaluation. Such analyses can provide an estimation of whether it is likely that the removal of existing uncertainty (by seeking additional data from additional studies, for example) would impact on the results in a way that would change decisions based on those results.

Cost-effectiveness model

A state transition (Markov) model was developed by the authors using Microsoft EXCEL. The structure was informed by clinical input. The model examines the progress of four hypothetical cohorts of women with HMB who are treated by one of four alternative strategies: Mirena coil; first-generation EA techniques; second-generation EA techniques; or hysterectomy. Given the reliance on secondary data and the nature of available data, the model-based economic evaluation takes the form of a cost–utility analysis and was carried out from the perspective of the NHS in a secondary care setting. An incremental approach was used for the reporting of the results.

Structure of the economic model

In the model, a cohort of 10,000 eligible women was compared for each strategy (Mirena coil; first-generation EA; second-generation EA; and hysterectomy). The starting age of women in the model is 42 years, and the model runs for a total of 10 years and assumes that all women will become menopausal at the age of 52 years, (the average age of menopause in the UK). These assumptions are those used by Garside *et al.*[16] Each model cycle is 1 month long and represents a typical menstrual cycle. The death rate from causes other than procedures for HMB was based on values for women in the Government Actuary's Department life tables of England and Wales for the years 1998–2008.[133]

The model is based on the clinical pathways presented in *Figure 16*. The pathway for women undergoing any EA technique (first or second generation) is shown in *Figure 17*. The pathway for women undergoing hysterectomy is shown in *Figure 18*. Health states are shown in boxes and arrows show the transitions that can occur. The health states and pathways are the same for both types of EA technique.

Definition of health states for endometrial ablation pathways

- Menorrhagia – all women in the cohort have pre-operative HMB.
- EA techniques – women undergo EA by first- or second-generation techniques.
- Complications – following EA, some women will experience severe postoperative complications. Perioperative complications are included in the EA state.
- Well – following EA, complications or treatment failure, women are satisfied with treatment.
- Symptomatic – following EA, complications or well, HMB may recur (treatment failure) at any time. Women may be retreated (repeat ablation), become 'well' or have a hysterectomy after initial or repeat ablation.
- Repeat ablation (RB) – if HMB recurs postoperatively, women may choose to have a second EA which occurs 6 months after the initial ablation. Only one repeat EA is permitted and it is always a first-generation technique (RB).
- Hysterectomy – if women become symptomatic after the first ablation, they may choose to have a hysterectomy. Hysterectomy is also an option after a failed repeat EA. This operation occurs 6 months after the decision.

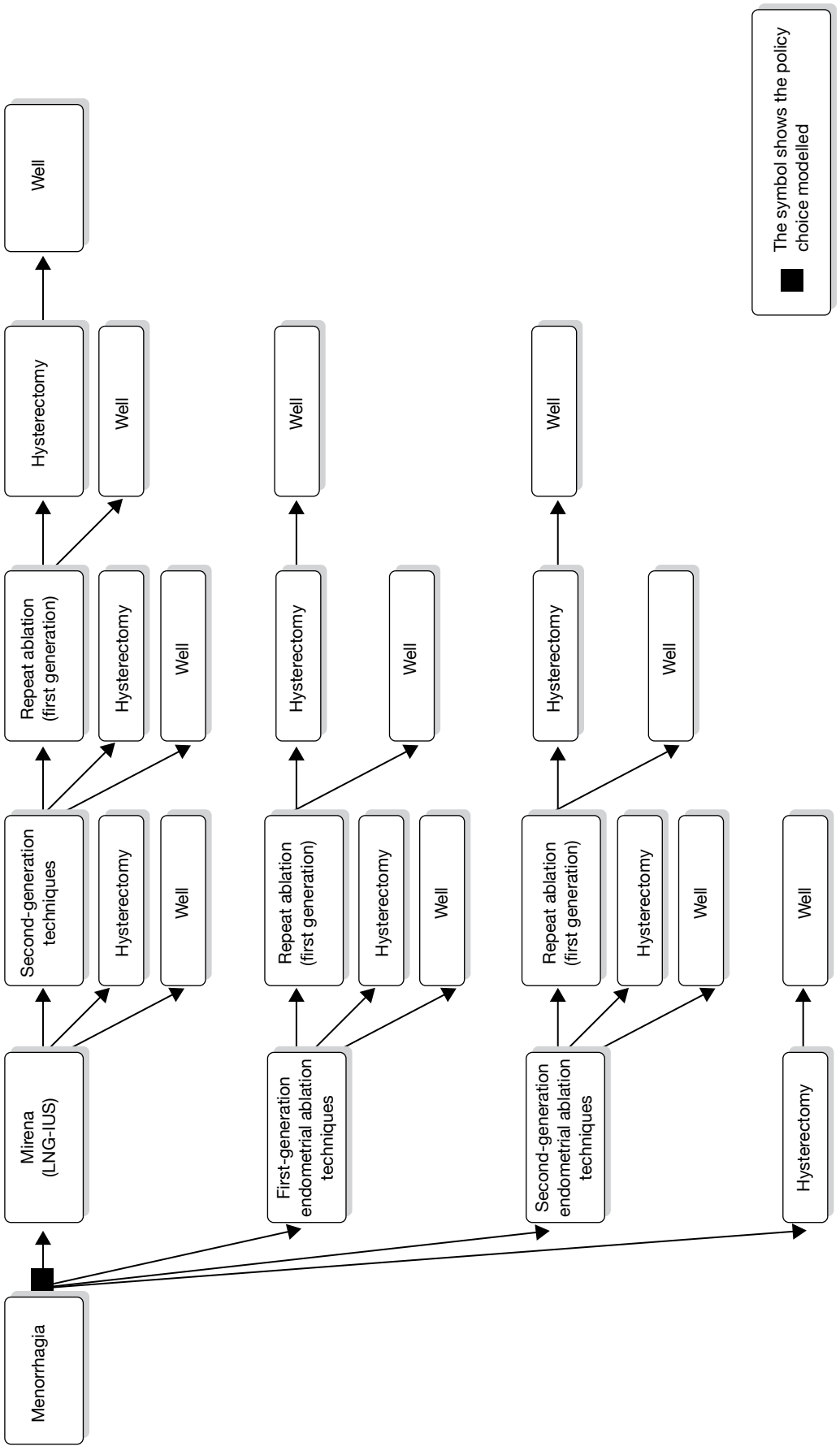


FIGURE 16 Summarised version of clinical pathway in the model.

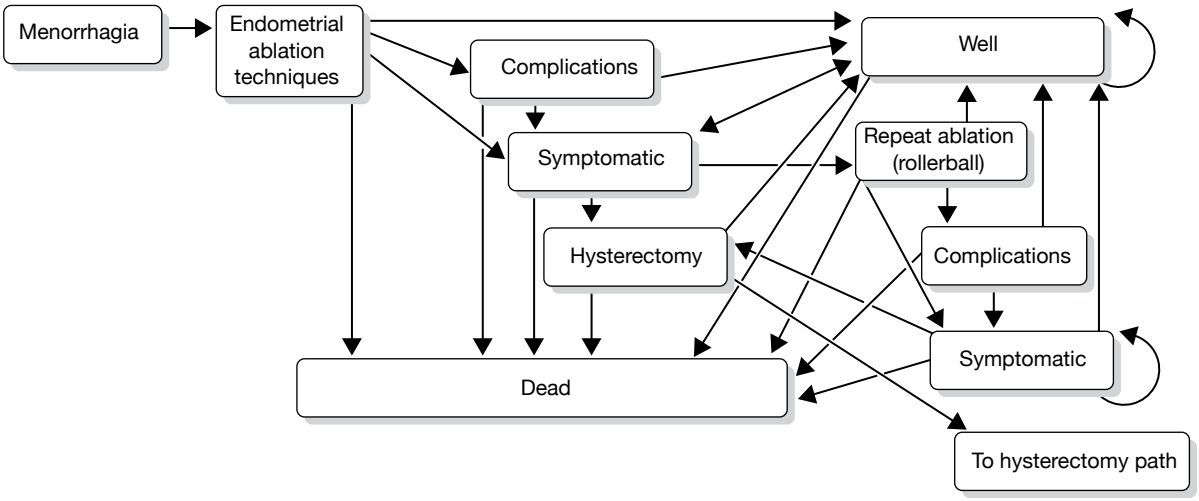


FIGURE 17 Clinical pathway for EA (first- or second-generation techniques).

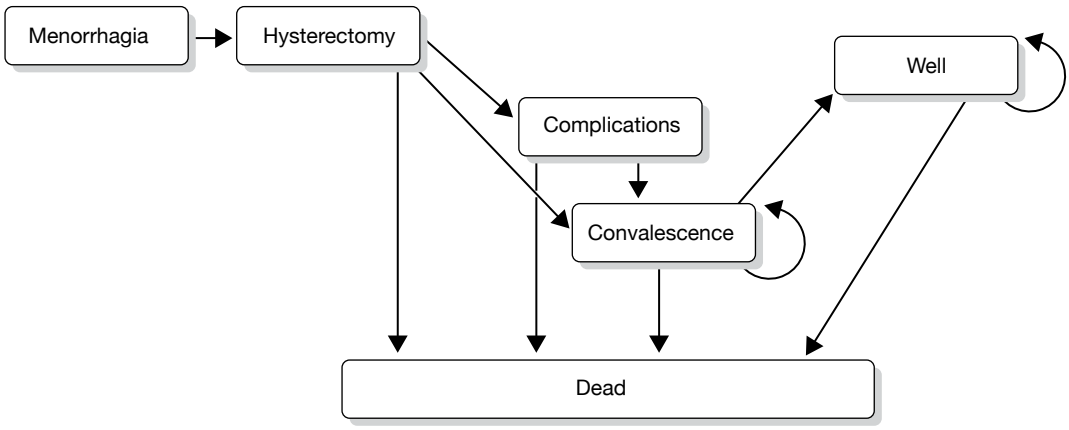


FIGURE 18 Clinical pathway for hysterectomy.

- Death – it is possible to die from natural causes. At hysterectomy and EA, women may also die as a direct result of the surgical procedure.

Definition of health states for hysterectomy pathways

- Menorrhagia – all women in the cohort have pre-operative HMB.
- Hysterectomy – all women undergo hysterectomy.
- Complications – following hysterectomy, some women will experience severe postoperative complications. The effects of these may last for 1 month. Operative complications are included in the hysterectomy state.
- Convalescence – following hysterectomy both with and without complications, a period of convalescence is experienced. This may last up to 3 months.
- Well – following convalescence women are satisfied with treatment.

Definition of health states in Mirena pathways

- Menorrhagia – all women in the cohort have pre-operative HMB.
- Mirena – all women have Mirena inserted.
- Well – following Mirena, women are satisfied with treatment.

- Symptomatic – following Mirena or well, HMB may recur (treatment failure) at any time. Women may be retreated, have a second-generation EA or remain symptomatic.

Clinical assumptions and adverse events

Mirena

The lifespan of Mirena is assumed to be 5 years.[15] If it is successful, treatment is repeated at 5 years. Treatment failure in Mirena is assumed to be more evident in the first year.[27,93] We also assumed a 2% insertion failure rate where procedure is repeated within a month.[27] If Mirena is unsuccessful, women move to the second-generation pathway. No adverse events associated with Mirena were available from the literature.

First- and second-generation endometrial ablation techniques and hysterectomy

Large national audits of hysterectomy and first-generation EA techniques were used as sources for perioperative and severe postoperative adverse events[5,47] and are presented in *Table 13*. Minor postoperative complications were not modelled. For second-generation EA techniques, complication rates of MEA were used.[134]

TABLE 13 Data used in the model

Description	*p*-value	Source		
Background mortality rate	0.00015	Life tables[133]		
Proportion of symptomatic women (post initial ablation) who have a repeat ablation	0.4	Cooper, 2001[28]		
Proportion of symptomatic women (post initial ablation) who have a hysterectomy	0.6	Cooper, 2001[28]		
	First-generation techniques[a]		***Second-generation techniques***	
	***p*-value**	**Source**	***p*-value**	**Source**
Operative complications	0.0445	Overton, 1997[47]	0.0028	Parkin, 2000[134]
Severe postoperative complications	0.0292	Overton, 1997[47]	0.0007	Parkin, 2000[134]
Death after operation	0.0002	Overton, 1997[47]	0	Parkin, 2000[134]
Severe complications following well	Fit by calibration to IPD meta-analysis		Fit by calibration to IPD meta-analysis	
Symptomatic following well	Fit by calibration to IPD meta-analysis		Fit by calibration to IPD meta-analysis	
Symptomatic following operative	Fit by calibration to IPD meta-analysis		Fit by calibration to IPD meta-analysis	
	LNG IUS		***Hysterectomy***[b]	
	***p*-value**	**Source**	***p*-value**	**Source**
Operative complications	–	–	0.0358	Maresh, 2002[5]
Severe postoperative complications	–	–	0.0102	Maresh, 2002[5]
Death after operation	–	–	0.0003	Maresh, 2002[5]
Proportion of women with LNG IUS in situ – year 1	0.6806[c]	Hurskainen, 2001[93]	–	–
Proportion of women with LNG IUS in situ – years 2–5	0.7037[c]	Hurskainen, 2004[27]	–	–
Insertion failure rate	0.0168	Hurskainen, 2004[27]	–	–

a Complication and mortality rates after repeat ablation (RB) were double those after initial ablation (MacLean-Fraser *et al.*[135] and professional estimate).
b Complication and mortality rates were adjusted for DUB population only.
c Original values as reported in the papers are presented. Failure per month in year 1 and years 2–5 were calculated.

Some of the data (specifically proportions in various health states at points in time after initial treatment) related to model outputs rather than model inputs. We applied a method that may be called 'probabilistic calibration' whereby model inputs for the relevant parameters were sampled uniformly across the plausible range, and cost and QALY outcomes were weighted according to the likelihood function comparing the model proportions in the various health states with the data.

Repeat ablation (rollerball)

If EA of any type fails, repeat ablation or hysterectomy is offered. In the model it is assumed that 60% of those with recurrent menorrhagia (symptomatic state) will have a repeat EA and 40% will have a hysterectomy.[28] Only one repeat ablation is offered, which is RB (first-generation EA). If the treatment fails a second time, only hysterectomy is available.

Complication rates in the repeat ablation are assumed to be double those incurred for the initial ablation.[135]

Repeat procedure

The transition probability for requiring a repeat procedure (ablation or hysterectomy) is time-dependent and is reduced by a constant factor each month. This reflects a decreasing hazard, which is obvious from the IPD data.

Resource use and costs

In order to calculate the costs of each of the procedures, a range of sources was used. All costs in the model are in UK £ (2008 value). Appropriate indices were used to inflate some of the costs that were obtained from the literature.[136] Costs are presented in *Table 14* and are discounted at 3.5% per annum.

Mirena

The cost of insertion of Mirena was estimated at £130.27. This procedure is assumed to be performed in a menstrual clinic as an outpatient procedure. The total cost includes those for the device, the initial consultation (10 minutes with a nurse and 30 minutes with a specialist registrar) and a sterile pack for use during Mirena insertion. Cost of discontinuation (£28.34) includes the cost of the consultation and the consumables (sterile pack) used for removal of the device.

First-generation techniques

The cost of first-generation EA was estimated at £1238. The source was a study which compared the costs of treating women with menorrhagia by hysterectomy or hysteroscopic surgery.[32] Costs in this study included pre-surgery treatment for EA, technical equipment (which varied for each method), hospital costs, gynaecological outpatient costs and retreatment. We excluded retreatment from our estimate because this is a separate procedure included in the model.

Second-generation techniques

The cost of second-generation techniques was estimated at £1101. The source for the costs of MEA and TBEA was the HTA report by Garside *et al.*[16] Statistical weights for the weighted cost mean were obtained from a study reporting NHS hospital episode statistics of EA from 1989–90 to 2005.[30]

TABLE 14 Table of costs

Description	Unit cost (£)[a]	Source
First-generation ablation techniques[b]	£1238	Cameron, 1996[32]
Second-generation ablation techniques[c]	£1101	Garside, 2004[16]
Repeat ablation (rollerball)	£1238	Cameron, 1996[32]
Hysterectomy	£2162	Cameron, 1996[32]
GP visit for referral to secondary care[d]	£46	PSSRU[137]
LNG IUS (Mirena)		
Total initial stage cost	£130.27	British National Formulary,[138] National Collaborating Centre for Women's and Children's Health CG44,[15] PSSRU[137]
Discontinuation	£28.34	
Adverse events[e]		
First-generation ablation techniques	£2161	National schedule of reference costs[139]
Second-generation ablation techniques	£1198	
Hysterectomy	£3008	

a Cost year is 2008.
b Weighted mean for TCRE and RB. Statistical weights assumed to be equal.
c Weighted mean for MEA and TBEA. Statistical weights calculated from Reid.[30]
d Cost of GP consultations was included as part of the referral to secondary care for a second-generation EA after Mirena failed and for a repeat ablation or a hysterectomy after both ablations failed.
e Values used for both 'immediate operative adverse events' and 'severe postoperative adverse events'.

Hysterectomy

The cost of hysterectomy was estimated at £2162 and the source was a study comparing the costs of treating women with menorrhagia by hysterectomy or hysteroscopic surgery.[32] For women who had a hysterectomy after a failed repeat ablation, we included an additional cost of a consultation with a GP for referral from primary to secondary care at £46.[136]

Repeat ablation (rollerball)

The cost of repeat ablation was the same as the cost of the first-generation techniques described above. We also included an additional cost of a GP consultation for referral from primary to secondary care, at £46.[136]

Adverse events

The source for the costs of adverse events was NHS reference costs (2009).[139] We used the same cost for both perioperative and severe postoperative complications for each of the procedures included in the model. The varying severity of complications of the two different types of EA and hysterectomy is reflected in costs as well. The cost of complications was £2161 for the first-generation techniques, £1198 for the second-generation techniques and £3008 for hysterectomy.

The MISTLETOE study was the source for perioperative and severe postoperative adverse events of the first-generation EA techniques.[47]

For second-generation EA techniques, complication rates of MEA were used as a proxy.[134] We chose not to use the complications rates from MISTLETOE for the second-generation EA techniques as well for the following reasons:

- The MISTLETOE study included radiofrequency and cryoablation only, and did not include MEA or TBEA, which are the commonest second-generation procedures used at the present time.
- Radiofrequency is no longer performed as it proved to be unsafe.
- The number of cryoablations in MISTLETOE is very low ($n = 36$) and there were no perioperative complications within that group.
- Severe postoperative complications in the MISTLETOE study are more frequent in the second-generation group than in the first-generation group, which is counterintuitive, as complications in second-generation techniques should be less frequent.

In using the complication rates of MEA from Parkin[134] ($n = 1400$), we acknowledge that this study might underestimate the true incidence of complications. All data reported in this study were from a specialist centre and reported complication rates might not be representative of other second-generation techniques. There were no adverse events associated with Mirena.

Utility values

The outcomes of treatment in the analysis are expressed in terms of QALYs gained for each procedure. Published evidence sources were used to identify the QoL weightings associated with each state in the model.[27,132,140] These values are described in detail in *Table 15*.

Heavy menstrual bleeding is associated with a QoL value of 0.50 – that is, patients who suffer from this illness have reported that they feel a loss in terms of QoL value that is equivalent to half a year at full health.[132] Treatment with Mirena is associated with a QoL value of 0.84.[27] After a successful treatment with Mirena, women move to the 'well' state, which according to the literature is also associated with a QoL value of 0.84.[27] EA is associated with a QoL value of 0.76, which captures convalescence as women are assumed to fully recover within 1 month (model cycle). After a successful first-generation ablation, women move to the 'well' state, associated with a QoL value of 0.73.[132] After a successful second-generation ablation, women move to the 'well' state, associated with a QoL value of 0.84, which is assumed to be the same as the 'well' state after a successful Mirena. This assumption was made in the absence of any available evidence and suggests that women's QoL value is 'better off' after a second-generation ablation compared to the QoL value after a first-generation ablation. We assumed that the 'well post second-generation ablation' state is the same as the 'well post LNG IUS' state and *not* same as the 'well post first-generation ablation' state. This was based on the fact that second-generation techniques (generally speaking), if successful, perform better, are less invasive and have fewer adverse events.

Hysterectomy is associated with a QoL value of 0.56 and 'convalescence post hysterectomy' with a QoL value of 0.74.[132] For the 'hysterectomy' state, we are assessing the mean QoL value for the month in which the hysterectomy is performed. We assumed that the 'hysterectomy' state utility value is 25% lower than the value for the health state 'convalescence post hysterectomy'; that is, zero QoL value for the 25% of that month that is the period of hospitalisation and then convalescence for the rest of the month. This assumption is in line with Garside *et al.*,[16] except we have reduced the hospitalisation to 25% of the month instead of 33% of the month because this is the duration of hospitalisation in our source for the cost of hysterectomy.[32] After a successful hysterectomy, women move to the 'well' state, which is associated with a QoL value of 0.88.[27]

Perioperative and severe postoperative complications of EA and hysterectomy are associated with a QoL value of 0.49.[140] If treatment is unsuccessful, women move to the symptomatic state with a QoL value of 0.50; in the model it is assumed that this state is equivalent to the HMB state in

TABLE 15 Health state utilities

	Utility weight (95% CI)	Source	Comment
Menorrhagia	0.50 (SE 0.04)	Sculpher, 1998[132]	–
Dead	0	–	By definition
Hysterectomy	0.56	–	Assumed ¼ less than 'convalescence post hysterectomy'
Well post hysterectomy	0.88 (0.75 to 0.95)[a]	Hurskainen, 2004[27]	–
Convalescence post hysterectomy	0.74 (SE 0.05)	Sculpher, 1998[132]	–
Severe complications post hysterectomy	0.49	Clegg, 2007[138]	–
First-generation EA	0.76 (SE 0.04)	Sculpher, 1998[132]	Includes 'convalescence post first-generation EA'
Well post first-generation EA	0.73 (SE 0.04)	Sculpher, 1998[132]	–
Severe complications post first-generation EA	0.49	–	Same as 'complications post hysterectomy'
Symptomatic post first-generation EA	0.50 (SE 0.04)	Sculpher, 1998[132]	Same as 'menorrhagia'
Second-generation EA	0.76 (SE 0.04)	–	Assumed same as 'first-generation EA'
Well post second-generation EA	0.84 (0.73 to 0.93)[a]	–	Same as 'Mirena'
Severe complications post second-generation EA	0.49	–	Same as 'complications post hysterectomy'
Symptomatic post second-generation EA	0.50 (SE 0.04)	Sculpher, 1998[132]	Same as 'menorrhagia'
Repeat ablation (RB)	0.76 (SE 0.04)	–	Includes 'convalescence post first-generation EA'
Well post repeat ablation	0.73 (SE 0.04)	Sculpher, 1998[132]	–
Severe complications post repeat ablation	0.49	–	Same as 'complications post hysterectomy'
Symptomatic post repeat ablation	0.50 (SE 0.04)	Sculpher, 1998[132]	Same as 'menorrhagia'
Mirena	0.84 (0.73 to 0.93)[a]	Hurskainen, 2004[27]	Same as 'well post Mirena'
Well post Mirena	0.84 (0.73 to 0.93)[a]	Hurskainen, 2004[27]	–
Symptomatic post Mirena	0.50 (SE 0.04)	Sculpher, 1998[132]	Same as 'menorrhagia'

SE, standard error
a Mean and width of CI were fitted.

terms of QoL value loss. The QoL value weightings associated with the repeat ablation (RB) are the same as the ones used for the first-generation techniques described above.

For costs, we aggregated according to the distributions at the start of each cycle. This is necessary to account for the full costs of initial treatment. When assessing total QALYs, we aggregated using Simpson's rule, which is an improvement on the half-cycle correction most commonly used (see, for example, Fröberg, 1969).[141]

Sensitivity analysis and reporting of results

All analyses are carried out from the perspective of the UK National Health Service and are reported in terms of incremental cost-effectiveness ratios (ICERs). Dominance in the results will exist if one strategy is found to be both cheaper and more effective (in terms of producing more QALYs). Two main analyses, Analysis 1 and Analysis 2, are carried out, and probabilistic sensitivity analysis (PSA) is carried out for both analyses. For reasons that will be explained in the discussion, Analysis 2 is assumed to be the base case. Additional alternative analyses carried out in addition to the PSA, including deterministic sensitivity analyses and subgroup analyses, are carried out on Analysis 2 only.

Uncertainty in the model parameters was assessed through PSA. In its most common form, PSA assigns to each input parameter a specific distribution and, by drawing randomly from those distributions, generates a large number of mean cost and effectiveness estimates. These estimates are then used to form an empirical joint distribution of the differences in cost and effectiveness between interventions.[142,143]

Value of information analysis

Where a decision is not robust to plausible variation in the input parameters, it is possible to estimate a statistic known as the expected value of perfect information (EVPI). This is determined as a function of the threshold ICER, which allows a conversion from QALYs to monetary value. The preferred decision under uncertainty is determined by maximising the mean net benefit across the distribution of input parameter values. For any specific parameter set which leads to the same decision, there is no value of information attached to those parameters. If, however, a parameter set leads to a change in the decision, then the value attached to that parameter set is the difference in net monetary benefit between the decision made under uncertainty and the decision made knowing those parameter values. The EVPI is obtained by calculating the value attached to each parameter set used in the PSA and averaging across all parameter sets, taking into account the weightings determined by the probabilistic calibration described in the previous section.

Subgroup analysis and deterministic sensitivity analysis

All subgroup and sensitivity analyses were carried out for Analysis 2 only, which was assumed to be the base case. Two subgroup analyses were considered appropriate a priori:

1. *A subgroup analysis on the basis of age*: This was planned specifically to analyse the data for women below the age of 40 years in our population in order to ascertain whether the results of the analysis, in terms of costs and satisfaction, were different for this subgroup compared with the women overall.

2. *A subgroup analysis on the basis of uterine cavity length*: The rationale for this subgroup analysis is that women with shorter uterine cavity length are more amenable to successful ablation (as opposed to hysterectomy). There was some evidence in the raw collated data which suggested the presence of a trend in support of this hypothesis and so a subgroup analysis was deemed appropriate to investigate this.

Three alternative one-way deterministic sensitivity analyses were considered appropriate a priori:

1. *Utilities are changed from the base case in which the mean utility values are used to an analysis based on reported median values*: We carried out the analyses using the median value for the utility scores from Sculpher.[132] In the base case we have used the mean scores for the utility values reported by Sculpher.[132] Use of the mean scores is most appropriate in economic evaluation. However, we noted that other published studies which used the same source for the utility values used the reported median values from the same study[132] and not the reported means, but without explanation or justification. We explored this to see if using the medians made any difference to the base-case results.
2. *Assumption that all women undergo hysterectomy after a failed repeat ablation*: In the base case it is assumed that all (100%) women who have a failed repeat ablation finally resort to a hysterectomy. In the sensitivity analysis this assumption is changed to assume that 10% of women remain symptomatic for the rest of the time period but do not seek further treatment. Thus, in this sensitivity analysis it was assumed that 90% of women chose hysterectomy after a second failed ablation and the remaining 10% chose to remain symptomatic.
3. *Costs associated with anaesthetic*: We considered varying the cost associated with anaesthetic. Seymour *et al.*[144] have presented the cost of MEA under local versus general anaesthesia. The cost was £440 for general and £428 for local anaesthesia. It had been intended to use the £440 GA estimate in the base case and the £428 LA estimate in the sensitivity analysis to explore any impact. However, on closer examination it became apparent that Seymour *et al.* excluded the cost of hormonal endometrial preparation (administered 5 weeks prior to MEA) and any pre-admission consultations because they were comparing different types of MEA, and these costs occurred in both types. Thus, we realise that the results from Seymour *et al.*[144] are not comparable to other reports of EA and hysterectomy used in the current model. Furthermore, by using the cost data from Garside *et al.*,[16] we have incorporated the balance between procedures undertaken under local versus general anaesthesia in the base-case analysis.

Results

Analysis 1

For Analysis 1, for the first-generation EA, second-generation EA and Mirena strategies, we assumed that repeat procedures (ablation or hysterectomy) are allowed at any age, but with a decreasing hazard.

The deterministic results are presented in *Table 16*.

TABLE 16 Results of deterministic analysis for Analysis 1

	Total costs (£)	Total QALYs	ICER (vs hysterectomy)
First-generation EA	30,040,000	64,485	Dominated
Second-generation EA	25,950,000	68,965	Dominated
Mirena	15,630,000	68,758	1600
Hysterectomy	23,000,000	73,332	–

The strategy of hysterectomy

The deterministic results show that hysterectomy dominates both first- and second-generation EA because the hysterectomy strategy is both less costly and produces more QALYs than either of the other two. The strategy of hysterectomy is more costly than the strategy of Mirena but also produces more QALYs. The incremental cost per additional QALY of hysterectomy compared with Mirena is £1600 per QALY.

The second-generation endometrial ablation strategy

Second-generation EA is both cheaper and produces more QALYs than the first-generation EA strategy and so is said to dominate it.

The Mirena strategy

The results show that Mirena is both less costly and produces more QALYs than first-generation EA and thus dominates it. However, although the Mirena strategy costs only half as much as the second-generation EA strategy, the second-generation EA is more effective at producing QALYs. Thus, the ICER for second-generation EA compared with Mirena is £50,000 per QALY. This means that for every woman who is treated with second-generation EA instead of Mirena there is an additional cost of £50,000 to achieve an additional QALY.

Consider for example a threshold ICER of £5000 per QALY; according to our model (not shown), the expected net monetary benefit per woman for hysterectomy is £34,386 at this valuation. This is higher than the expected net benefit for any of the other three options and so hysterectomy is the preferred option given parameter uncertainty. However, replications of the model accounting for approximately 26% of the probability favour different options. So the model probability that hysterectomy is the preferred option is 74%. This is the probability shown on the cost-effectiveness acceptability frontier (CEAF) (*Figure 19*). The limitation of the CEAF as a decision aid is that it does not tell us, in the remaining 26% of cases, whether changing to another option would make a large or a small difference to the expected net benefit.

Figure 20 shows the difference between the expected net benefit of the optimal strategy allowing for perfect information and the expected net benefit of the optimal strategy given current information for any given threshold ICER. If the preferred option could be chosen after completely resolving parameter uncertainty, then the estimated net benefit per woman would be

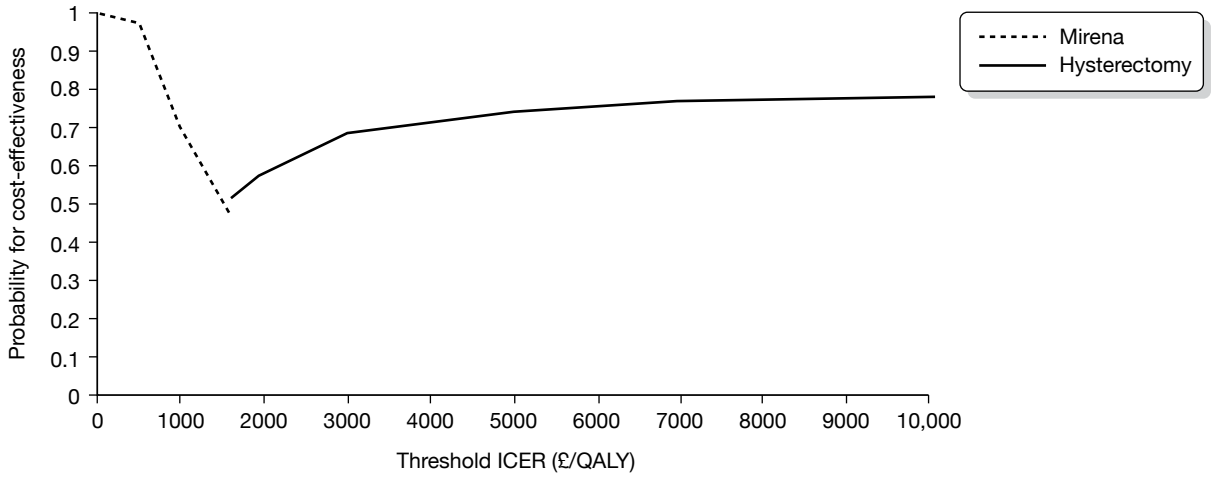


FIGURE 19 Cost-effectiveness acceptability frontier for Analysis 1.

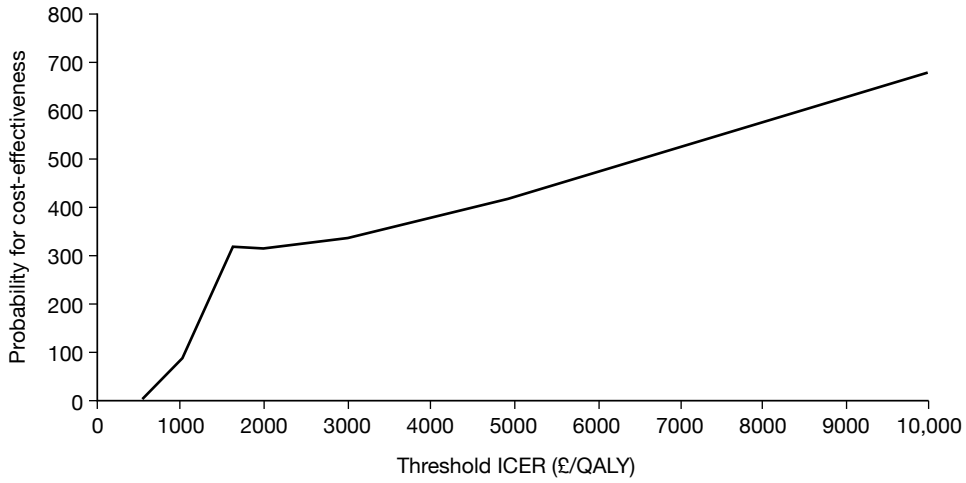


FIGURE 20 Probability that the preferred option is cost-effective for any given threshold ICER.

£34,812 (from the model, not shown). The additional £426 is known as the 'per woman' EVPI. Similar remarks apply at other threshold ICERs. At thresholds below 1600 per QALY (i.e. the kink in *Figure 19*), the preferred option is Mirena, and so the CEAF shows the probability that Mirena is the preferred option. As the threshold approaches zero the decision is effectively to choose the least costly option – this is Mirena in all replications of the model. So the probability shown on the CEAF goes to one, and the EVPI becomes zero. As the threshold ICER becomes very large, the decision is effectively the option with the highest QALY output. There is some residual uncertainty in the model and so the probability on the CEAF does not actually approach one. Had the EVPI been measured in health units, the height of the EVPI curve would tend to a fixed non-zero limit. However, because it is measured in monetary terms, the curve has an increasing slope.

In summary, for every replication in this model, Mirena was the least costly option. For very low-threshold ICERs, Mirena is preferred with certainty (so EVPI tends to zero and CEAF tends to one). In some replications of the model (accounting for 18% of the probability), hysterectomy was not the most effective option (in terms of total number of QALYs). Therefore, the CEAF never goes above 82% and the EVPI does not go to zero.

At an ICER of £1600, the preferred option changes from coil to hysterectomy, so there is a discontinuity in the CEAF and a discontinuity in the gradient of the EVPI curve.

Analysis 2

For Analysis 2, we assumed that if symptoms do not recur within 2 years of the initial ablation, then they are unlikely to do so later, and therefore no repeat procedure takes place thereafter. Thus, we have to limit the time as to when a repeat procedure (ablation or hysterectomy) may occur to 2 years. Similarly, as regards to repeat ablations, if women are not symptomatic within 4.5 years of the initial ablation (2 + 2 years + 6 months from decision to repeat procedure), it is assumed they will never become symptomatic (*Table 17*).

First-generation EA is dominated by all the other strategies.

TABLE 17 Deterministic results for Analysis 2

	Total costs (£)	Total QALYs	ICER (vs hysterectomy)
First-generation EA	23,590,000	63,745	Dominated
Second-generation EA	19,470,000	69,678	970
Mirena	16,150,000	68,566	1440
Hysterectomy	23,000,000	73,332	–

Second-generation endometrial ablation strategy

Second-generation EA is cheaper than all the other strategies except for Mirena; however, it does produce more QALYs than the Mirena strategy. The ICER for second-generation EA versus Mirena is £2980 per additional QALY.

The Mirena strategy

Mirena dominates the first-generation EA strategy but does not dominate second-generation EA.

The hysterectomy strategy

Hysterectomy is the strategy which produces the most QALYS. As this strategy is both cheaper and produces more QALYs than the first-generation EA strategy, it is considered to dominate the latter. Hysterectomy is more expensive but produces more QALYs than the Mirena strategy and the ICER representing the value of the additional benefit of this strategy compared with Mirena is £1440 per additional QALY. Despite being more costly, the hysterectomy strategy produces more QALYs than the second-generation EA strategy and the cost per unit of benefit for this comparison is £970 per additional QALY.

The detailed explanation of the interpretation of *Figures 21 and 22* is similar to the explanation given for the comparable figures for Analysis 1.

In summary, *Figure 21* presents the probability that the preferred option is cost-effective for any given threshold ICER.

In this model, hysterectomy dominates the graph. For all but a few replications, and the few account for a negligible probability, Mirena was the least costly option and for very low-threshold ICERs Mirena is preferred with certainty (so EVPI tends to zero and CEAF tends to one); however, as the threshold ICER increases hysterectomy becomes the preferred option.

However, it is clear that in some replications of the model (accounting for 20% of the probability), hysterectomy was not the most effective option (in terms of total number of QALYs). Therefore, the CEAF never goes above 80% and the EVPI does not go to zero, which would suggest that there is value in finding out whether if the uncertainty was removed hysterectomy would remain the most cost-effective option.

At an ICER of £1440, the preferred option changes from the Mirena coil to hysterectomy, so there is a discontinuity both in the CEAF as well as in the gradient of the EVPI curve presented in *Figure 21*.

Figure 22 presents the difference between the expected net benefit of the optimal strategy allowing for perfect information and the expected net benefit of the optimal strategy given current information for any given threshold ICER.

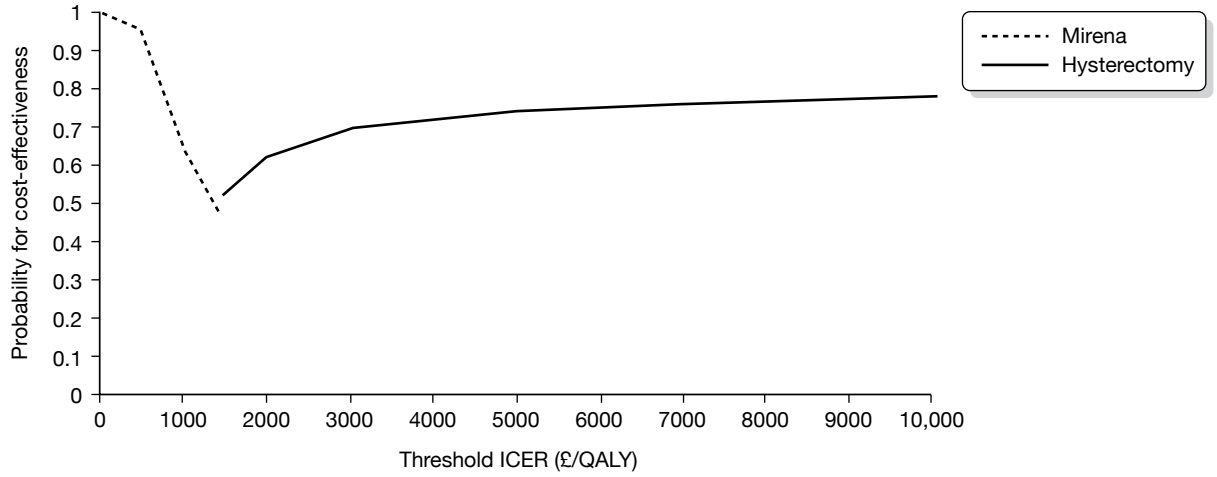


FIGURE 21 Cost-effectiveness acceptability frontier for Analysis 2.

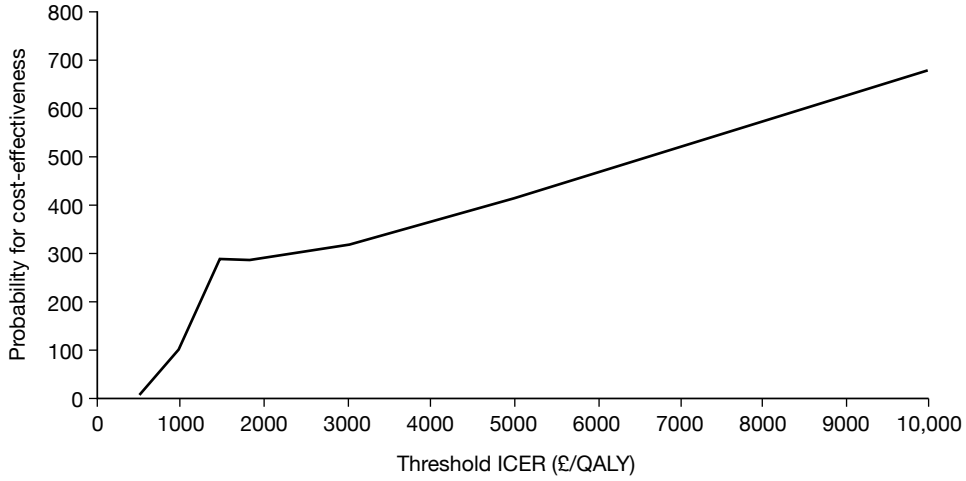


FIGURE 22 Per-woman expected value of perfect information.

Results of the subgroup analysis and one-way deterministic sensitivity analysis for Analysis 2

Subgroup analysis

On the basis of uterine cavity length

We carried out deterministic analysis assuming (1) that all women had the shortest uterine cavity length and (2) that all women had the longest uterine cavity length. A corresponding PSA was also carried out.

The results of the deterministic analysis are presented in *Table 18*.

The results are broadly similar to those presented in the base case analysis. First-generation ablation is no longer 'dominated' by hysterectomy and the incremental cost per additional QALY of hysterectomy compared with other strategies is now slightly higher than in the base case. However, there is unlikely to be a change in decision based on these results and hysterectomy

TABLE 18 Summary of deterministic sensitivity analysis for uterine cavity length (Analysis 2)

	Total costs (£)	Total QALYs	ICER (vs hysterectomy)
Short uterine cavity length			
First-generation EA	21,356,000	63,143	161
Second-generation EA	19,264,000	69,582	996
Mirena	15,667,000	68,201	1429
Hysterectomy	23,000,000	73,332	–
Long uterine cavity length			
First-generation EA	20,104,000	62,809	275
Second-generation EA	17,986,000	69,655	1364
Mirena	15,158,000	68,558	1642
Hysterectomy	23,000,000	73,332	–

continues to be likely the most cost-effective strategy. The results of the PSA serve to reinforce the results of the deterministic analysis presented above, and so are not presented.

Sensitivity analysis

Utilities are changed from the base case which used mean utility values to median utility values

In the base case analysis the mean value for the health state of 'well' post ablation was lower (0.73) than the mean value for the health state of first-generation ablation (0.76). Furthermore, the health state of 'well' post Mirena coil is not available in the report by Sculpher[132] and so a value was taken from a different study by Hurskainen *et al.*[27] As a result the health state for 'well' post Mirena coil is lower than the health state of 'well' post repeat ablation. Thus, although mean values are most appropriate, the use of mean values for utilities that are available in alternative published sources in our base case presents some slightly counterintuitive results.

Other studies have tackled this issue by using median values from different studies rather than the mean, which leads to a more intuitive set of values being used. But there is no other sensible justification for using median rather than mean values.

We carried out two sensitivity analyses to explore the impact of using the median reported values as opposed to the means.

A. QALY1

In this analysis we used the 'median' QoL values from Sculpher[132] and the value of 'well' post second EA and set this equal to 'well' post first-generation EA (i.e. 'well' post second-generation EA = 'well' post first-generation EA). It should be noted that 'well' in this instance is not the same as 'well' post Mirena coil that we used in the base case

The deterministic results of this sensitivity analysis are presented in *Table 19*.

The key change in results compared with the base case (Analysis 2) is that there is a clear shift away from the hysterectomy strategy in favour of the second-generation EA strategy as this

TABLE 19 Deterministic sensitivity analysis where mean values are substituted by median values (QALY1)

	Total costs (£)	Total QALYs	ICER (vs hysterectomy)
First-generation EA	23,588,000	74,218	2225
Second-generation EA	19,466,000	74,402	Dominates
Mirena	16,151,000	71,089	2391
Hysterectomy	23,000,000	73,954	–

now dominates both first-generation EA and hysterectomy. In the base case (Analysis 2), first-generation EA was dominated by second-generation EA, but the latter strategy did not dominate the hysterectomy strategy as it now does. The ICER for second-generation EA versus the Mirena strategy is £1000 per additional QALY in this sensitivity analysis.

B. QALY2

In this analysis we used the 'median' QoL values from Sculpher[132] and we set the value of well post second-generation EA = well post first-generation EA = well post Mirena coil.

The deterministic results from this sensitivity analysis are presented in *Table 20*.

These results are very similar to those presented in sensitivity analysis A. Once again the strategy of second-generation EA dominates first-generation EA and hysterectomy in contrast to the base case. The ICER for second-generation EA versus Mirena is £3624 per additional QALY.

Assumption that all (100%) women undergo hysterectomy after a failed repeat ablation

In the sensitivity analysis we change this assumption and assume that 10% of women remain symptomatic and only 90% follow the hysterectomy strategy after a failed repeat ablation

The deterministic results for this sensitivity analysis are presented in *Table 21*.

In this analysis, although the total cost and QALYs for each strategy are different to those in the base case, the overall result is unchanged from the base case.

Two sensitivity analyses deemed worthy of investigation before the main analysis was undertaken were not carried out for the following reasons:

1. *Subgroup analysis on women below 40 years of age.* On scrutinising the data it was clear that women below the age of 40 years had a 100% satisfaction rate with the Mirena coil strategy and therefore it was very clear that any subgroup analysis on the basis of age would be dominated by the strategy of coil. The base case uses 'all ages of women' – and so to specifically analyse the subgroup of women older than 40 years would prove fruitless because that result was captured by the main analysis.
2. *Sensitivity analysis – varying the cost of anaesthetic.* In the base case analyses, the cost of second-generation techniques (MEA and TBEA) was based on the estimates used by Garside *et al.*[16] The cost estimate from Garside *et al.*[16] already takes into account local versus general anaesthesia. It was thus considered that any recalculations/re-estimations to explore the minor difference in costs of LA versus GA observed in Seymour *et al.*[142] would add more uncertainty.

TABLE 20 Deterministic sensitivity analysis where mean values are substituted by median values (QALY2)

	Total costs (£)	Total QALYs	ICER (vs hysterectomy)
First-generation EA	23,588,000	74,218	2225
Second-generation EA	19,466,000	74,402	Dominates
Mirena	16,151,000	73,488	14,683
Hysterectomy	23,000,000	73,954	–

TABLE 21 Deterministic sensitivity analysis in which 90% of women are assumed to follow the hysterectomy strategy

	Total costs (£)	Total QALYs	ICER (vs hysterectomy)
First-generation EA	23,424,000	63,589	Dominated
Second-generation EA	19,323,000	69,542	970
Mirena	16,059,000	68,508	1440
Hysterectomy	23,000,000	73,332	–

Discussion

Statement of principal findings

The results of base case Analysis 2 show that the strategy of hysterectomy is most cost-effective. Hysterectomy dominates first-generation EA and, although more expensive, produces more QALYs than the other strategies. The ICER for hysterectomy compared with Mirena is £1440 per additional QALY. Compared with the second-generation EA strategy, the ICER for hysterectomy is £970 per additional QALY. These results suggest that hysterectomy would be considered the most cost-effective strategy in the light of the acceptable thresholds used by NICE, which tends to accept new technologies if the ICER is within £20,000 per additional QALY.

The results of this study were robust to all the sensitivity and subgroup analysis that were carried out with the exception of the sensitivity analysis carried out on the QALY data. The results of the main analysis reported in this study are based on an analysis that used the reported values of QALYs that are available in the published literature, specifically the 'mean' reported QALY values. When we carried out the same analysis using the median QALY values, the results changed and second-generation EA became the most cost-effective strategy, dominating both the first-generation ablation strategy and the hysterectomy strategy. In the sensitivity analysis the second-generation EA strategy was more expensive than the Mirena strategy but also produced more QALYs. The ICERs that resulted suggested that second-generation EA would be considered the most cost-effective strategy.

Strengths and limitations

The major strength of the economic component of this study is that it is based on a state-of-the-art Markov model which was informed by data from an IPD meta-analysis of randomised trials. A multidisciplinary team including economists, expert clinicians and statisticians provided input into the model structure primarily based on the evidence in the literature. All assumptions used in the model were based on the available evidence as far as possible. Assumptions were agreed with the team before the analysis was carried out and without knowledge of how these assumptions would affect the results.

In terms of limitations, not all aspects of outcome have been included because of the limited time scale in our model and the lack of long-term data. For strategies such as hysterectomy, for

instance, there is a finality associated with the procedure which may have an effect on women and, in the long term, have implications on QoL that have not been captured in our model. These include the possibility of long-term complications such as urinary incontinence for which surgery is required. Furthermore, the utilities used reflect the satisfaction of the outcomes only. It is clear that once women have had a hysterectomy their satisfaction is high since in contrast to the other interventions they experience no bleeding at all. But the utility measure does not capture the anxiety prior to hysterectomy associated with major surgery and GA. It is conceivable that such anxiety may lead to decisions that avoid the strategy and to try other options for as long as possible.

In addition, the fact that the complexity of the model contributed to a long running time has some effect on the extent of sensitivity analyses that were undertaken. Both the main analysis and sensitivity/subgroup analysis had a long model running time and required laborious recalibration, which meant that additional sensitivity and subgroup analyses were not undertaken without serious consideration.

Strengths and weaknesses and assumptions in relation to other studies

With regard to the utility values that have been used in the current study the following points are worth noting.

First, the paper by Sculpher[132] has been extensively referenced in the literature. Other studies (including Garside *et al.*[16]) have used the median values from that paper and not the means. However, for economic evaluation it is the mean values that are most appropriate to use.[145,146]

There are minor inconsistencies in the results presented by Sculpher[132] when looking at the QALY values for 'well post TCRE' and 'convalescence post TCRE' compared with the 'well post hysterectomy' and 'convalescence post hysterectomy' states: the mean QALY value for the 'well post TCRE' state is 0.73 while the mean QALY value for the 'convalescence post TCRE' state is 0.76. One would expect that 'well post TCRE' state would have a higher value than the 'convalescence post TCRE' state, which is the case for the median QALY values. This is also true for the mean and median QALY values of the 'well post hysterectomy' and 'convalescence post hysterectomy' states. Despite these apparent inconsistencies, we used the mean values as they were presented in the literature and considered it possible that this might be explained by the greater initial relief that a woman might experience after a hysterectomy than after a TCRE, although no explanation was put forward in the original literature.

Comparison with other studies

Our results are consistent with those reported by Garside *et al.*,[16] who compared MEA and TBEA with TCRE, RB and hysterectomy. A state-transition (Markov) model was used and assumed a hypothetical cohort of 1000 patients for a period of 10 years. Our model used the same assumptions as Garside *et al.*[16] regarding the age of women entering the model, who were assumed to be 42 years; the model ran for 10 years and women became menopausal at age 52 years.

Garside *et al.* concluded that hysterectomy is cost-effective compared with MEA and TBEA. They found that TBEA dominated all other ablation techniques.

In addition, when compared with hysterectomy, MEA and TBEA were found to be less costly but provided fewer QALYs. Garside *et al.*'s ICER for hysterectomy compared with TBEA was £2410 per QALY, and for hysterectomy compared with MEA their ICER was £2108 per QALY. The authors found that, when comparing MEA and TBEA with TCRE, RB and hysterectomy,

the model was highly sensitive to the utility values associated with being well following ablation. Garside *et al.*[16] recommended that results are interpreted with caution owing to the sensitivity of the model to the utility values used.

The results of our study do not concur with the result of the trial by Hurskainen *et al.*[27] from Finland, which compared Mirena with hysterectomy. Mirena was found to be cost-effective at 5 years when compared with hysterectomy. There was no statistically significant difference in QoL scores at 5 years, as measured by the EQ-5D instrument, between the two treatment groups. Mean direct costs in the Mirena arm remained significantly lower ($1892) than in the hysterectomy arm ($2787) despite 40% of women in the Mirena arm going on to have a hysterectomy.

Economic modelling undertaken to inform NICE guidelines on HMB (NICE CG44, 2007)[15] shows that Mirena is cost-effective when compared with both hormonal and non-hormonal treatment. It generates more QALYs at a lower cost than any other medical or surgical treatment strategy considered. This analysis also considered surgery as a comparator treatment. The surgical strategy produced fewer QALYs at a higher cost than Mirena. The NICE model assumed that, within the 5-year lifespan of Mirena, some women who experienced failure would move straight to hysterectomy (based on data from Hurskainen *et al.*[27]). In contrast, in our study the assumption (based on advice from clinical colleagues) was that all women who experienced failure with the Mirena strategy would follow the second-generation EA pathway in the first instance.

However, while hysterectomy clearly comes out on top, the available long-term follow-up data on Mirena use are so inconsistent that we have to be cautious in our interpretation.

Meaning of the study

The results of this study suggest that hysterectomy is more cost-effective than either ablation or Mirena. These results are based on 'mean' reported values of utilities in published studies and are highly sensitive to the utility values that are used. When 'median' values for utility estimates are used, the strategy of second-generation ablation becomes the most cost-effective strategy. To the current authors there is no clear justification for using median values but they have been used in similar prominent studies without such justification. We assume that this is because some of the 'mean' values reported in the literature appear inconsistent. Replacing 'mean' values that appear inconsistent with 'median' values without clear justification will bias any results. The clear sensitivity of the results to the utility values serves to highlight the importance of using appropriate and robust data for utilities.

The study shows that first-generation ablation techniques are less cost-effective than second-generation techniques whatever utility values are used. Based on available data, Mirena does not come through as a cost-effective strategy compared with second-generation ablation or hysterectomy.

Unanswered questions for future research

The current study has used a state-of-the-art model, data from an IPD meta-analysis and all available data on QoL associated with available interventions and the outcomes for alternative treatments for HMB. One of the main causes of uncertainty regards the utility values associated with alternative interventions and their success. There would be little value in future studies comparing the outcomes and costs of any alternative interventions to treat women for menorrhagia without undertaking a comprehensive study to investigate the QoL associated with

the outcomes of the alternative interventions. Future studies should also explore the preferences and a priori anxieties associated with hysterectomy and the alternatives. Our study suggests that current available data are not robust enough.

Finally, both the preferences of women and clinicians perhaps need to be considered, as do the economic consequences of hysterectomy in terms of long-term morbidity such as pelvic floor dysfunction. Many clinicians believe (rightly or wrongly) that there is no one-size-fits-all approach to HMB and individual choices can determine perceived success; thus, evidence on preferences is also required.

Chapter 5

Interpretation of available evidence and consensus regarding treatment

Data from the IPD meta-analysis suggest that more women are dissatisfied following first-generation EA than hysterectomy. However, it is important to note that dissatisfaction rates are low after all treatments and hysterectomy is associated with an increased hospital stay and recovery period. In the absence of head-to-head trials, indirect estimates suggest hysterectomy is also preferable to second-generation EA in terms of patient satisfaction. In terms of cost-effectiveness, hysterectomy is considered the best strategy, but it carries a higher risk of complications and is perceived as a final option by gynaecological experts and consumers. Dissatisfaction rates were comparable between first- and second-generation techniques, although second-generation techniques were cheaper, quicker and associated with a faster recovery and fewer complications. There are few comparisons of Mirena versus more invasive procedures. The few data available suggest that Mirena is potentially cheaper and more effective than first-generation ablation techniques, with rates of satisfaction similar to those of second-generation techniques. Owing to small, imprecise trials with relatively high levels of non-compliance, the evidence to suggest that hysterectomy is preferable to Mirena is currently so limited that definitive conclusions cannot yet be made.

Observational data indicate that at 7 years a quarter of women face further gynaecological surgery after EA while an initial hysterectomy for HMB is more likely to lead to further surgery for stress urinary incontinence. The incidence of endometrial cancer following EA is reassuringly low at 0.02%. The type of hysterectomy has an influence on future risk of surgery, with vaginal hysterectomy associated with a higher chance of further surgery for urinary incontinence and pelvic floor prolapse than hysterectomy carried out through the abdominal route.

A summary of the results on effectiveness and cost-effectiveness was sent electronically to 15 national experts (minimal-access gynaecological surgeons) along with a short questionnaire (see *Appendix 9*) to encourage a rapid response. After two mailings, responses were received from 10 clinicians. Their responses are summarised in *Table 22*. Mirena was offered as first-line treatment and second-generation EA as second-line treatment by 9 out of 10 responders, while hysterectomy was considered the final port of call for women with HMB in the absence of demonstrable organic pathology. It is also clear from the responses that such a simplistic approach was not considered appropriate by some of the clinicians, who felt that often the choice of treatment depended on which intervention had been used before. As *Table 22* suggests, some of the clinicians were also keen to incorporate the patients' own preferences. One in particular (Clinician G) indicated that patients should choose any one of the three options in the context of first-line treatment for HMB.

The letter to the clinicians along with a summary of their views was sent electronically to three consumers. All three agreed with the order in which the three treatments were prioritised by the clinicians. Two of them made further comments highlighting potential problems associated with a rigid clinical algorithm and pointed out other factors such as age and fertility status which could have a bearing on the choice of treatments. Both argued for a degree of flexibility in order to accommodate the needs and preferences of individual women (*Table 23*).

TABLE 22 Clinicians' responses to queries regarding the treatment of women with HMB with failed oral medical treatment and no obvious clinical abnormalities

	First-line treatment	Second-line treatment	Third-line treatment
Clinician A	Mirena or second-generation EA	Hysterectomy/second-generation EA	Hysterectomy
Clinician B	Mirena	Second-generation EA	Hysterectomy
Clinician C	Mirena	Second-generation EA	Hysterectomy
Clinician D	Mirena	Second-generation EA	Hysterectomy
Clinician E	Mirena	Second-generation EA but will consider hysterectomy	Will depend on previous treatment
Clinician F	Mirena	Second-generation EA	Hysterectomy but may consider repeat ablation
Clinician G	Would offer patient choice of Mirena, ablation and hysterectomy	Hysterectomy	Hysterectomy – (relevant if patient chose Mirena as first option)
Clinician H	Mirena	Second-generation EA	Hysterectomy
Clinician I	Mirena	Second-generation EA (hysterectomy if the patient wishes)	Hysterectomy
Clinician J	Mirena	Second-generation EA	Hysterectomy

TABLE 23 Consumer responses to clinician comments

	Agree with clinicians' choice of treatment as Mirena, second-generation EA and hysterectomy	Comments
Consumer 1	Yes	'I agree with the findings'
Consumer 2	Yes	'I would definitely go for the least invasive initially but would be looking for some assurance as to effectiveness. However, if I were older, and fertility were not an issue, then I might want to go straight to second-generation ablation. For me, hysterectomy would always be a last resort because it is major surgery. I think I would be happy if the professional presented the choices in the terms outlined'
Consumer 3	Yes	'I do not think that any one of the above (treatments) could be suggested as an outright "winner" in consumer terms. Leaving aside the women with obvious health problems which might dictate the treatment for HMB, the ultimate choice should be made by the woman in full knowledge of the facts. The age of the woman could well influence their choice – younger women might favour Mirena because it is reversible. There are also issues for some women with hysterectomy meaning loss of fertility/presumed femininity and some women are not happy with surgery'

Conclusion

An IPD meta-analysis of randomised trials as well as the results of a cost-effectiveness analysis favour hysterectomy in women with HMB. Interpretation of these results needs to take into account a number of issues. The limited evidence on the effectiveness of Mirena, concerns about the long-term consequences of hysterectomy and individual preferences of women and gynaecologists are factors that influence the choice of treatment. While hysterectomy results in significantly fewer women being dissatisfied than those undergoing EA, it is worth noting that rates of satisfaction were very high for all treatment modalities. Although economic models used suggest that hysterectomy is the most cost-effective treatment option for HMB, any decision to promote this procedure must balance the morbidity associated with it against the ease of Mirena insertions in the community, and the ability to perform second-generation ablative procedures outwith the traditional theatre setting. The latter could potentially free up theatre time in secondary care which could be used for other procedures. A key reason for the higher success rates associated with hysterectomy is the definitive nature of the procedure. Failure rates

for Mirena remain to be formally established, but those associated with EA are well known. Around a quarter of all women who undergo EA will require subsequent gynaecological surgery, with just under a fifth requiring a hysterectomy. Endometrial cancer rates following EA are very low, although longer term follow-up will be necessary to confirm this. Mirena protects against endometrial hyperplasia and hence endometrial cancer rates should be low. It is clear that clinical experts and consumers considered ease of access to treatment, degree of invasiveness, long-term consequences and patient autonomy to be important determinants. Expert clinical opinion favours offering the least invasive treatment, that is, Mirena first, followed by ablation, with hysterectomy reserved for women in whom the first two options have failed. This approach is endorsed by lay consumers, although they are anxious that women have the opportunity to choose the option that is best for them.

Chapter 6

Conclusion

This study has produced data on the clinical effectiveness and cost-effectiveness of different modalities of treatment of HMB and highlighted the risk of further surgery following EA and hysterectomy. It has also exposed gaps in the literature – especially with regards to the clinical effectiveness of Mirena in comparison with EA and hysterectomy and long-term follow-up data in women using it for HMB.

Despite a longer hospital stay and time to resumption of normal activities, more women were satisfied after hysterectomy than after first-generation EA. In the absence of head-to-head trials, indirect estimates suggest that hysterectomy is also preferable to second-generation EA in terms of patient satisfaction. Dissatisfaction rates were comparable between first- and second-generation techniques, although second-generation techniques were cheaper, quicker and associated with faster recovery and fewer complications. There are few comparisons of Mirena versus more invasive procedures. The few data available suggest that Mirena is potentially cheaper and more effective than first-generation ablation techniques with rates of satisfaction that are similar to first- and second-generation techniques. Owing to a paucity of trials, evidence to suggest that hysterectomy is preferable to Mirena is currently so limited that definitive conclusions cannot yet be made.

A quarter of women undergoing EA as an initial treatment are likely to face further gynaecological surgery (mainly repeat ablation or hysterectomy) for persistent menstrual problems. However, hysterectomy is more likely to lead to future surgery for stress urinary incontinence. Thus, in comparison with hysterectomy, the lower morbidity associated with EA needs to be balanced against the chance of repeat surgery for the same symptoms, although the risk of long-term pelvic floor problems may be less.

The cost-effectiveness analysis identified the strategy of opting for hysterectomy as the most cost-effective one. Hysterectomy is both cheaper as well as more effective than first-generation EA. In comparison with second-generation EA and Mirena, hysterectomy costs more but produces more QALYs. The ICER for hysterectomy is £1440 per additional QALY compared with Mirena and £970 per additional QALY compared with second-generation EA. These results suggest that hysterectomy would be considered the most cost-effective strategy in light of the acceptable thresholds used by NICE, which tends to accept new technologies if the ICER is within £20,000 per additional QALY.

Our review provides evidence that hysterectomy reduces dissatisfaction compared with EA and this information should be used as part of a consultation with women making a choice about treatment options when initial drug treatment fails to control HMB. EA is satisfactory for a very high proportion of women, but, if complete cessation of bleeding is sought, then hysterectomy may be offered. A decision to opt for hysterectomy needs also to take into account the invasive nature of the procedure and its potential for short- and long-term morbidity in some women. Relatively few trials have evaluated the evidence of effectiveness of Mirena. These are small, imprecise and have relatively high levels of compliance. Thus, we concur with a recent NICE recommendation that women should be offered Mirena before more invasive procedures. We have highlighted the benefits and risks associated with the three broad strategies for the treatment of HMB, and, while supportive of the existing NICE guideline on this subject, our

results underline the need for a degree of flexibility in accommodating women's preferences. Hysterectomy may be the most cost-effective strategy, but, owing to its invasive nature and higher risk of complications, is considered a final option by gynaecological experts and consumers who are swayed by other considerations such as ease of access to treatment, degree of invasiveness, long-term consequences and patient autonomy.

Acknowledgements

Individual patient data meta-analysis

We would like to thank all authors of identified trials for sending us their trial data, and both the British and European Societies of Gynaecological Endoscopy for their support and help with the review.

Members of the Collaborative Group for Individual Patient Data Meta-analysis

The following authors provided us with IPD from their trials (name, organisation, country):

- J Abbott, University of New South Wales, Sydney, Australia
- J Barrington, Torbay Hospital, South Devon, UK
- S Bhattacharya, University of Aberdeen, Aberdeen Maternity Hospital, Aberdeen, UK
- MY Bongers, Maxima Medical Centre, Veldhoven, the Netherlands
- J-L Brun, Hôpital Universitarie Pellegrin, Bordeaux, France
- R Busfield, data supplied by M Sowter, Auckland Obstetric Centre, New Zealand
- TJ Clark, Birmingham Women's Hospital, Birmingham, UK
- J Cooper (2004 trial), data supplied by Microsulis Medical Ltd, Hampshire, UK
- KG Cooper, Aberdeen Royal Infirmary, Aberdeen, UK
- SL Corson (2001 trial), data supplied by Boston Scientific Corporation, Marlborough, MA, USA
- K Dickersin, Johns Hopkins Bloomberg School of Public Health, USA
- N Dwyer, Weston General Hospital, Weston Super Mare, UK
- M Gannon, Midland Regional Hospital, Mullingar, Ireland
- J Hawe, Countess of Chester Hospital, Chester, UK
- R Hurskainen, University of Helsinki, Finland
- WR Meyer, data supplied by Ethicon, a Johnson & Johnson Company, NJ, USA
- H O'Connor, Coombe Women's Hospital, Dublin, Ireland
- S Pinion, Aberdeen Royal Infirmary, Aberdeen, UK
- AM Sambrook, Aberdeen Royal Infirmary, Aberdeen, UK
- WH Tam, Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong, China
- IAA van Zon-Rabelink, Medical Spectrum Twente, Enschede, the Netherlands
- E Zupi, Tor Vergata University, Rome, Italy.

Chapter 3

We would like to acknowledge Dr Venkat Timmaraju's contribution towards data cleaning and analysis and Dr Amalraj's contribution towards reanalysis of the data set.

Interpretation of evidence and consensus panel

We would like to thank all the clinicians and consumers for participating in our consensus panel.

Contributions of authors

S Bhattacharya (corresponding author), Professor of Reproductive Medicine, University of Aberdeen, conceived the idea for the project, co-ordinated the project, contributed to the design of the different components, interpreted results, edited *Chapters 2* and *4* and wrote the first draft of the final report.

LJ Middleton, Medical Statistician, Birmingham Clinical Trials Unit, University of Birmingham, developed the protocol, reviewed papers for systematic review, performed the statistical analysis, wrote the initial draft of *Chapter 2* and edited the final report.

A Tsourapas, Research Fellow, University of Birmingham, who, under the supervision of PB and TR, constructed the model, identified and collected the required data on costs and effectiveness required for the model, and carried out the analysis for *Chapter 4*.

A Lee, Chair in Medical Statistics, University of Aberdeen, helped to develop the protocol, led data analysis for *Chapter 3*, contributed to its current draft and helped edit the final report.

R Champaneria, Research Associate, Birmingham Clinical Trials Unit, University of Birmingham, carried out literature searches and retrieved the identified papers, acted as the group secretariat for the collection and amalgamation of IPD, and commented on the final report.

JP Daniels, Research Fellow, Birmingham Clinical Trials Unit, University of Birmingham, developed the protocol, co-ordinated the IPD meta-analysis and commented on the final report.

T Roberts, Professor of Health Economics, University of Birmingham, developed the protocol, designed the model-based cost-effectiveness analysis, wrote the first draft of *Chapter 4*, and edited the final report.

NH Hilken, IT Co-ordinator, Birmingham Clinical Trials Unit, University of Birmingham, developed the master database for IPD meta-analysis.

P Barton, Lecturer, University of Birmingham, developed the protocol, designed the model-based cost-effectiveness analysis and commented on the final report.

R Gray, Professor and Unit Director, Birmingham Clinical Trials Unit, University of Birmingham, developed the protocol, oversaw IPD meta-analysis and commented on the final report

KS Khan, Professor of Obstetrics, Gynaecology and Clinical Epidemiology, Academic Department of Obstetrics & Gynaecology, University of Birmingham, conceived the idea for the project, developed the protocol, contributed to the IPD meta-analysis and commented on the final report.

P Chien, Consultant Gynaecologist, Ninewells Hospital, Dundee, contributed to the design of the protocol, interpreted results and commented on the final report.

P O'Donovan, Professor and Consultant Obstetrician and Gynaecologist, Department of Obstetrics & Gynaecology, Bradford Royal Infirmary, contributed to design of the protocol, interpreted results and commented on the final report.

KG Cooper, Consultant Gynaecologist, Department of Obstetrics & Gynaecology, Aberdeen Royal Infirmary, conceived the idea for the project, developed the protocol, interpreted results, wrote the initial draft of *Chapter 3* and edited the final report.

Publication

1. Middleton LJ, Champaneria R, Daniels JP, Bhattacharya S, Cooper KG, Hilken NH, *et al.* Hysterectomy, endometrial destruction, and levonorgestrel releasing intrauterine system (Mirena) for heavy menstrual bleeding: systematic review and meta-analysis of data from individual patients. *BMJ* 2010;**341**:c3929.

References

1. Rees MC. Role of menstrual blood loss measurements in management of complaints of excessive menstrual bleeding. *Br J Obstet Gynaecol* 1991;**98**:327–8.
2. National Statistics Online. *Census 2001*. 2002. URL: www.statistics.gov.uk/census2001/pop2001/england_wales.asp (accessed 1 February 2010).
3. Coulter A, Peto V, Jenkinson C. Quality of life and patient satisfaction following treatment for menorrhagia. *Fam Pract* 1994;**11**:394–401.
4. Vessey MP, Villard-Mackintosh L, McPherson K, Coulter A, Yeates D. The epidemiology of hysterectomy: findings in a large cohort study. *Br J Obstet Gynaecol* 1992;**99**:402–7.
5. Maresh MJ, Metcalfe MA, McPherson K, Overton C, Hall V, Hargreaves J, *et al.* The VALUE national hysterectomy study: description of the patients and their surgery. *BJOG* 2002;**109**:302–12.
6. Royal College of Obstetricians and Gynaecologists. *The initial management of menorrhagia. Evidence based guidelines No. 1*. London: Royal College of Obstetricians and Gynaecologists; 1998.
7. Chimbira TH, Anderson AB, Turnbull A. Relation between measured menstrual blood loss and patient's subjective assessment of loss, duration of bleeding, number of sanitary towels used, uterine weight and endometrial surface area. *Br J Obstet Gynaecol* 1980;**87**:603–9.
8. Fraser IS, McCarron G, Markham R. A preliminary study of factors influencing perception of menstrual blood loss volume. *Am J Obstet Gynecol* 1984;**149**:788–93.
9. Hallberg L, Hogdahl AM, Nilsson L, Rybo G. Menstrual blood loss – a population study. Variation at different ages and attempts to define normality. *Acta Obstet Gynecol Scand* 1966;**45**:320–51.
10. Hallberg L, Nilsson L. Determination of menstrual blood loss. *Scand J Clin Lab Invest* 1964;**16**:244–8.
11. Gannon MJ, Day P, Hammadieh N, Johnson N. A new method for measuring menstrual blood loss and its use in screening women before endometrial ablation. *Br J Obstet Gynaecol* 1996;**103**:1029–33.
12. Higham JM, O'Brien PM, Shaw RW. Assessment of menstrual blood loss using a pictorial chart. *Br J Obstet Gynaecol* 1990;**97**:734–9.
13. Reid PC, Coker A, Coltart R. Assessment of menstrual blood loss using a pictorial chart: a validation study. *BJOG* 2000;**107**:320–2.
14. Wyatt KM, Dimmock PW, Walker TJ, O'Brien PM. Determination of total menstrual blood loss. *Fertil Steril* 2001;**76**:125–31.
15. National Institute for Health and Clinical Excellence. *CG44 Heavy Menstrual Bleeding: full guideline; NICE Guideline January 2007*. URL: http://guidance.nice.org.uk/CG44/Guidance/pdf/English (accessed 1 February 2010).
16. Garside R, Stein K, Wyatt K, Round A, Price A. The effectiveness and cost-effectiveness of microwave and thermal balloon endometrial ablation for heavy menstrual bleeding: a systematic review and economic modelling. *Health Technol Assess* 2004;**8**(3).

17. Lethaby A, Hickey M. Endometrial destruction techniques for heavy menstrual bleeding. *Cochrane Database Syst Rev* 2002;**2**:CD001501.

18. Coulter A, Long A, Kelland J, O'Meara S, Sculpher M, Song F, *et al.* Managing menorrhagia. *Qual Health Care* 1995;**4**:218–26.

19. Clark TJ, Khan KS, Foon R, Pattison H, Bryan S, Gupta JK. Quality of life instruments in studies of menorrhagia: a systematic review. *Eur J Obstet Gynecol Reprod Biol* 2002;**104**:96–104.

20. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;**30**:473–83.

21. Jenkinson C, Peto V, Coulter A. Making sense of ambiguity: evaluation in internal reliability and face validity of the SF 36 questionnaire in women presenting with menorrhagia. *Qual Health Care* 1996;**5**:9–12.

22. Lamping DL, Rowe P, Clarke A, Black N, Lessof L. Development and validation of the Menorrhagia Outcomes Questionnaire. *Br J Obstet Gynaecol* 1998;**105**:766–79.

23. Shaw RW, Brickley MR, Evans L, Edwards MJ. Perceptions of women on the impact of menorrhagia on their health using multi-attribute utility assessment. *Br J Obstet Gynaecol* 1998;**105**:1155–9.

24. Crow R, Gage H, Hampson S, Hart J, Kimber A, Storey L, *et al.* The measurement of satisfaction with healthcare: implications for practice from a systematic review of the literature. *Health Technol Assess* 2002;**6**(32).

25. Milsom I, Andersson K, Andersch B, Rybo G. A comparison of flurbiprofen, tranexamic acid, and a levonorgestrel-releasing intrauterine contraceptive device in the treatment of idiopathic menorrhagia. *Am J Obstet Gynecol* 1991;**164**:879–83.

26. Lahteenmaki P, Haukkamaa M, Puolakka J, Riikonen U, Sainio S, Suvisaari J, *et al.* Open randomised study of use of levonorgestrel releasing intrauterine system as alternative to hysterectomy. *BMJ* 1998;**316**:1122–6.

27. Hurskainen R, Teperi J, Rissanen P, Aalto AM, Grenman S, Kivela A, *et al.* Clinical outcomes and costs with the levonorgestrel-releasing intrauterine system or hysterectomy for treatment of menorrhagia – randomized trial 5-year follow-up. *JAMA* 2004;**291**:1456–63.

28. Cooper KG, Jack SA, Parkin DE, Grant AM. Five-year follow up of women randomised to medical management or transcervical resection of the endometrium for heavy menstrual loss: clinical and quality of life outcomes. *BJOG* 2001;**108**:1222–8.

29. Department of Health. *Hospital episode statistics 2000–1*. London: Office for National Statistics; 2002.

30. Reid PC. Endometrial ablation in England – coming of age? An examination of hospital episode statistics 1989/1990 to 2004/2005. *Eur J Obstet Gynecol Reprod Biol* 2007;**135**:191–4.

31. Lethaby A, Shepperd S, Cooke I, Farquhar C. Endometrial resection and ablation versus hysterectomy for heavy menstrual bleeding. *Cochrane Database Syst Rev* 1999;**2**:CD000329.

32. Cameron IM, Mollison J, Pinion SB, Atherton-Naji A, Buckingham K, Torgerson D. A cost comparison of hysterectomy and hysteroscopic surgery for the treatment of menorrhagia. *Eur J Obstet Gynecol Reprod Biol* 1996;**70**:87–92.

33. Farrell SA, Kieser K. Sexuality after hysterectomy. *Obstet Gynecol* 2000;**95**:1045–51.

34. Brown JS, Sawaya G, Thom DH, Grady D. Hysterectomy and urinary incontinence: a systematic review. *Lancet* 2000;**356**:535–9.

35. Kuh DL, Wadsworth M, Hardy R. Women's health in midlife: the influence of the menopause, social factors and health in earlier life. *Br J Obstet Gynaecol* 1997;**104**:923–33.

36. Duffy S, Reid PC, Smith JH, Sharp F. In vitro studies of uterine electrosurgery. *Obstet Gynecol* 1991;**78**:213–20.

37. Duffy S, Reid PC, Sharp F. In-vivo studies of uterine electrosurgery. *Br J Obstet Gynaecol* 1992;**99**:579–82.

38. Goldrath MH, Fuller TA, Segal S. Laser photovaporization of endometrium for the treatment of menorrhagia. *Am J Obstet Gynecol* 1981;**140**:14–19.

39. Davis JA. Hysteroscopic endometrial ablation with the neodymium-YAG laser. *Br J Obstet Gynaecol* 1989;**96**:928–32.

40. Magos AL, Baumann R, Turnbull AC. Transcervical resection of endometrium in women with menorrhagia. *BMJ* 1989;**298**:1209–12.

41. Gannon MJ, Holt EM, Fairbank J, Fitzgerald M, Milne MA, Crystal AM, *et al.* A randomised trial comparing endometrial resection and abdominal hysterectomy for the treatment of menorrhagia. *BMJ* 1991;**303**:1362–4.

42. Dwyer N, Hutton J, Stirrat GM. Randomised controlled trial comparing endometrial resection with abdominal hysterectomy for the surgical treatment of menorrhagia. *Br J Obstet Gynaecol* 1993;**100**:237–43.

43. Pinion SB, Parkin DE, Abramovich DR, Naji A, Alexander DA, Russell IT, *et al.* Randomised trial of hysterectomy, endometrial laser ablation, and transcervical endometrial resection for dysfunctional uterine bleeding. *BMJ* 1994;**309**:979–83.

44. O'Connor H, Broadbent JA, Magos AL, McPherson K. Medical Research Council randomised trial of endometrial resection versus hysterectomy in management of menorrhagia. *Lancet* 1997;**349**:897–901.

45. Crosignani PG, Vercellini P, Apolone G, De Giorgi O, Cortesi I, Meschia M. Endometrial resection versus vaginal hysterectomy for menorrhagia: long-term clinical and quality-of-life outcomes. *Obstet Gynecol* 1997;**177**:95–101.

46. A randomised trial of endometrial ablation versus hysterectomy for the treatment of dysfunctional uterine bleeding: outcome at four years. Aberdeen Endometrial Ablation Trials Group. *Br J Obstet Gynaecol* 1999;**106**:360–6.

47. Overton C, Hargreaves J, Maresh M. A national survey of the complications of endometrial destruction for menstrual disorders: the MISTLETOE study. Minimally Invasive Surgical Techniques – Laser, EndoThermal or Endoresection. *Br J Obstet Gynaecol* 1997;**104**:1351–9.

48. A Scottish audit of hysteroscopic surgery for menorrhagia: complications and follow up. Scottish Hysteroscopy Audit Group. *Br J Obstet Gynaecol* 1995;**102**:249–54.

49. Arieff AI, Ayus JC. Endometrial ablation complicated by fatal hyponatremic encephalopathy. *JAMA* 1993;**270**:1230–2.

50. Rosenberg MK. Hyponatremic encephalopathy after rollerball endometrial ablation. *Anesth Analg* 1995;**80**:1046–8.

51. Loffer FD. Three-year comparison of thermal balloon and rollerball ablation in treatment of menorrhagia. *J Am Assoc Gynecol Laparosc* 2001;**8**:48–54.

52. Loffer FD, Grainger D. Five-year follow-up of patients participating in a randomized trial of uterine balloon therapy versus rollerball ablation for treatment of menorrhagia. *J Am Assoc Gynecol Laparosc* 2002;**9**:429–35.

53. Meyer WR, Walsh BW, Grainger DA, Peacock LM, Loffer FD, Steege JF. Thermal balloon and rollerball ablation to treat menorrhagia: a multicenter comparison. *Obstet Gynecol* 1998;**92**:98–103.

54. Cooper KG, Bain C, Parkin DE. Comparison of microwave endometrial ablation and transcervical resection of the endometrium for treatment of heavy menstrual loss: a randomised trial. *Lancet* 1999;**354**:1859–63.

55. Bain C, Cooper KG, Parkin DE. Microwave endometrial ablation versus endometrial resection: a randomized controlled trial. *Obstet Gynecol* 2002;**99**:983–7.

56. Cooper J, Gimpelson R, Laberge P, Galen D, Garza-Leal JG, Scott J, *et al.* A randomized, multicenter trial of safety and efficacy of the NovaSure system in the treatment of menorrhagia. *J Am Assoc Gynecol Laparosc* 2002;**9**:418–28.

57. Weber AM. Endometrial ablation. *Obstet Gynecol* 2002;**99**:969–70.

58. Sowter MC, Lethaby A, Singla AA. Pre-operative endometrial thinning agents before endometrial destruction for heavy menstrual bleeding. *Cochrane Database Syst Rev* 2002;**3**:CD001124.

59. Vancaillie T. Electrocoagulation of the endometrium. In Lewis BV, Magos AL, editors. *Endometrial ablation*. Edinburgh: Churchill Livingstone; 1993. pp. 133–42.

60. Kochli OR. Endometrial ablation in the year 2000 – do we have more methods than indications? *Contrib Gynecol Obstet* 2000;**20**:91–120.

61. Lethaby A, Hickey M, Garry R. Endometrial destruction techniques for heavy menstrual bleeding. *Cochrane Database Syst Rev* 2005;**4**:CD001501.

62. Sharp NC, Cronin N, Feldberg I, Evans M, Hodgson D, Ellis S. Microwaves for menorrhagia: a new fast technique for endometrial ablation. *Lancet* 1995;**346**:1003–4.

63. Hodgson DA, Feldberg IB, Sharp N, Cronin N, Evans M, Hirschowitz L. Microwave endometrial ablation: development, clinical trials and outcomes at three years. *Br J Obstet Gynaecol* 1999;**106**:684–94.

64. Abbott J, Hawe J, Hunter D, Garry R. A double-blind randomized trial comparing the Cavaterm and the NovaSure endometrial ablation systems for the treatment of dysfunctional uterine bleeding. *Fertil Steril* 2003;**80**:203–8.

65. Bongers MY, Bourdrez P, Mol BW, Heintz AP, Brolmann HA. Randomised controlled trial of bipolar radio-frequency endometrial ablation and balloon endometrial ablation. *BJOG* 2004;**111**:1095–102.

66. Rybo G. Clinical and experimental studies on menstrual blood loss. *Acta Obstet Gynecol Scand* 1966;**45**(Suppl. 7):1–23.

67. Bain C, Cooper KG, Parkin DE. A partially randomised patient preference trial of microwave endometrial ablation using local anaesthesia and intravenous sedation or general anaesthesia: a pilot study. *Gynaecol Endoscopy* 2001;**10**:223–8.

68. Cooper KG, Bain C, Lawrie L, Parkin DE. A randomised comparison of microwave endometrial ablation with transcervical resection of the endometrium; follow up at a minimum of five years. *BJOG* 2005;**112**:470–5.

69. Lethaby A, Farquhar C. Treatments for heavy menstrual bleeding. *BMJ* 2003;**327**:1243–4.

70. Abbott J., Garry R. The surgical management of menorrhagia. *Hum Reprod Update* 2002;**8**:78.

71. Reid PC, Mukri F. Trends in number of hysterectomies performed in England for menorrhagia: examination of health episode statistics, 1989 to 2002–3. *BMJ* 2005;**330**:938–9.

72. Lethaby AE, Cooke I, Rees M. Progesterone or progestogen-releasing intrauterine systems for heavy menstrual bleeding. *Cochrane Database Syst Rev* 2005;**4**:CD002126.

73. Marjoribanks J, Lethaby A, Farquhar C. Surgery versus medical therapy for heavy menstrual bleeding. *Cochrane Database Syst Revs* 2006;**2**:CD003855.

74. Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data: is there a difference? *Lancet* 1993;**341**:418–22.

75. Higgins JPT, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1.* Oxford: John Wiley & Sons Ltd; 2008.

76. Stewart LA, Tierney JF. To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Eval Health Prof* 2002;**25**:76–97.

77. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, *et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009;**62**:e1–34.

78. Zupi E, Zullo F, Marconi D, Sbracia M, Pellicano M, Solima E, *et al.* Hysteroscopic endometrial resection versus laparoscopic supracervical hysterectomy for menorrhagia: a prospective randomized trial. *Obstet Gynecol* 2003;**188**:7–12.

79. Corson SL. A multicenter evaluation of endometrial ablation by Hydro ThermAblator and rollerball for treatment of menorrhagia. *J Am Assoc Gynecol Laparosc* 2001;**8**:359–67.

80. Crosignani PG, Vercellini P, Mosconi P, Oldani S, Cortesi I, De Giorgi O. Levonorgestrel-releasing intrauterine device versus hysteroscopic endometrial resection in the treatment of dysfunctional uterine bleeding. *Obstet Gynecol* 1997;**90**:257–63.

81. Barrington JW, Arunkalaivanan AS, Abdel-Fattah M. Comparison between the levonorgestrel intrauterine system (LNG-IUS) and thermal balloon ablation in the treatment of menorrhagia. *Eur J Obstet Gynecol Reprod Biol* 2003;**108**:72–4.

82. Group E. EuroQol – a new facility for the treatment of health-related quality of life. *Health Policy* 1990;**16**:199–208.

83. Riley RD, Lambert PC, Staessen JA, Wang J, Gueyffier F, Thijs L, *et al.* Meta-analysis of continuous outcomes combining individual patient data and aggregate data. *Stat Med* 2008;**27**:1870–93.

84. Early Breast Cancer Trialists' Collaborative Group. *Treatment of early breast cancer: worldwide evidence, 1985–1990*. Oxford, UK: Oxford University Press; 1990.

85. Higgins JPT, Thompson SG, Deeks JD, Altman DG. Measuring inconsistency in meta-analysis. *BMJ* 2003;**327**:557–60.

86. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, *et al.* Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. *Br J Cancer* 1977;**35**:1–39.

87. Fleiss JL. The statistical basis of meta-analysis. *Stat Methods Med Res* 1993;**2**:121–45.

88. Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R, *et al.* Indirect comparisons of competing interventions. *Health Technol Assess* 2005;**9**(26).

89. Hasselblad V. Meta-analysis of multitreatment studies. *Med Decis Making* 1998;**18**:37–43.

90. Whitehead A. *Meta-analysis of controlled clinical trials.* Chichester, UK: John Wiley & Sons; 2002.

91. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986;**42**:121–30.

92. Dickersin K, Munro MG, Clark M, Langenberg P, Scherer R, Frick K, *et al.* Hysterectomy compared with endometrial ablation for dysfunctional uterine bleeding: a randomized controlled trial. *Obstet Gynecol* 2007;**110**:1279–89.

93. Hurskainen R, Teperi J, Rissanen P, Aalto AM, Grenman S, Kivela A, *et al.* Quality of life and cost-effectiveness of levonorgestrel-releasing intrauterine system versus hysterectomy for treatment of menorrhagia: a randomised trial. *Lancet* 2001;**357**:273–7.

94. Hawe J, Abbott J, Hunter D, Phillips G, Garry R. A randomised controlled trial comparing the Cavaterm endometrial ablation system with the Nd:YAG laser for the treatment of dysfunctional uterine bleeding. *BJOG* 2003;**110**:350–7.

95. van Zon-Rabelink IA, Vleugels MPH, Merkus HMWM, de Graaf R. Efficacy and satisfaction rate comparing endometrial ablation by rollerball electrocoagulation to uterine balloon thermal ablation in a randomised controlled trial. *Eur J Obstet Gynecol Reprod Biol* 2004;**114**:97–103.

96. Soysal ME, Soysal SK, Vicdan K. Thermal balloon ablation in myoma-induced menorrhagia under local anesthesia. *Gynecol Obstet Invest* 2001;**51**:128–33.

97. Romer T. [Therapy of recurrent menorrhagia – Cavaterm balloon coagulation versus roller-ball endometrium coagulation – a prospective randomized comparative study]. *Zentralbl Gynakol* 1998;**120**:511–14.

98. Duleba AJ, Heppard MC, Soderstrom RM, Townsend DE. A randomized study comparing endometrial cryoablation and rollerball electroablation for treatment of dysfunctional uterine bleeding. *J Am Assoc Gynecol Laparosc* 2003;**10**:17–26.

99. Cooper J, Anderson TL, Fortin CA, Jack SA, Plentl MB. Microwave endometrial ablation vs. rollerball electroablation for menorrhagia: a multicenter randomized trial. *J Am Assoc Gynecol Laparosc* 2004;**11**:394–403.

100. Perino A, Castelli A, Cucinella G, Bionda A, Pane A, Venezia R. A randomized comparison of endometrial laser intrauterine thermotherapy and hysteroscopic endometrial resection. *Fertil Steril* 2004;**82**:731–4.

101. Corson SL, Brill AI, Brooks PG, Cooper JM, Indman PD, Liu JH, *et al.* One-year results of the vesta system for endometrial ablation. *J Am Assoc Gynecol Laparosc* 2000;**7**:489–97.

102. Pellicano M, Guida M, Acunzo G, Cirillo D, Bifulco G, Nappi C. Hysteroscopic transcervical endometrial resection versus thermal destruction for menorrhagia: a prospective randomized trial on satisfaction rate. *Obstet Gynecol* 2002;**187**:545–50.

103. Brun JL, Burlet G, Galand B, Quereux C, Bernard P. Cavaterm thermal balloon endometrial ablation versus hysteroscopic endometrial resection to treat menorrhagia: the French, multicenter, randomized study. *J Minim Invasive Gynecol* 2006;**13**:424–30.

104. Malak KA. Management of menorrhagia with the levonorgestrel intrauterine system versus endometrial resection. *Gynaecol Surg* 2006;**3**:275–80.

105. Kittelsen N, Istre O. A randomized study comparing levonorgestrel intrauterine system (LNG IUS) and transcervical resection of the endometrium (TCRE) in the treatment of menorrhagia: preliminary results. *Gynaecol Endosc* 1998;**7**:61–5.

106. Soysal M, Soysal S, Ozer S. A randomized controlled trial of levonorgestrel releasing IUD and thermal balloon ablation in the treatment of menorrhagia. *Zentralbl Gynakol* 2002;**124**:213–19.

107. Busfield RA, Farquhar CM, Sowter MC, Lethaby A, Sprecher M, Yu Y, *et al.* A randomised trial comparing the levonorgestrel intrauterine system and thermal balloon ablation for heavy menstrual bleeding. *BJOG* 2006;**113**:257–63.

108. Shaw RW, Symonds IM, Tamizian O, Chaplain J, Mukhopadhyay S. Randomised comparative trial of thermal balloon ablation and levonorgestrel intrauterine system in patients with idiopathic menorrhagia. *Aust NZ J Obstet Gynaecol* 2007;**47**:335–40.

109. Tam WH, Yuen PM, Shan Ng DP, Leung PL, Lok IH, Rogers MS. Health status function after treatment with thermal balloon endometrial ablation and levnonorgestrel intrauterine system for idiopathic menorrhagia: a randomised study. *Gynecol Obstet Invest* 2006;**62**:84–8.

110. Bhattacharya S, Cameron IM, Parkin DE, Abramovich DR, Mollison J, Pinion SB, *et al.* A pragmatic randomised comparison of transcervical resection of the endometrium with endometrial laser ablation for the treatment of menorrhagia. *BJOG* 1997;**104**:601–7.

111. Clark TJ. A randomised controlled trial to compare the effectiveness of outpatient endometrial ablation techniques (NovaSure versus Thermachoice) in the treatment of menorrhagia. *Unpublished* 2008.

112. Sambrook AM, Cooper KG, Campbell MK, Cook JA. Clinical outcomes from a randomised comparison of microwave endometrial ablation with thermal balloon endometrial ablation for the treatment of heavy menstrual bleeding. *BJOG* 2009;**116**:1038–45.

113. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999;**354**:1896–900.

114. Brown PM, Farquhar CM, Lethaby A, Sadler LC, Johnson NP. Cost-effectiveness analysis of levonorgestrel intrauterine system and thermal balloon ablation for heavy menstrual bleeding. *BJOG* 2006;**113**:797–803.

115. Cooper KG, Jack SA, Parkin DE, Grant AM. Five-year follow up of women randomised to medical management or transcervical resection of the endometrium as treatment for heavy menses. *Br J Obstet Gynaecol* 2001;**108**:1222–8.

116. Sambrook AM, Bain C, Parkin DE, Cooper KG. A randomised comparison of microwave endometrial ablation with transcervical resection of the endometrium: follow up at a minimum of 10 years. *BJOG* 2009;**116**:1033–7.

117. Stewart LA, Clarke MJ. Practical methodology of meta-analyses (overviews) using updated individual patient data. Cochrane Working Group. *Stat Med* 1995;**14**:2057–79.

118. Spilsbury K, Hammond I, Bulsara M, Semmens JB. Morbidity outcomes of 78,577 hysterectomies for benign reasons over 23 years. *BJOG* 2008;**115**:1473–83.

119. Cole S. Scottish maternity and neonatal records. In Chalmers T, McIlwraith GM, editors. *Perinatal audit and surveillance, Proceedings of the Eighth Study Group of the RCOG London.* London: Royal College of Obstetricians and Gynaecologists; 1980. pp. 39–51.

120. Morris R, Carstairs V. Which deprivation? A comparison of selected deprivation indexes. *J Public Health Med* 1991;**13**:318–26.

121. Kuh D, Stirling S. Socioeconomic variation in admission for diseases of female genital system and breast in a national cohort aged 15–43. *BMJ* 1995;**311**:840–3.

122. Luoto R, Keskimaki I, Reunanen A. Socioeconomic variations in hysterectomy: evidence from a linkage study of the Finnish hospital discharge register and population census. *J Epidemiol Community Health* 1997;**51**:67–73.

123. Thakar R, Sultan AH. Hysterectomy and pelvic organ dysfunction. *Best Pract Res Clin Obstet Gynaecol* 2005;**19**:403–18.

124. McPherson K, Herbert A, Judge A, Clarke A, Bridgman S, Maresh M, *et al.* Self-reported bladder function five years post-hysterectomy. *J Obstet Gynaecol* 2005;**25**:469–75.

125. Krogh RA, Lauszus FF, Guttorm E, Rasmussen K. Surgery and cancer after endometrial resection. Long-term follow-up on menstrual bleeding and hormone treatment by questionnaire and registry. *Arch Gynecol Obstet* 2009;**280**:911–16.

126. Loft A, Lidegaard O, Tabor A. Incidence of ovarian cancer after hysterectomy: a nationwide controlled follow up. *Br J Obstet Gynaecol* 1997;**104**:1296–301.

127. Kilkku P, Hirvonen T, Gronroos M. Supra-vaginal uterine amputation vs. abdominal hysterectomy: the effects on urinary symptoms with special reference to pollakisuria, nocturia and dysuria. *Maturitas* 1981;**3**:197–204.

128. Virtanen H, Makinen J, Tenho T, Kiilholma P, Pitkanen Y, Hirvonen T. Effects of abdominal hysterectomy on urinary and sexual symptoms. *Br J Urol* 1993;**72**:868–72.

129. Gimbel H. Total or subtotal hysterectomy for benign uterine diseases? A meta-analysis. *Acta Obstet Gynecol Scand* 2007;**86**:133–44.

130. Okaro EO, Jones KD, Sutton C. Long term outcome following laparoscopic supracervical hysterectomy. *BJOG* 2001;**108**:1017–20.

131. Blandon RE, Bharucha AE, Melton LJ, 3rd, Schleck CD, Babalola EO, Zinsmeister AR, *et al.* Incidence of pelvic floor repair after hysterectomy: a population-based cohort study. *Am J Obstet Gynecol* 2007;**197**:664, e1–7.

132. Sculpher M. A cost–utility analysis of abdominal hysterectomy versus transcervical endometrial resection for the surgical treatment of menorrhagia. *Int J Technol Assess Health Care* 1998;**14**:302–19.

133. Government Actuary's Department. *Life tables*. 2009. URL: www.gad.gov.uk/demography_Data/life_Tables/ (accessed September 2009).

134. Parkin DE. Microwave endometrial ablation (MEA (TM)): a safe technique? Complication data from a prospective series of 1400 cases. *Gynaecol Endosc* 2000;**9**:385–8.

135. MacLean-Fraser E, Penava D, Vilos GA. Perioperative complication rates of primary and repeat hysteroscopic endometrial ablations. *J Am Assoc Gynecol Laparoscop* 2002;**9**:175–7.

136. Curtis L. *Unit costs of health and social care*. Canterbury: Personal Social Services Unit, University of Kent; 2008.

137. PSSRU. URL: www.Pssru.Ac.Uk/ (accessed September 2009).

138. British Medical Association and Royal Pharmaceutical Society of Great Britain. *British national formulary.* No. 58, September 2009. London: BMA and RPS; 2009.

139. Department of Health. *NHS reference costs 2008–2009.* 2010. URL: www.dh.gov.uk/en/publicationsandstatistics/publications/publicationspolicyandguidance/DH_111591 (accessed September 2009).

140. Clegg JP, Guest JF, Hurskainen R. Cost–utility of levonorgestrel intrauterine system compared with hysterectomy and second generation endometrial ablative techniques in managing patients with menorrhagia in the UK. *Curr Med Res Opin* 2007;**23**:1637–48.

141. Fröberg CE. *Introduction to numerical analysis.* Reading, MA: Addison-Wesley; 1969.

142. Briggs A, Sculpher M, Buxton M. Uncertainty in the economic-evaluation of health-care technologies – the role of sensitivity analysis. *Health Econ* 1994;**3**:95–104.

143. Briggs AH. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* 2000;**17**:479–500.

144. Seymour J, Wallage S, Graham W, Parkin D, Cooper K. The cost of microwave endometrial ablation under different anaesthetic and clinical settings. *BJOG* 2003;**110**:922–6.

145. Torrance GW. Measurement of health state utilities for economic appraisal. *J Health Econ* 1986;**5**:1–30.

146. Thompson SG, Barber JA. How should cost data in pragmatic randomised trials be analysed? *Br Med J* 2000;**320**:1197–200.

147. Dickersin K, Munro M, Langenberg P, Scherer R, Frick KD, Weber AM, *et al.* Surgical Treatments Outcomes Project for Dysfunctional Uterine Bleeding (STOP-DUB): design and methods. *Control Clin Trials* 2003;**24**:591–609.

148. Sculpher MJ, Dwyer N, Byford S, Stirrat GM. Randomised trial comparing hysterectomy and transcervical endometrial resection: effect on health related quality of life and costs two years after surgery. *BJOG* 1996;**103**:142–9.

149. van Zon-Rabelink IA, Vleugels MP, Merkus HM, de Graaf R. Endometrial ablation by rollerball electrocoagulation compared to uterine balloon thermal ablation. Technical and safety aspects. *Eur J Obstet Gynecol Reprod Biol* 2003;**110**:220–3.

150. Rauramo I, Elo I, Istre O. Long-term treatment of menorrhagia with levonorgestrel intrauterine system versus endometrial resection. *Obstet Gynecol* 2004;**104**:1314–21.

Appendix 1

Full electronic search strategy used in the systematic review

Search strategy for population

#1 menorrhagia/all subheadings
#2 hypermenorrhea/all subheadings
#3 excessive NEAR ("menstrual bleeding" OR "menstrual blood loss")
#4 dysfunctional NEAR ("uterine bleeding" OR "menstrual bleeding")
#5 heavy NEAR ("menstrual bleeding" OR "menstrual blood loss")
#6 "iron deficient anaemia"
#7 (#3 OR #4 OR #5 OR #6) in TI, AB
#8 #1 OR #2 OR #7

Search strategy for interventions

Hysterectomy

#1 EXPLODE "hysterectomy"/all sub-headings
#2 "vaginal hysterectomy"/all sub-headings
#3 "total abdominal hysterectomy"
#4 "subtotal abdominal hysterectomy"
#5 "laparoscopic hysterectomy"
#6 #1 OR #2 OR #3 OR #4 OR #5

Ablation

#1 EXPLODE "hysteroscopy"/all sub-headings
#2 ("transcervical resection") NEAR "endometrium"
#3 "TCRE"
#4 "endometrial ablation"
#5 "laser ablation"
#6 "electrosurgery"
#7 "rollerball"
#8 "thermal balloon"
#9 "hypertherm$"
#10 "thermotherapy"
#11 "photodynamic therapy"
#12 "phototherapy"
#13 "cryoablation"
#14 "microwave ablation"
#15 "radiofrequency"
#16 "saline irrigation"
#17 "laser interstitial"
#18 "Thermachoice"
#19 "Cavaterm"

#20 “ELITT”
#21 “Vesta”
#22 “Novasure”
#23 “Microsulis”
#24 “Cryogen”

Mirena

#1 EXPLODE “contraceptive”/all sub-headings
#2 “Mirena® coil”/all sub-headings
#3 “levonorgestrel”
#4 “intra uterine device”
#5 #1 OR #2 OR #3 OR #4

Search strategy for randomised controlled trials

#1 Randomized Controlled Trial IN PT.
#2 Controlled Clinical Trial IN PT.
#3 Randomized Controlled Trials IN SH
#4 Random Allocation IN SH.
#5 Double Blind Method IN SH
#6 Single Blind Method IN SH
#7 (#1 OR #2 OR #3 OR #4 OR #5 OR #6)
#8 Animal in SH NOT Human in SH.
#9 #7 not # 8
#10 Clinical Trial IN PT.
#11 EXPLODE Clinical Trials/all sub-headings
#12 (clin$NEAR trial$) IN TI, AB
#13 ((singl$OR doubl$OR trebl$OR tripl$) NEAR (blind$OR mask$)) IN TI, AB
#14 Placebos IN SH
#15 placebo$IN TI, AB
#16 random$IN TI, AB
#17 Research Design IN SH
#18 #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17
#19 #18 NOT #8
#20 #19 NOT #9
#21 Comparative Study IN SH
#22 EXPLORE Evaluation Studies/all-sub-headings
#23 Follow Up Studies IN SH
#24 Prospective Studies IN SH
#25 (control$OR prospectiv$OR volunteer$) IN TI, AB
#26 #21 OR #22 OR #23 OR #24 OR #25
#27 #26 NOTt #8
#28 #27 NOT (#9 OR #20)
#29 #9 OR #20 OR #28

Appendix 2

Characteristics of studies included in the systematic review of randomised trials comparing hysterectomy, endometrial ablation and Mirena for heavy menstrual bleeding

Paper/number of women randomised	Patients	Intervention	Stated key outcome measures	Patient satisfaction and how it was measured	IPD received?
Hysterectomy vs first-generation EA					
Dickersin *et al.*, 2007[92] (design and methods paper also published[147]) Raw data available *n* = 237	Women with DUB. Up to 3 fibroids allowed, must each be smaller than 3 cm	EA vs hysterectomy	Major problem solved (primary outcome) Resolution of problem Bleeding Pain Fatigue QoL Adverse events Reoperation rate Follow-up reported at 12 months, 2 and 5 years; IPD at 6 months, 3 and 4 years also received	Women were asked if their major problem was solved from baseline Answers were given using the following scale: Yes No	Yes
Zupi *et al.*, 2003[78] Raw data available *n* = 203	Women with HMB. Fibroids excluded	TCRE vs laparoscopic supracervical hysterectomy	Primary outcome unclear Duration of hospitalisation Period of convalescence Perioperative complications Resumption of usual activities QoL Follow-up reported at 3 months, 1 and 2 years	No comparable measure	Yes
Crosignani *et al.*,1997[45] *n* = 92	Women with HMB < 50 years old with a mobile uterus smaller than a 12-week pregnancy. Fibroids excluded if > 3 cm	TCRE vs vaginal hysterectomy	Satisfaction (primary outcome) Improvement in menstrual blood loss Operating time Complications Postoperative hospital stay Resumption of usual activities Resumption of work activities QoL Follow-up reported at 2 years	Women were asked how satisfied they were with their operation Answers were given using the following scale: Very satisfied Satisfied Uncertain Dissatisfied Very dissatisfied	No

continued

Paper/number of women randomised	Patients	Intervention	Stated key outcome measures	Patient satisfaction and how it was measured	IPD received?
O'Connor *et al.*, 1997[44] Raw data available *n*=202	Women with symptomatic HMB. Fibroids excluded if larger than 5 cm	TCRE vs abdominal + vaginal hysterectomy	Satisfaction (primary outcome) Need for further surgery QoL Duration of surgery Duration of hospital stay Operative and postoperative complications Resumption of work activities Resumption of usual activities Resumption of sexual activities Follow-up reported at 3 months, then 1, 2 and 3 years	Women were asked how satisfied they were with their treatment Answers were given using the following scale: Very satisfied Satisfied Not sure Dissatisfied Very dissatisfied	Yes
Pinion *et al.*,1994[43] Raw data available *n*=204	Women who would have otherwise had a hysterectomy for HMB. IPD showed that fibroids were included; exact eligibility details regarding this parameter not given in paper	TCRE + laser vs abdominal hysterectomy	Satisfaction (primary outcome) Operative complications Postoperative recovery Relief of menstrual symptoms Relief of other symptoms Follow-up reported at 6 and 12 months	Women were asked how satisfied they were with their treatment Answers were given using the following scale: Very satisfied Moderately satisfied Dissatisfied Very dissatisfied	Yes
Dwyer *et al.*, 1993[42] (health economics papers also published[132,148]) Raw data available *n*=200	Women needing surgical treatment for HMB. IPD showed that fibroids were included; exact eligibility details regarding this parameter not given in paper	TCRE vs abdominal hysterectomy	Satisfaction (primary outcome) Postoperative complications Duration of operation Length of hospital stay Resumption of work activities Resumption of usual activities Resumption of sexual activities Changes in pre-menstrual symptoms QoL Need for further surgery Total health service resource cost Follow-up reported at 4 months and 2 years	Women were asked how satisfied they were with their operation Answers were given using the following scale: Very satisfied Quite satisfied Not very satisfied Dissatisfied	Yes
Gannon *et al.*, 1991[41] Raw data available *n*=54	Women with HMB. Fibroids excluded	TCRE vs abdominal hysterectomy	Primary outcome unclear Length of operating time Hospitalisation Recovery Cost of surgery Change in menstrual blood loss Postoperative complications Need for further surgery Resource cost of surgery Follow-up reported at 12 months	No comparable measure	Yes

Paper/number of women randomised	Patients	Intervention	Stated key outcome measures	Patient satisfaction and how it was measured	IPD received?
Hysterectomy vs Mirena					
Hurskainen *et al.*, 2001[93] (5-year follow-up study also published[27]) Raw data available *n* = 236	Women with HMB. Fibroids excluded	Mirena vs hysterectomy (abdominally, vaginally or laparoscopically)	QoL (EQ-5D) (primary outcome) QoL (SF-36) Cost-effectiveness Adverse events General health (visual analogue scale, VAS) Anxiety/depression Sexual functioning Follow-up reported at 12 months and 5 years; IPD at 6 months also received	No comparable measure	Yes
First- vs second-generation EA					
Brun *et al.*, 2006[103] Raw data available *n* = 62	Women with HMB unresponsive to medical treatment. Submucous fibroids excluded, other fibroids included (further details not given)	TCRE vs thermal balloon (Cavaterm)	Amenorrhoea rate (primary outcome) Satisfaction PBAC (Higham blood loss) score Operative time Discharge time Complication rate Resumption of normal activities Follow-up reported at 6 and 12 months; IPD at 3 months also received	Refers to 'satisfaction rate' Answers were given using the following scale: Excellent Good Moderate Bad	Yes
Cooper *et al.*, 2004[99] Raw data available *n* = 322	Women with documented HMB due to benign causes. Fibroids excluded if > 3 cm	RB vs microwave	Satisfaction Amenorrhoea rate Duration of procedure Anaesthesia Type of anaesthesia Device-related complications Adverse events Dysmenorrhoea QoL questionnaire (SF-36) Acceptability of treatment Follow-up reported at 3, 6 and 12 months	Women were asked how satisfied they were with their treatment Answers were given using the following scale: Very satisfied Satisfied Dissatisfied	Yes
Perino *et al.*, 2004[100] *n* = 116	Women with abnormal uterine bleeding. Not stated if fibroids were excluded	TCRE vs ELITT	Amenorrhoea rate (primary outcome) Satisfaction Bleeding status Intraoperative complication rate Duration of procedure Pain Further treatment with hysterectomy Follow-up reported at 12 months and 3 years	Refers to 'patient satisfaction' Answers were given using the following scale: Very satisfied Satisfied Dissatisfied	No

continued

Paper/number of women randomised	Patients	Intervention	Stated key outcome measures	Patient satisfaction and how it was measured	IPD received?
Duleba *et al.*, 2003[98] *n* = 279	Women with HMB due to benign causes. Fibroids excluded if > 2 cm	RB vs endometrial cryoablation	PBAC (Higham blood loss) score (primary outcome) Satisfaction Bleeding Pain Adverse events Anaesthesia Pre-menstrual symptoms Follow-up reported at 12 months	Women were asked how satisfied they were with the outcome of the procedure Answers were given using the following scale: Very Slightly Not at all	No
Hawe *et al.*, 2003[94] Raw data available *n* = 72	Women with DUB requesting conservative surgical management of their condition. Fibroids excluded	Nd:Yag laser vs thermal balloon (Cavaterm)	Amenorrhoea rate (primary outcome) Satisfaction Effect on blood loss QoL Sexual activity Acceptability of procedure Follow-up reported at 6 and 12 months	Women were asked how satisfied they were with their treatment Answers were given using the following scale: Very satisfied Moderately satisfied Dissatisfied Very dissatisfied	Yes
van Zon-Rabelink *et al.*, 2004[95] (technical safety report also published[149]) Raw data available *n* = 139	Women with DUB. IPD showed that fibroids were included; exact eligibility details regarding this parameter not given in paper	RB vs thermal balloon	PBAC (Higham blood loss) score (primary outcome) Satisfaction QoL Menstrual status Follow-up reported at 6 and 12 months and 2 years	Refers to 'patient satisfaction' Answers were given using the following scale: Satisfied Not satisfied	Yes
Cooper *et al.*, 2002[56] *n* = 265	Women with symptomatic HMB. Fibroids excluded	Wire loop resection + RB vs bipolar radiofrequency (NovaSure)	PBAC (Higham blood loss) score (primary outcome) Satisfaction Procedure time Sedation Intraoperative complications Postoperative complications Follow-up reported at 6 and 12 months	Women were asked how satisfied they were with the outcome of the procedure No precise information was given on the scale used to answer this question and IPD were not received. Percentage of women very satisfied or satisfied was quoted	No
Pellicano *et al.*, 2002[102] *n* = 82	Women with HMB unresponsive to medical treatment. Fibroids excluded	TCRE vs thermal destruction (Cavaterm)	Satisfaction (primary outcome) Operative time Discharge time Complication rate Reintervention rate Resumption of normal activities Follow-up reported at 3 and 12 months and 2 years	Women were asked about the improvement of their health state after the procedure Answers were given using the following scale: Excellent Good Moderate No improvement	No

Paper/number of women randomised	Patients	Intervention	Stated key outcome measures	Patient satisfaction and how it was measured	IPD received?
Corson, 2001[79] *n* = 276	Women with HMB due to benign causes. Fibroids excluded if > 4 cm	RB vs HA	PBAC (Higham blood loss) score (primary outcome) Amenorrhoea rate Adverse events Need for further surgery Operative complications Follow-up reported at 6 and 12 months	No comparable measure	Yes
Soysal *et al.*, 2001[96] *n* = 96	Menorrhagic women over 40 with a mobile myomatous uterus smaller than 12-week pregnancy. Fibroids excluded if > 3 cm	RB vs thermal balloon	PBAC (Higham blood loss) score (primary outcome) Satisfaction Duration of procedure Complication rates Postoperative pain scores Amenorrhoea rates Follow-up reported at 12 months	Women were asked how satisfied they were with their operation Answers were given using the following scale: Very satisfied Satisfied Dissatisfied	No
Corson *et al.*, 2000[101] *n* = 276	Women with HMB, without organic uterine disease, who failed or poorly tolerated medical therapy. Fibroids excluded if > 2 cm	TCRE + RB vs thermal balloon (Vesta)	PBAC (Higham blood loss) score (primary outcome) Amenorrhoea Adverse events QoL Follow-up reported at 12 months and 2 years	No comparable measure.	No
Cooper *et al.*, 1999[54] (2-year[55] and 5-year[115] follow-up study also published) Raw data available *n* = 263	Women referred for EA surgery. Fibroids included; exact eligibility details regarding this parameter not given in paper	TCRE + RB vs microwave	Satisfaction (primary outcome) Acceptability of treatment Menstrual status QoL Morbidity Duration of procedure Intraoperative complications Postoperative pain relief Postoperative stay. Absence from work Follow-up reported at 12 months, 2 years, 5 years and 10 years	Women were asked how satisfied they were with their treatment Answers were given using the following scale: Totally satisfied Generally satisfied Fairly satisfied Fairly dissatisfied Generally dissatisfied Totally dissatisfied	Yes
Meyer *et al.*, 1998[53] Raw data available *n* = 275	Women with HMB. Fibroids excluded	RB vs thermal balloon (ThermaChoice)	PBAC (Higham blood loss) score (primary outcome) Satisfaction Improvement in dysmenorrhoea symptoms Inability to work Complication rate Duration of procedure Requirement for additional surgery Follow-up reported at 3, 6 and 12 months	Women were asked how satisfied they were with their treatment Answers were given using the following scale: Very satisfied Satisfied Not satisfied	Yes

continued

Paper/number of women randomised	Patients	Intervention	Stated key outcome measures	Patient satisfaction and how it was measured	IPD received?
Romer, 1998[97] *n* = 20	Women with recurrent therapy for refractory HMB. Fibroids excluded (intrauterine abnormalities excluded, so assumed this included fibroids)	RB vs thermal balloon (Cavaterm)	Amenorrhoea rate (primary outcome) Hypomenorrhoea rate Follow-up reported at 12 months	No comparable measure	No
Mirena vs first-generation EA					
Malak, 2006[104] *n* = 60	Women with excessive uterine bleeding. Up to 3 fibroids allowed, must each be < 3 cm	TCRE vs Mirena	Primary outcome unclear PBAC (Higham blood loss) score LNG IUS discontinuation rate Effect of menstrual bleeding on general well-being, work performance, physical activity and sexual activity assessed using VAS Follow-up reported at 12 months	No comparable measure	No
Kittelsen and Istre, 1998[105] (long-term follow-up paper also published[150]) *n* = 60	Women with HMB. Fibroids excluded	TCRE vs Mirena	Primary outcome unclear QoL Additional treatments received Adverse events Follow-up reported at 12 months, 2 years and 3 years	No comparable measure	No
Crosignani *et al.*, 1997[80] *n* = 70	Women with DUB. Fibroids excluded	TCRE vs Mirena	Primary outcome unclear Satisfaction Reduction in menstrual bleeding Health-related QoL Amenorrhoea rates Additional treatments Adverse events Follow-up reported at 6 and 12 months	Women were asked how satisfied they were with their treatment Answers were given using the following scale: Very satisfied Satisfied Uncertain Dissatisfied	No
Mirena vs second-generation EA					
Shaw *et al.*, 2007[108] *n* = 66	Women with HMB. Fibroids excluded	Thermal balloon vs Mirena	PBAC (Higham blood loss) score (primary outcome) Satisfaction Continuation with treatment Hysterectomy rates Follow-up reported at 3, 6, 9 and 12 months, and 2 years	Women were asked for their perception of their treatment effect Answers were given using the following scale: Very good Good Poor	No

Paper/number of women randomised	Patients	Intervention	Stated key outcome measures	Patient satisfaction and how it was measured	IPD received?
Tam *et al.*, 2006[109] Raw data available *n* = 44	Women with excessive menstrual bleeding attending the outpatient gynaecology clinic. IPD showed that fibroids were included; exact eligibility details regarding this parameter not given in paper	Thermal balloon vs Mirena	Primary outcome unclear Health status function SF-36 Follow-up reported at 12 months; IPD at 6 months also received	No comparable measure	Yes
Busfield *et al.*, 2006[107] (cost-effectiveness paper carried out by Brown *et al.*, 2006[114]) Raw data available *n* = 79	Women with HMB. Fibroids excluded if > 3 cm	Thermal balloon vs Mirena	PBAC (Higham blood loss) score (primary outcome) Satisfaction QoL Menstrual symptoms Adverse events Treatment failures Follow-up reported at 3, 6 and 12 months, and 2 years	Women were asked if the menstrual symptoms had been successfully treated Answers were given using the following scale: Definitely yes Probably yes Not sure Probably no Definitely no	Yes
Barrington *et al.*, 2003[81] Raw data available *n* = 50	Women with HMB. Fibroids excluded	Thermal balloon vs Mirena	Primary outcome unclear PBAC (Higham blood loss) score Amenorrhoea Follow-up reported at 6 months	No comparable measure	Yes
Soysal *et al.*, 2002[106] *n* = 72	Women with dysfunctional HMB. Fibroids excluded if > 2 cm	Thermal balloon vs Mirena	PBAC (Higham blood loss) score (primary outcome) Satisfaction Health-related QoL Additional treatments Adverse events Follow-up reported at 12 months	Women were asked about their degree of satisfaction/ recommendation Answers were given using the following scale: Highly recommends Recommends Did not know Did not recommend	No

Appendix 3

Quality of studies included in the systematic review of randomised trials comparing hysterectomy, endometrial destruction and Mirena for heavy menstrual bleeding

Paper	Was randomisation adequate?	Was the target population described adequately?	Was the sample size calculation reported?	Were the two populations comparable at baseline?	Was an ITT analysis reported?	Was the follow-up > 80%?	Was compliance with allocated treatment > 80% in both arms at 12 months?
Hysterectomy vs first-generation EA							
Dickersin *et al.*, 2007[92]	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zupi *et al.*, 2003[78]	Unclear, not stated	Yes	Yes	Yes	No	No	Yes
Crosignani *et al.*, 1997[45]	Yes	Yes	Yes	Yes	Yes	Yes	Yes
O'Connor *et al.*, 1997[44]	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pinion *et al.*, 1994[43]	Unclear, not stated	No	Yes	Yes	Yes	Yes	No
Dwyer *et al.*, 1993[42]	Unclear, not stated	Yes	Yes	Yes	No	Yes	Yes
Gannon *et al.*, 1991[41]	Unclear, not stated	Yes	No	Yes	No	Yes	Yes
Hysterectomy vs Mirena							
Hurskainen *et al.*, 2001[93]	Yes	Yes	Yes	Yes	Yes	Yes	No
First- vs second-generation EA							
Brun *et al.*, 2006[103]	Yes	Yes	Yes	Yes	No	No	Yes
Cooper *et al.*, 2004[99]	Unclear, not stated	Yes	Yes	Yes	Yes	Yes	Yes
Perino *et al.*, 2004[100]	Unclear, not stated	Yes	Yes	Yes	Yes	Yes	Yes
Duleba *et al.*, 2003[98]	Unclear, not stated	Yes	Yes	Yes	No	Yes	Yes
Hawe *et al.*, 2003[94]	Yes	Yes	Yes	Yes	Yes	Yes	Yes
van Zon-Rabelink *et al.*, 2004[95]	Unclear, not stated	Yes	Yes	Yes	Yes	No	Yes
Cooper *et al.*, 2002[56]	Unclear, not stated	Yes	Yes	Yes	Yes	Yes	Yes

continued

Paper	Was randomisation adequate?	Was the target population described adequately?	Was the sample size calculation reported?	Were the two populations comparable at baseline?	Was an ITT analysis reported?	Was the follow-up > 80%?	Was compliance with allocated treatment > 80% in both arms at 12 months?
Pellicano *et al.*, 2002[102]	Unclear, not stated	Yes	No	Yes	No	Yes	Yes
Corson, 2001[79]	Unclear, not stated	Yes	Yes	Yes	No	Yes	Yes
Soysal *et al.*, 2001[96]	Yes	Yes	No	Yes	No	Yes	Yes
Corson *et al.*, 2000[101]	Unclear, not stated	Yes	Yes	Yes	No	Yes	Yes
Cooper *et al.*, 1999[54]	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Meyer *et al.*, 1998[53]	Unclear, not stated	Yes	Yes	Yes	No	Yes	Yes
Romer, 1998[97]	Unclear, not stated	No	No	Yes	No	Yes	No
Mirena vs first -generation EA							
Malak *et al.*, 2006[104]	Yes	Yes	No	Yes	No	Yes	Yes
Kittelsen, 1998[105]	Unclear, not stated	Yes	No	Yes	No	Yes	No
Crosignani *et al.*, 1997[80]	Yes	Yes	No	Yes	No	Yes	Yes
Mirena vs second-generation EA							
Shaw *et al.*, 2007[108]	Yes	Yes	Yes	Yes	No	No	No
Tam *et al.*, 2006[109]	Unclear, not stated	Yes	No	Yes	No	No	No
Busfield *et al.*, 2006[107]	Yes	Yes	Yes	Yes	No	No	No
Barrington *et al.*, 2003[81]	Unclear, not stated	Yes	No	No	No	Yes	Yes
Soysal *et al.*, 2002[106]	Yes	Yes	Yes	Yes	Yes	Yes	Yes

	Time point	Trials (no.)	WMD (95% CI)	OR (95% CI)	*p*-value	Hetero (*p*)/ I^2 (%)
Duration surgery (minutes)	–	6 (850)	32 (30 to 34)	–	< 0.0001	< 0.0001/99
Duration hospital stay (days)	–	7 (1066)	3.0 (2.9 to 3.1)	–	< 0.0001	< 0.0001/99
Surgery pain score (0–10)	–	2 (367)	2.5 (2.2 to 2.9)	–	< 0.0001	0.8/0
Return to work (days)	–	6 (725)	14 (13 to 16)	–	< 0.0001	< 0.0001/98
Return normal activities (days)	–	5 (770)	5.2 (4.7 to 5.7)	–	< 0.0001	< 0.0001/98
Return sexual activity (days)	–	2 (302)	36 (31 to 41)	–	< 0.0001	< 0.0001/99
Proportion dyspareunia	6 months	1 (166)	–	0.71 (0.39 to 1.31)	0.3	–
	12 months	2 (322)	–	0.87 (0.51 to 1.48)	0.6	0.2/47
SF-36 general health (absolute)	12 months	1 (181)	–9.8 (–13.9 to –5.7)	–	< 0.0001	–
SF-36 physical function (absolute)		1 (181)	–1.2 (–5.3 to 2.9)	–	0.6	–
SF-36 role physical (absolute)		1 (181)	–0.8 (–5.0 to 3.4)	–	0.7	–
SF-36 role emotional (absolute)		1 (181)	–3.9 (–8.2 to 0.4)	–	0.08	–
SF-36 mental health (absolute)		1 (181)	–2.7 (–6.8 to 1.4)	–	0.2	–
SF-36 social function (absolute)		1 (181)	–21.2 (–24.7 to –17.8)	–	< 0.0001	–
SF-36 vitality (absolute)		1 (181)	–11.3 (–14.8 to –7.8)	–	< 0.0001	–
SF-36 pain (absolute)		1 (181)	–1.5 (–6.1 to 3.1)	–	0.5	–
SF-36 general health (absolute)	2 years	2 (225)	–6.5 (–12.1 to –0.9)	–	0.02	0.4/0
SF-36 physical function (absolute)		2 (221)	–2.8 (–7.4 to 1.8)	–	0.2	0.8/0
SF-36 role physical (absolute)		2 (223)	–1.3 (–10.4 to 7.9)	–	0.8	0.4/0
SF-36 role emotional (absolute)		2 (224)	–7.6 (–16.2 to 1.1)	–	0.09	0.7/0
SF-36 mental health (absolute)		2 (221)	–2.8 (–7.4 to 1.8)	–	0.2	0.6/0
SF-36 social function (absolute)		2 (221)	–7.1 (–12.5 to –1.8)	–	0.009	0.5/0
SF-36 vitality (absolute)		2 (222)	–5.0 (–10.5 to 0.5)	–	0.07	0.08/67
SF-36 pain (absolute)		2 (225)	–8.4 (–14.9 to –2.0)	–	0.01	0.6/0
SF-36 general health (change)	12 months	1 (181)	–9.6 (–13.5 to –5.7)	–	< 0.0001	–
SF-36 physical function (change)		1 (181)	–1.0 (–5.0 to 3.0)	–	0.6	–
SF-36 role physical (change)		1 (181)	0.1 (–4.1 to 4.3)	–	1.0	–
SF-36 role emotional (change)		1 (181)	–4.4 (–8.4 to –0.4)	–	0.03	–
SF-36 mental health (change)		1 (181)	–1.0 (–4.9 to 2.9)	–	0.6	–
SF-36 social function (change)		1 (181)	–24 (–27 to –21)	–	< 0.0001	–
SF-36 vitality (change)		1 (181)	–13 (–16 to –9)	–	< 0.0001	–
SF-36 pain (change)		1 (181)	–2.2 (–7.3 to 2.9)	–	0.4	–
EQ-5D (absolute)	6 months	1 (220)	–0.09 (–0.16 to –0.02)	–	0.02	–
	12 months	1 (210)	0.00 (–0.08 to 0.08)	–	1.0	–
	2 years	1 (213)	–0.02 (–0.09 to 0.05)	–	0.6	–
	3 years	1 (157)	0.04 (–0.05 to 0.13)	–	0.4	–
	4 years	1 (98)	–0.01 (–0.12 to 0.10)	–	0.9	–
EQ-5D (change)	6 months	1 (220)	–0.09 (–0.18 to –0.00)	–	0.05	–
	12 months	1 (210)	–0.03 (–0.13 to 0.07)	–	0.5	–
	2 years	1 (213)	–0.03 (–0.12 to 0.06)	–	0.5	–
	3 years	1 (157)	0.06 (–0.05 to 0.17)	–	0.3	–
	4 years	1 (96)	–0.01 (–0.16 to 0.14)	–	0.9	–

continued

	Time point	Trials (no.)	WMD (95% CI)	OR (95% CI)	*p*-value	Hetero (*p*)/ I^2 (%)
		Trials	Frequency			
Repeat EA	6 months	3	11/318 (3%)			
	12 months	3	17/248 (7%)			
	2 years	2	13/222 (6%)			
	3 years	2	15/189 (8%)			
	4 years	1	1/48 (2%)			
	5 years	1	2/123 (2%)			
Hysterectomy after EA	6 months	3	11/305 (4%)			
	12 months	4	27/271 (10%)			
	2 years	3	38/246 (15%)			
	3 years	2	33/194 (17%)			
	4 years	1	23/59 (39%)			
	5 years	1	42/123 (34%)			

Appendix 4

Pooled results for hysterectomy versus first-generation endometrial ablation

Appendix 4.1 Hysterectomy versus first-generation endometrial ablation – complications

	Trials	Frequency (hysterectomy: max. 530; first-generation EA: max. 585)	OR (95% CI)[a]	*p*-value	Hetero (*p*)/ I^2 (%)
Periprocedure complications					
Anaesthesia problems (hysterectomy, first-generation EA)	7	3; 0	10.9 (1.08 to 111)	0.04	0.7/0
Excessive bleeding (hysterectomy, first-generation EA)	7	10; 10	1.03 (0.42 to 2.53)	1.0	0.7/0
Injury surrounding organs (hysterectomy)	7	3	–	–	–
Uterine perforation (first-generation EA)	7	11	–	–	–
Fluid overload (first-generation EA)	7	21	–	–	–
Visceral damage (first-generation EA)	7	1	–	–	–
Cervical laceration (first-generation EA)	7	4	–	–	–
Procedure abandoned (first-generation EA)	7	2	–	–	–
Converted to hysterectomy (first-generation EA)	7	14	–	–	–
Further complications (< 1 month)					
Urinary tract infection (hysterectomy, first-generation EA)	7	43; 9	4.38 (2.48 to 7.75)	< 0.0001	0.6/0
Deep-vein thrombosis (hysterectomy, first-generation EA)	7	2; 0	6.96 (0.43 to 112)	0.2	–
Excessive bleeding (hysterectomy)	7	9	–	–	–
Embolism (hysterectomy)	7	2	–	–	–
Further bleeding (first-generation EA)	7	0	–	–	–
Sepsis (first-generation EA)	7	9	–	–	–
Pyrexia (first-generation EA)	7	5	–	–	–
Endometriosis (first-generation EA)	7	1	–	–	–
Abdominal pain (first-generation EA)	7	0	–	–	–
Foul discharge (first-generation EA)	7	0	–	–	–
Visceral damage (first-generation EA)	7	0	–	–	–

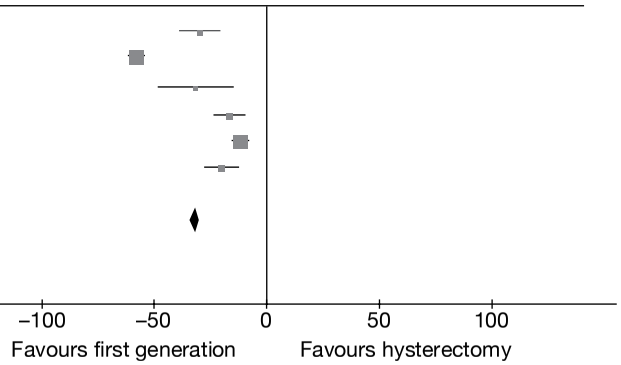
a < 0 favours hysterectomy, > 0 favours first-generation EA.

Appendix 4.2 Hysterectomy versus first-generation endometrial ablation

Duration of surgery (minutes)

	First generation			Hysterectomy				Mean difference	Mean difference
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 99% CI**	**IV, Fixed, 99% CI**
b. Zupi 2003[78]	41.7	19.2	89	71.5	28.1	92	6.3%	−29.80 (−38.99 to −20.61)	
c. Crosignani 1997[45]	13	3.7	41	71	8.9	44	37.6%	−58.00 (−61.76 to −54.24)	
d. O'Connor 1997[44]	34.3	30.1	100	65.9	42	52	1.9%	−31.60 (−48.49 to −14.71)	
e. Pinion 1994[43]	44.8	13.7	102	61.4	21.9	83	10.5%	−16.60 (−23.71 to −9.49)	
f. Dwyer 1993[42]	36.2	9.9	98	47.9	11.2	97	35.0%	−11.70 (−15.60 to −7.80)	
g. Gannon 1991[41]	29.4	8	27	49.5	13	25	8.8%	−20.10 (−27.88 to −12.32)	
Total (95% CI)			**457**			**393**	**100.0%**	**−31.86 (−33.61 to −30.10)**	

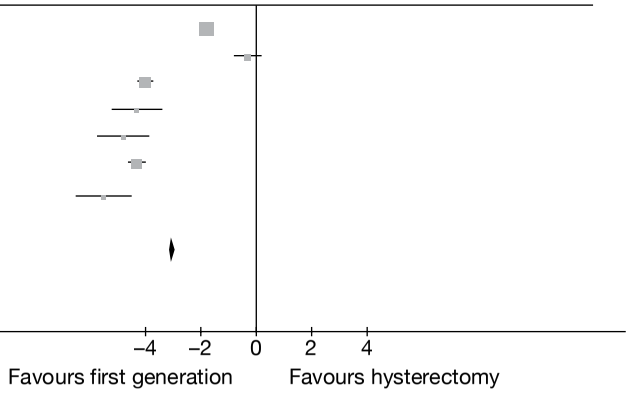
Heterogeneity: $\chi^2 = 543.45$, df = 5 ($p < 0.00001$); $I^2 = 99\%$
Test for overall effect: $z = 35.58$ ($p < 0.00001$)



Duration of hospital stay (days)

	First generation			Hysterectomy				Mean difference	Mean difference
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 99% CI**	**IV, Fixed, 99% CI**
a. Dickersin 2007[92]	0.1	0.3	110	1.9	1	118	36.0%	−1.80 (−2.05 to −1.55)	
b. Zupi 2003[78]	1.3	1.1	89	1.6	1.5	92	8.8%	−0.30 (−0.80 to 0.20)	
c. Crosignani 1997[45]	1	0.1	41	5	0.74	44	26.4%	−4.00 (−4.29 to −3.71)	
d. O'Connor 1997[44]	1.1	0.5	80	5.4	2.4	47	2.7%	−4.30 (−5.21 to −3.39)	
e. Pinion 1994[43]	2.5	2.9	105	7.3	2.3	97	2.5%	−4.80 (−5.75 to −3.85)	
f. Dwyer 1993[42]	2.1	0.7	98	6.4	1	94	21.4%	−4.30 (−4.62 to −3.98)	
g. Gannon 1991[41]	1.6	1.4	25	7.1	1.4	26	2.2%	−5.50 (−6.51 to −4.49)	
Total (95% CI)			**548**			**518**	**100.0%**	**−3.01 (−3.12 to −2.89)**	

Heterogeneity: $\chi^2 = 611.56$, df = 6 ($p < 0.00001$); $I^2 = 99\%$
Test for overall effect: $z = 51.94$ ($p < 0.00001$)



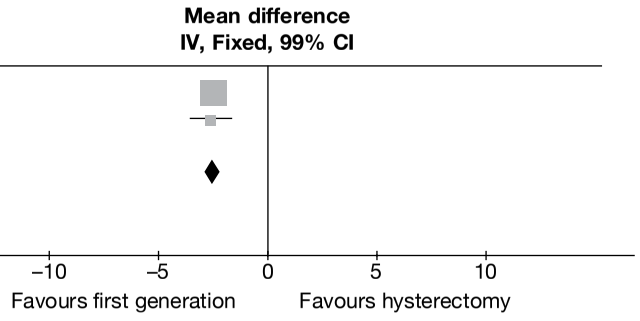
DOI: 10.3310/hta15190

Surgery pain score

Study or subgroup	First generation Mean	SD	Total	Hysterectomy Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
b. Zupi 2003[78]	3.8	0.6	89	6.3	1.9	92	76.7%	–2.50 (–3.04 to –1.96)
e. Pinion 1994[43]	2.6	2.1	89	5.2	3	97	23.3%	–2.60 (–3.57 to –1.63)
Total (95% CI)			**178**			**189**	**100.0%**	**–2.52 (–2.88 to –2.17)**

Heterogeneity: $\chi^2 = 0.05$, df = 1 ($p = 0.82$); $I^2 = 0\%$

Test for overall effect: $z = 13.85$ ($p < 0.00001$)

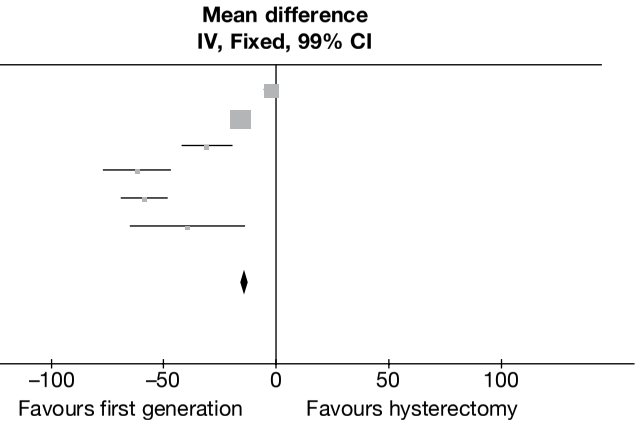


Time to return to work (days)

Study or subgroup	First generation Mean	SD	Total	Hysterectomy Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
b. Zupi 2003[78]	22.4	7.7	89	24.5	9.8	92	36.0%	–2.10 (–5.47 to 1.27)
c. Crosignani 1997[45]	14	3.7	41	30	5.9	44	54.7%	–16.00 (–18.73 to –13.27)
d. O'Connor 1997[44]	20.9	21.4	84	51.6	22.4	36	3.2%	–30.70 (–42.04 to –19.36)
e. Pinion 1994[43]	30.6	28.5	78	92.4	40.1	67	1.8%	–61.80 (–76.91 to –46.69)
f. Dwyer 1993[42]	18.1	15.2	85	76.7	33.2	81	3.8%	–58.60 (–69.01 to –48.19)
g. Gannon 1991[41]	21.4	26.5	13	60.8	25.9	15	0.6%	–39.40 (–65.00 to –13.80)
Total (95% CI)			**390**			**335**	**100.0%**	**–14.04 (–15.57 to –12.50)**

Heterogeneity: $\chi^2 = 295.51$, df = 5 ($p < 0.00001$); $I^2 = 98\%$

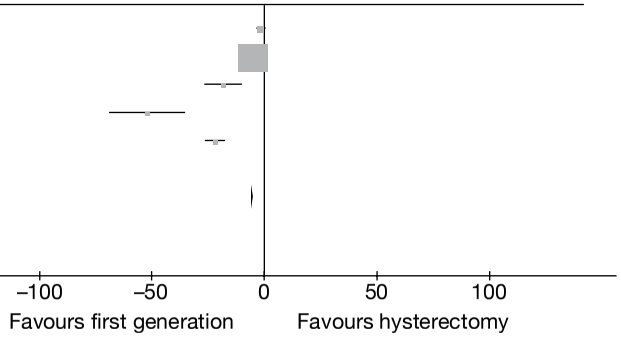
Test for overall effect: $z = 17.90$ ($p < 0.00001$)



Time to return to normal activities (days)

Study or subgroup	First generation Mean	SD	Total	Hysterectomy Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
b. Zupi 2003[78]	8.8	4.3	89	10.3	6.5	92	10.6%	−1.50 (−3.60 to 0.60)
c. Crosignani 1997[45]	8	1.5	41	13	1.1	44	86.2%	−5.00 (−5.74 to −4.26)
d. O'Connor 1997[44]	13.3	19.5	104	31.6	17.5	44	0.7%	−18.30 (−26.69 to −9.91)
e. Pinion 1994[43]	32.9	34.1	88	85	46.8	74	0.2%	−52.10 (−68.95 to −35.25)
f. Dwyer 1993[42]	8.3	8.3	98	30.2	15	96	2.3%	−21.90 (−26.40 to −17.40)
Total (95% CI)			**420**			**350**	**100.0%**	**−5.19 (−5.71 to −4.67)**

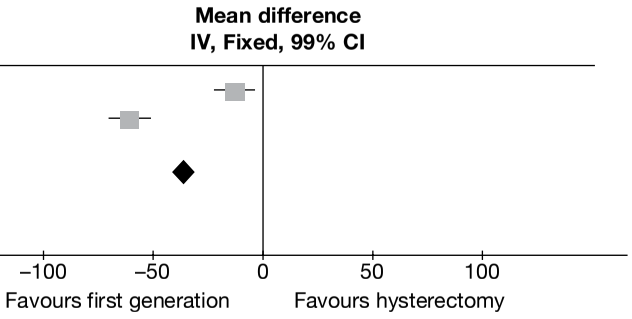
Heterogeneity: $\chi^2 = 180.09$, df = 4 ($p < 0.00001$); $I^2 = 98\%$
Test for overall effect: $z = 19.47$ ($p < 0.00001$)



Time to return to sexual activity (days)

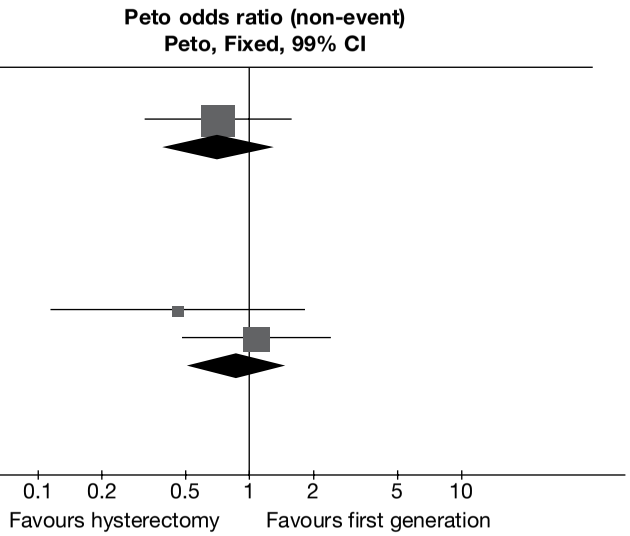
Study or subgroup	First generation Mean	SD	Total	Hysterectomy Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
d. O'Connor 1997[44]	28	23.4	90	40.9	17.1	41	51.7%	−12.90 (−22.26 to −3.54)
f. Dwyer 1993[42]	17.1	10.9	90	77.8	32.2	81	48.3%	−60.70 (−70.38 to −51.02)
Total (95% CI)			**180**			**122**	**100.0%**	**−36.01 (−41.13 to −30.89)**

Heterogeneity: $\chi^2 = 83.58$, df = 1 ($p < 0.00001$); $I^2 = 99\%$
Test for overall effect: $z = 13.78$ ($p < 0.00001$)



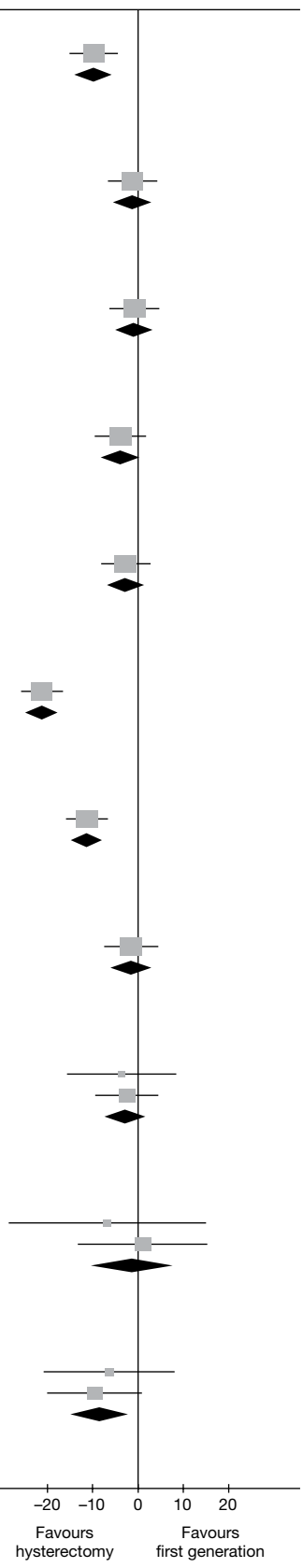
Proportion with dyspareunia

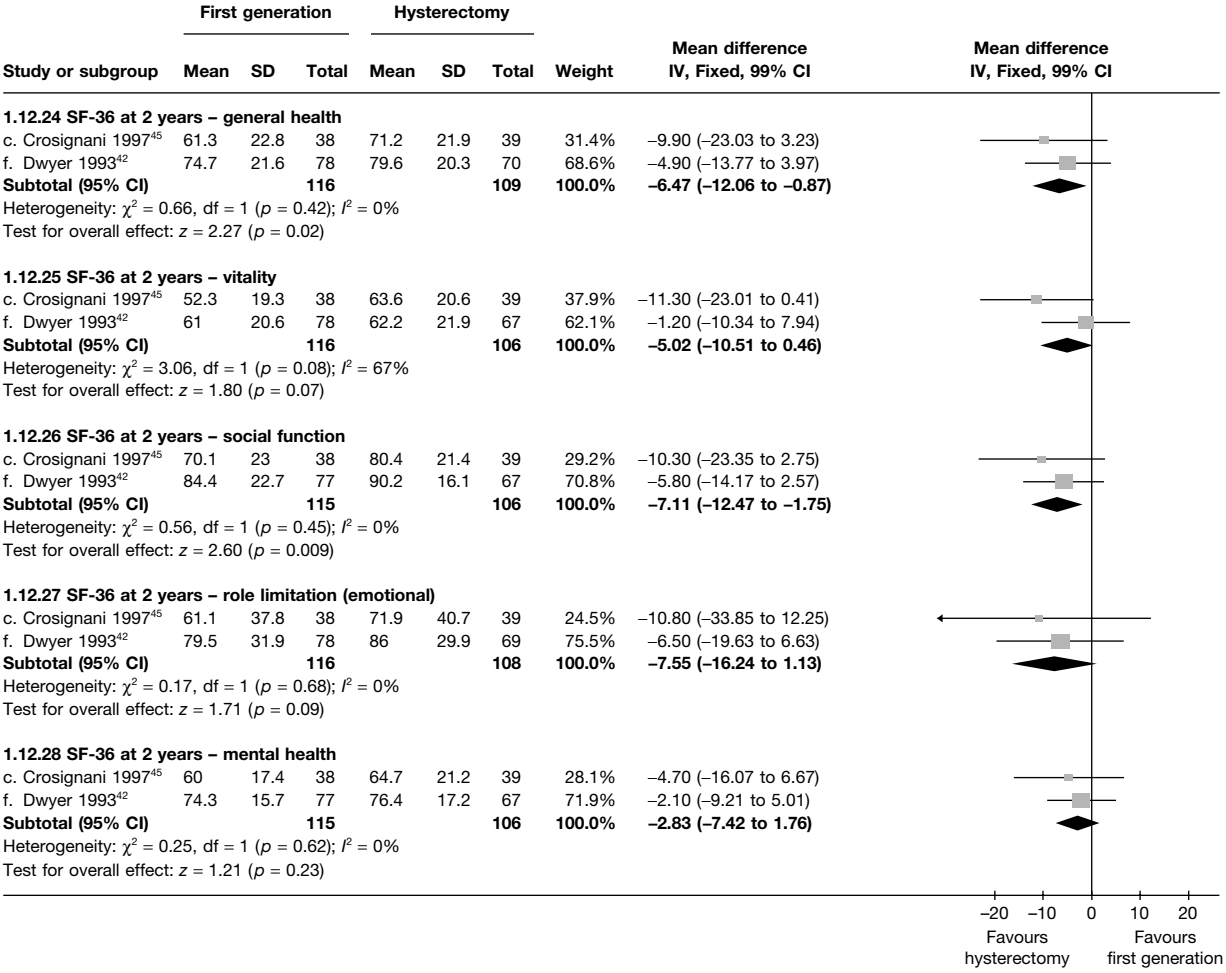
Study or subgroup	First generation Events	First generation Total	Hysterectomy Events	Hysterectomy Total	Weight	Peto odds ratio (non-event) Peto, Fixed, 99% CI
1.11.1 6 months						
e. Pinion 1994[43]	48	89	35	77	100.0%	0.71 (0.32 to 1.59)
Subtotal (95% CI)		**89**		**77**	**100.0%**	**0.71 (0.39 to 1.31)**
Total events	48		35			
Heterogeneity: not applicable						
Test for overall effect: $z = 1.09$ ($p = 0.28$)						
1.11.2 1 year						
d. O'Connor 1997[44]	15	108	3	50	25.6%	0.46 (0.12 to 1.83)
e. Pinion 1994[43]	36	85	35	79	74.4%	1.08 (0.48 to 2.43)
Subtotal (95% CI)		**193**		**129**	**100.0%**	**0.87 (0.51 to 1.48)**
Total events	51		38			
Heterogeneity: $\chi^2 = 1.89$, df = 1 ($p = 0.17$); $I^2 = 47\%$						
Test for overall effect: $z = 0.51$ ($p = 0.61$)						



SF-36 scores (absolute values)

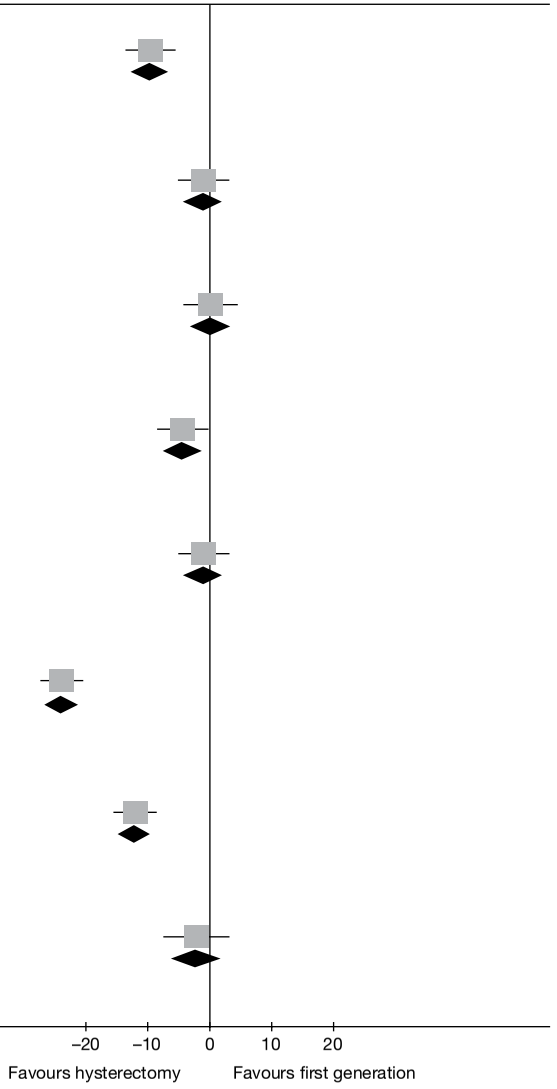
Study or subgroup	First generation Mean	SD	Total	Hysterectomy Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.12.13 SF-36 at 1 year – general health								
b. Zupi 2003[78]	59.6	13.7	89	69.4	14.2	92	100.0%	–9.80 (–15.14 to –4.46)
Subtotal (95% CI)			**89**			**92**	**100.0%**	**–9.80 (–13.86 to –5.74)**
Heterogeneity: not applicable								
Test for overall effect: $z = 4.73$ ($p < 0.00001$)								
1.12.14 SF-36 at 1 year – physical function								
b. Zupi 2003[78]	66.4	15.1	89	67.6	13.2	92	100.0%	–1.20 (–6.64 to 4.24)
Subtotal (95% CI)			**89**			**92**	**100.0%**	**–1.20 (–5.34 to 2.94)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.57$ ($p = 0.57$)								
1.12.15 SF-36 at 1 year – role limitation (physical)								
b. Zupi 2003[78]	61.3	14.8	89	62.1	13.9	92	100.0%	–0.80 (–6.30 to 4.70)
Subtotal (95% CI)			**89**			**92**	**100.0%**	**–0.80 (–4.99 to 3.39)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.37$ ($p = 0.71$)								
1.12.16 SF-36 at 1 year – role limitation (emotional)								
b. Zupi 2003[78]	64.2	14.4	89	68.1	15.2	92	100.0%	–3.90 (–9.57 to 1.77)
Subtotal (95% CI)			**89**			**92**	**100.0%**	**–3.90 (–8.21 to 0.41)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.77$ ($p = 0.08$)								
1.12.17 SF-36 at 1 year – mental health								
b. Zupi 2003[78]	60.5	14.8	89	63.2	13.6	92	100.0%	–2.70 (–8.15 to 2.75)
Subtotal (95% CI)			**89**			**92**	**100.0%**	**–2.70 (–6.84 to 1.44)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.28$ ($p = 0.20$)								
1.12.18 SF-36 at 1 year – social function								
b. Zupi 2003[78]	67.3	12.7	89	88.5	11.5	92	100.0%	–21.20 (–25.84 to –16.56)
Subtotal (95% CI)			**89**			**92**	**100.0%**	**–21.20 (–24.73 to –17.67)**
Heterogeneity: not applicable								
Test for overall effect: $z = 11.76$ ($p < 0.00001$)								
1.12.19 SF-36 at 1 year – vitality								
b. Zupi 2003[78]	61	12.8	89	72.3	11.3	92	100.0%	–11.30 (–15.93 to –6.67)
Subtotal (95% CI)			**89**			**92**	**100.0%**	**–11.30 (–14.82 to –7.78)**
Heterogeneity: not applicable								
Test for overall effect: $z = 6.29$ ($p < 0.00001$)								
1.12.20 SF-36 at 1 year – pain								
b. Zupi 2003[78]	58.6	17	89	60.1	14	92	100.0%	–1.50 (–7.47 to 4.47)
Subtotal (95% CI)			**89**			**92**	**100.0%**	**–1.50 (–6.05 to 3.05)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.65$ ($p = 0.52$)								
1.12.21 SF-36 at 2 years – physical function								
c. Crosignani 1997[45]	84.4	21	38	88	20.1	39	24.8%	–3.60 (–15.67 to 8.47)
f. Dwyer 1993[42]	89.4	17.9	77	91.9	14.4	67	75.2%	–2.50 (–9.44 to 4.44)
Subtotal (95% CI)			**115**			**106**	**100.0%**	**–2.77 (–7.35 to 1.80)**
Heterogeneity: $\chi^2 = 0.04$, df = 1 ($p = 0.84$); $I^2 = 0\%$								
Test for overall effect: $z = 1.19$ ($p = 0.24$)								
1.12.22 SF-36 at 2 years – role limitation (physical)								
c. Crosignani 1997[45]	67.4	36.3	38	74.1	37.9	39	30.1%	–6.70 (–28.48 to 15.08)
f. Dwyer 1993[42]	82.9	33.2	76	81.8	33.8	70	69.9%	1.10 (–13.20 to 15.40)
Subtotal (95% CI)			**114**			**109**	**100.0%**	**–1.25 (–10.35 to 7.85)**
Heterogeneity: $\chi^2 = 0.59$, df = 1 ($p = 0.44$); $I^2 = 0\%$								
Test for overall effect: $z = 0.27$ ($p = 0.79$)								
1.12.23 SF-36 at 2 years – pain								
c. Crosignani 1997[45]	69.6	27	38	75.9	21.9	39	34.4%	–6.30 (–20.75 to 8.15)
f. Dwyer 1993[42]	73.5	26.3	80	83	23.1	68	65.6%	–9.50 (–19.96 to 0.96)
Subtotal (95% CI)			**118**			**107**	**100.0%**	**–8.40 (–14.85 to –1.95)**
Heterogeneity: $\chi^2 = 0.21$, df = 1 ($p = 0.64$); $I^2 = 0\%$								
Test for overall effect: $z = 2.55$ ($p = 0.01$)								





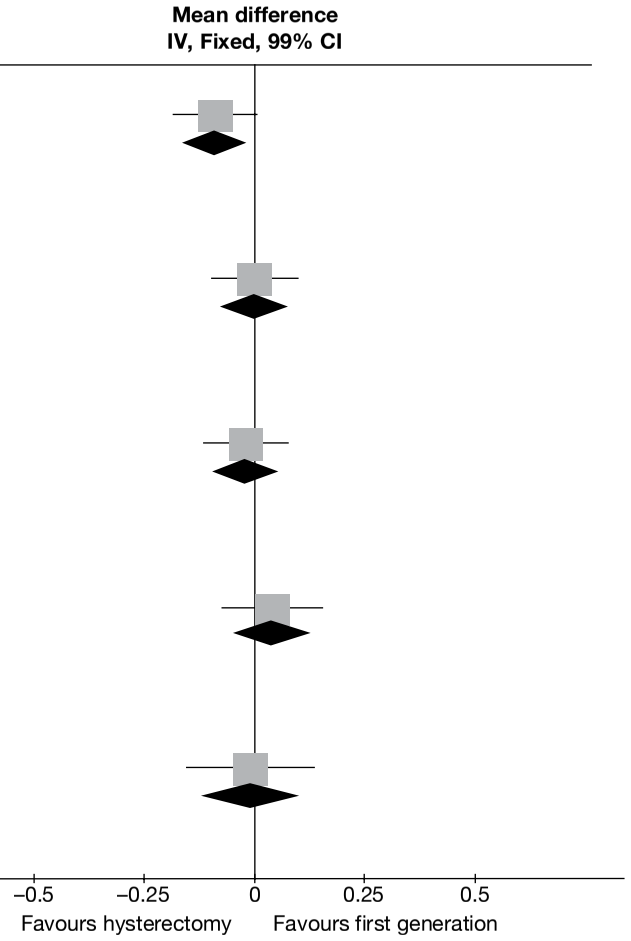
SF-36 scores (change from baseline)

Study or subgroup	First generation Mean	SD	Total	Hysterectomy Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.13.13 SF-36 at 1 year – general health								
b. Zupi 2003[78]	7.7	13.2	89	17.3	13.3	92	100.0%	–9.60 (–14.67 to –4.53)
Subtotal (95% CI)			**89**			**92**	**100.0%**	**–9.60 (–13.46 to –5.74)**
Heterogeneity: not applicable								
Test for overall effect: $z = 4.87$ ($p < 0.00001$)								
1.13.14 SF-36 at 1 year – physical function								
b. Zupi 2003[78]	3.8	14.8	89	4.8	12.2	92	100.0%	–1.00 (–6.20 to 4.20)
Subtotal (95% CI)			**89**			**92**	**100.0%**	**–1.00 (–4.96 to 2.96)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.50$ ($p = 0.62$)								
1.13.15 SF-36 at 1 year – role limitation (physical)								
b. Zupi 2003[78]	3	14	89	2.9	14.7	92	100.0%	0.10 (–5.40 to 5.60)
Subtotal (95% CI)			**89**			**92**	**100.0%**	**0.10 (–4.08 to 4.28)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.05$ ($p = 0.96$)								
1.13.16 SF-36 at 1 year – role limitation (emotional)								
b. Zupi 2003[78]	3.4	13.4	89	7.8	13.8	92	100.0%	–4.40 (–9.61 to 0.81)
Subtotal (95% CI)			**89**			**92**	**100.0%**	**–4.40 (–8.36 to –0.44)**
Heterogeneity: not applicable								
Test for overall effect: $z = 2.18$ ($p = 0.03$)								
1.13.17 SF-36 at 1 year – mental health								
b. Zupi 2003[78]	2.4	13.7	89	3.4	13.3	92	100.0%	–1.00 (–6.17 to 4.17)
Subtotal (95% CI)			**89**			**92**	**100.0%**	**–1.00 (–4.94 to 2.94)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.50$ ($p = 0.62$)								
1.13.18 SF-36 at 1 year – social function								
b. Zupi 2003[78]	10.9	11.9	89	34.9	10.7	92	100.0%	–24.00 (–28.34 to –19.66)
Subtotal (95% CI)			**89**			**92**	**100.0%**	**–24.00 (–27.30 to –20.70)**
Heterogeneity: not applicable								
Test for overall effect: $z = 14.25$ ($p < 0.00001$)								
1.13.19 SF-36 at 1 year – vitality								
b. Zupi 2003[78]	4.3	12	89	16.9	10.8	92	100.0%	–12.60 (–16.98 to –8.22)
Subtotal (95% CI)			**89**			**92**	**100.0%**	**–12.60 (–15.93 to –9.27)**
Heterogeneity: not applicable								
Test for overall effect: $z = 7.42$ ($p < 0.00001$)								
1.13.20 SF-36 at 1 year – pain								
b. Zupi 2003[78]	1.5	18.2	89	3.7	16.7	92	100.0%	–2.20 (–8.89 to 4.49)
Subtotal (95% CI)			**89**			**92**	**100.0%**	**–2.20 (–7.29 to 2.89)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.85$ ($p = 0.40$)								



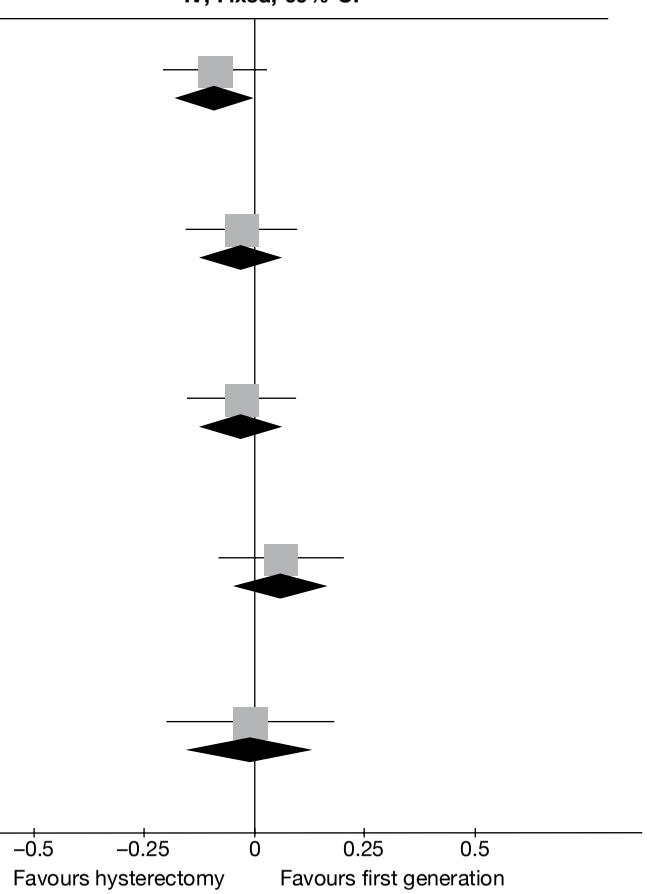
EQ-5D scores (absolute values)

Study or subgroup	First generation Mean	SD	Total	Hysterectomy Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.14.2 6 months								
a. Dickersin 2007[92]	0.69	0.31	111	0.78	0.24	109	100.0%	–0.09 (–0.19 to 0.01)
Subtotal (95% CI)			**111**			**109**	**100.0%**	**–0.09 (–0.16 to –0.02)**
Heterogeneity: not applicable								
Test for overall effect: $z = 2.41$ ($p = 0.02$)								
1.14.3 12 months								
a. Dickersin 2007[92]	0.74	0.29	107	0.74	0.27	103	100.0%	0.00 (–0.10 to 0.10)
Subtotal (95% CI)			**107**			**103**	**100.0%**	**0.00 (–0.08 to 0.08)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.00$ ($p = 1.00$)								
1.14.4 2 years								
a. Dickersin 2007[92]	0.74	0.27	106	0.76	0.28	107	100.0%	–0.02 (–0.12 to 0.08)
Subtotal (95% CI)			**106**			**107**	**100.0%**	**–0.02 (–0.09 to 0.05)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.53$ ($p = 0.60$)								
1.14.5 3 years								
a. Dickersin 2007[92]	0.76	0.25	75	0.72	0.31	82	100.0%	0.04 (–0.08 to 0.16)
Subtotal (95% CI)			**75**			**82**	**100.0%**	**0.04 (–0.05 to 0.13)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.89$ ($p = 0.37$)								
1.14.6 4 years								
a. Dickersin 2007[92]	0.74	0.29	47	0.75	0.27	51	100.0%	–0.01 (–0.16 to 0.14)
Subtotal (95% CI)			**47**			**51**	**100.0%**	**–0.01 (–0.12 to 0.10)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.18$ ($p = 0.86$)								



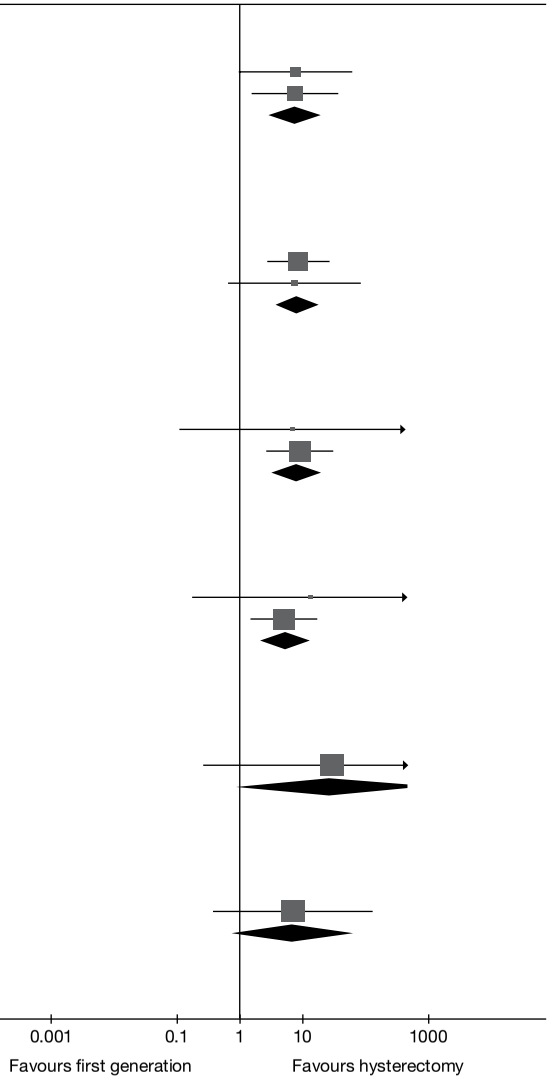
EQ-5D scores (change from baseline)

Study or subgroup	First generation Mean	SD	Total	Hysterectomy Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.15.1 6 months								
a. Dickersin 2007[92]	0.14	0.34	111	0.23	0.34	109	100.0%	–0.09 (–0.21 to 0.03)
Subtotal (95% CI)			**111**			**109**	**100.0%**	**–0.09 (–0.18 to –0.00)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.96$ ($p = 0.05$)								
1.15.2 12 months								
a. Dickersin 2007[92]	0.16	0.35	107	0.19	0.36	103	100.0%	–0.03 (–0.16 to 0.10)
Subtotal (95% CI)			**107**			**103**	**100.0%**	**–0.03 (–0.13 to 0.07)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.61$ ($p = 0.54$)								
1.15.3 2 years								
a. Dickersin 2007[92]	0.18	0.33	106	0.21	0.37	107	100.0%	–0.03 (–0.15 to 0.09)
Subtotal (95% CI)			**106**			**107**	**100.0%**	**–0.03 (–0.12 to 0.06)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.62$ ($p = 0.53$)								
1.15.4 3 years								
a. Dickersin 2007[92]	0.22	0.36	75	0.16	0.33	82	100.0%	0.06 (–0.08 to 0.20)
Subtotal (95% CI)			**75**			**82**	**100.0%**	**0.06 (–0.05 to 0.17)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.09$ ($p = 0.28$)								
1.15.5 4 years								
a. Dickersin 2007[92]	0.18	0.38	47	0.19	0.35	51	100.0%	–0.01 (–0.20 to 0.18)
Subtotal (95% CI)			**47**			**51**	**100.0%**	**–0.01 (–0.16 to 0.14)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.14$ ($p = 0.89$)								



Proportion requiring repeat endometrial ablation

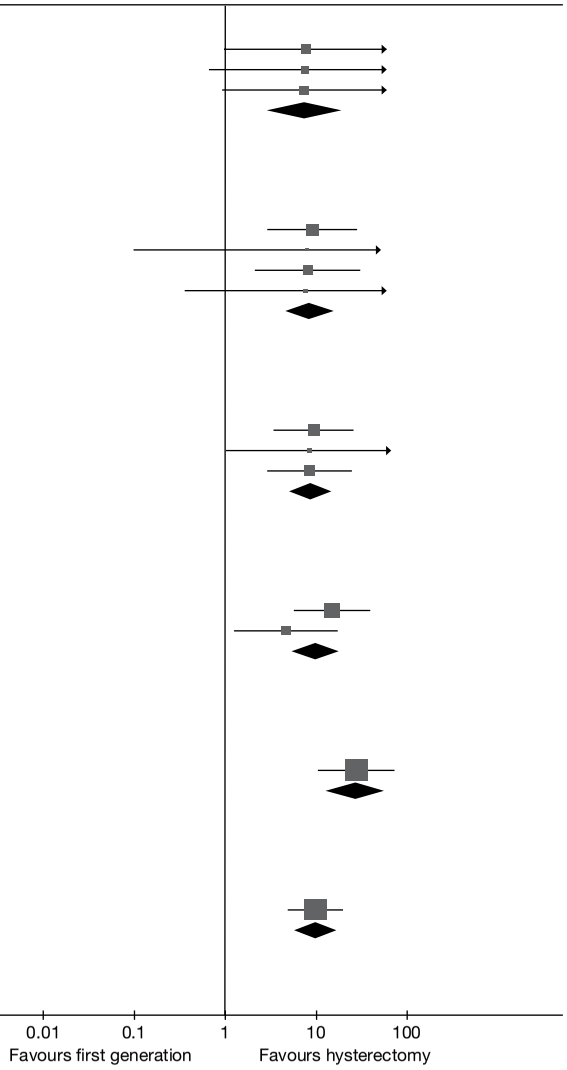
Study or subgroup	First generation Events	First generation Total	Hysterectomy Events	Hysterectomy Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.16.1 6 months						
a. Dickersin 2007[92]	0	123	0	114		Not estimable
e. Pinion 1994[43]	4	96	0	97	36.7%	7.71 (0.57 to 103.39)
f. Dwyer 1993[42]	7	99	0	95	63.3%	7.56 (1.04 to 54.65)
Subtotal (95% CI)		**318**		**306**	**100.0%**	**7.61 (2.30 to 25.21)**
Total events	11		0			
Heterogeneity: $\chi^2 = 0.00$, df = 1 ($p = 0.99$); $I^2 = 0\%$						
Test for overall effect: $z = 3.32$ ($p = 0.0009$)						
1.16.2 1 year						
a. Dickersin 2007[92]	0	123	0	114		Not estimable
e. Pinion 1994[43]	14	97	0	97	81.9%	8.54 (2.05 to 35.52)
g. Gannon 1991[41]	3	28	0	26	18.1%	7.42 (0.36 to 154.22)
Subtotal (95% CI)		**248**		**237**	**100.0%**	**8.32 (3.12 to 22.22)**
Total events	17		0			
Heterogeneity: $\chi^2 = 0.01$, df = 1 ($p = 0.91$); $I^2 = 0\%$						
Test for overall effect: $z = 4.23$ ($p < 0.0001$)						
1.16.3 2 years						
a. Dickersin 2007[92]	1	123	0	114	8.1%	6.87 (0.04 to 1190.59)
f. Dwyer 1993[42]	12	99	0	97	91.9%	8.15 (1.76 to 37.68)
Subtotal (95% CI)		**222**		**211**	**100.0%**	**8.04 (2.63 to 24.56)**
Total events	13		0			
Heterogeneity: $\chi^2 = 0.01$, df = 1 ($p = 0.93$); $I^2 = 0\%$						
Test for overall effect: $z = 3.66$ ($p = 0.0003$)						
1.16.4 3 years						
a. Dickersin 2007[92]	1	73	0	114	7.7%	12.96 (0.07 to 2544.72)
d. O'Connor 1997[44]	14	116	0	56	92.3%	4.98 (1.08 to 22.95)
Subtotal (95% CI)		**189**		**170**	**100.0%**	**5.36 (1.75 to 16.38)**
Total events	15		0			
Heterogeneity: $\chi^2 = 0.20$, df = 1 ($p = 0.65$); $I^2 = 0\%$						
Test for overall effect: $z = 2.95$ ($p = 0.003$)						
1.16.5 4 years						
a. Dickersin 2007[92]	1	48	0	114	100.0%	29.22 (0.10 to 8234.09)
Subtotal (95% CI)		**48**		**114**	**100.0%**	**29.22 (0.40 to 2137.30)**
Total events	1		0			
Heterogeneity: not applicable						
Test for overall effect: $z = 1.54$ ($p = 0.12$)						
1.16.6 5 years						
a. Dickersin 2007[92]	2	123	0	114	100.0%	6.92 (0.18 to 267.25)
Subtotal (95% CI)		**123**		**114**	**100.0%**	**6.92 (0.43 to 111.58)**
Total events	2		0			
Heterogeneity: not applicable						
Test for overall effect: $z = 1.36$ ($p = 0.17$)						



DOI: 10.3310/hta15190

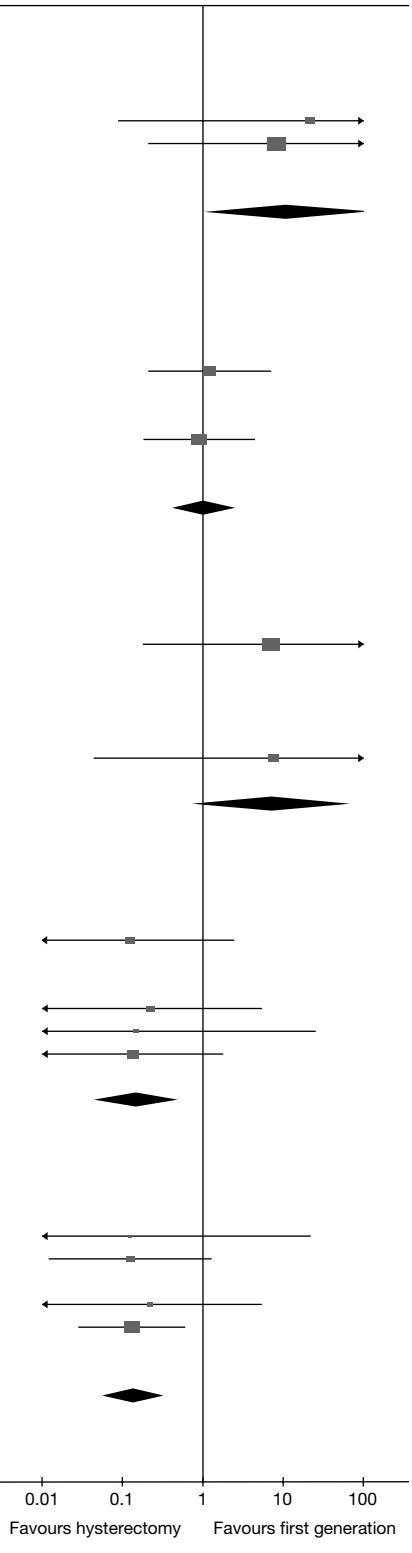
Proportion requiring hysterectomy after endometrial ablation

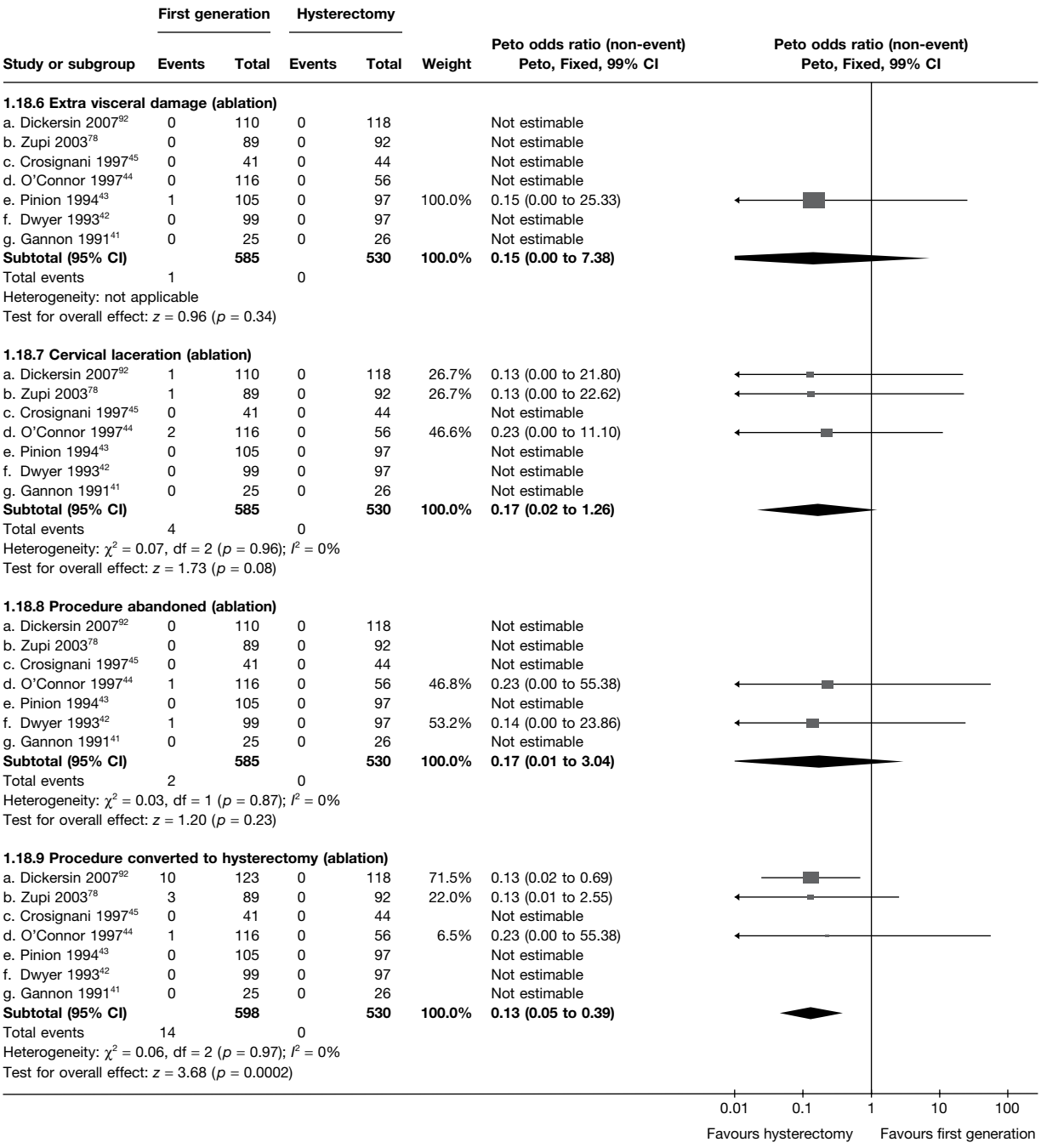
Study or subgroup	First generation Events	First generation Total	Hysterectomy Events	Hysterectomy Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.17.2 6 months						
a. Dickersin 2007[92]	4	110	0	114	36.4%	7.88 (0.59 to 105.41)
e. Pinion 1994[43]	3	96	0	97	27.4%	7.63 (0.38 to 151.65)
f. Dwyer 1993[42]	4	99	0	97	36.3%	7.47 (0.56 to 100.16)
Subtotal (95% CI)		**305**		**308**	**100.0%**	**7.66 (2.33 to 25.17)**
Total events	11		0			
Heterogeneity: $\chi^2 = 0.00$, df = 2 ($p = 1.00$); $I^2 = 0\%$						
Test for overall effect: $z = 3.35$ ($p = 0.0008$)						
1.17.3 1 year						
a. Dickersin 2007[92]	14	105	0	114	51.3%	9.19 (2.22 to 38.05)
c. Crosignani 1997[45]	1	41	0	44	3.9%	7.95 (0.05 to 1377.52)
e. Pinion 1994[43]	10	97	0	97	37.2%	8.15 (1.54 to 43.22)
g. Gannon 1991[41]	2	28	0	28	7.7%	7.67 (0.19 to 302.89)
Subtotal (95% CI)		**271**		**283**	**100.0%**	**8.62 (3.97 to 18.69)**
Total events	27		0			
Heterogeneity: $\chi^2 = 0.03$, df = 3 ($p = 1.00$); $I^2 = 0\%$						
Test for overall effect: $z = 5.45$ ($p < 0.00001$)						
1.17.4 2 years						
a. Dickersin 2007[92]	18	106	0	114	47.1%	9.49 (2.68 to 33.63)
c. Crosignani 1997[45]	4	41	0	44	10.9%	8.58 (0.62 to 118.47)
f. Dwyer 1993[42]	16	99	0	97	42.0%	8.54 (2.23 to 32.63)
Subtotal (95% CI)		**246**		**255**	**100.0%**	**8.98 (4.64 to 17.39)**
Total events	38		0			
Heterogeneity: $\chi^2 = 0.02$, df = 2 ($p = 0.99$); $I^2 = 0\%$						
Test for overall effect: $z = 6.51$ ($p < 0.00001$)						
1.17.5 3 years						
a. Dickersin 2007[92]	21	78	0	114	64.8%	15.63 (4.66 to 52.40)
d. O'Connor 1997[44]	12	116	0	56	35.2%	4.88 (0.95 to 25.16)
Subtotal (95% CI)		**194**		**170**	**100.0%**	**10.37 (4.95 to 21.76)**
Total events	33		0			
Heterogeneity: $\chi^2 = 2.17$, df = 1 ($p = 0.14$); $I^2 = 54\%$						
Test for overall effect: $z = 6.19$ ($p < 0.00001$)						
1.17.6 4 years						
a. Dickersin 2007[92]	23	59	0	114	100.0%	28.85 (8.58 to 97.07)
Subtotal (95% CI)		**59**		**114**	**100.0%**	**28.85 (11.46 to 72.63)**
Total events	23		0			
Heterogeneity: not applicable						
Test for overall effect: $z = 7.14$ ($p < 0.00001$)						
1.17.7 5 years						
a. Dickersin 2007[92]	42	123	0	114	100.0%	10.30 (4.29 to 24.71)
Subtotal (95% CI)		**123**		**114**	**100.0%**	**10.30 (5.29 to 20.04)**
Total events	42		0			
Heterogeneity: not applicable						
Test for overall effect: $z = 6.86$ ($p < 0.00001$)						



Number of patients with adverse events – periprocedure

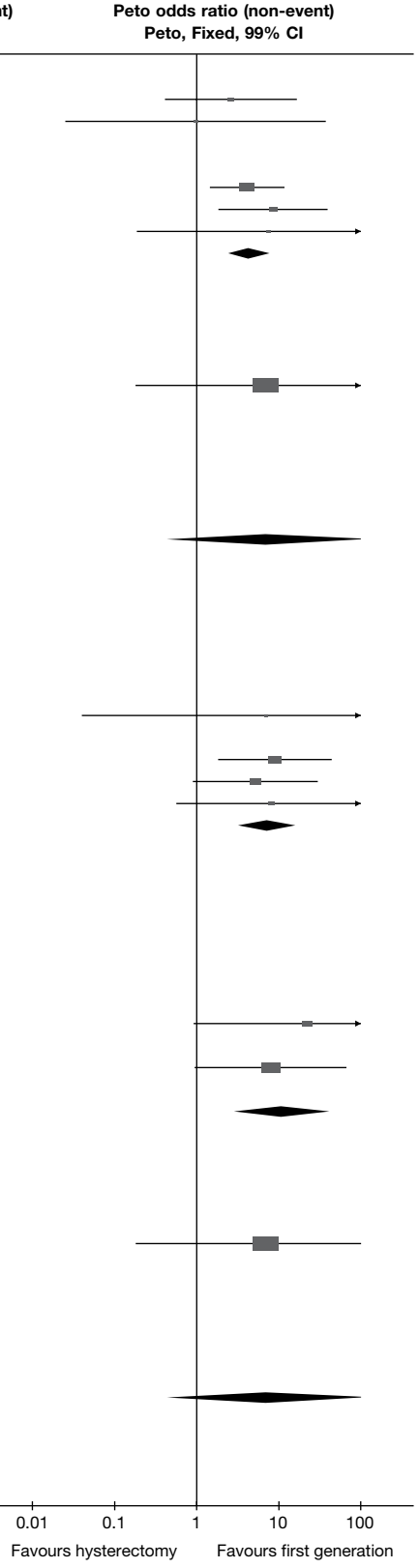
Study or subgroup	First generation Events	First generation Total	Hysterectomy Events	Hysterectomy Total	Weight	Peto odds ratio (non-event) Peto, Fixed, 99% CI
1.18.1 Anaesthesia problems (both)						
a. Dickersin 2007[92]	0	110	0	118		Not estimable
b. Zupi 2003[78]	0	89	0	92		Not estimable
c. Crosignani 1997[45]	0	41	0	44		Not estimable
d. O'Connor 1997[44]	0	116	1	56	30.7%	21.57 (0.09 to 5262.64)
e. Pinion 1994[43]	0	105	2	97	69.3%	8.11 (0.21 to 313.44)
f. Dwyer 1993[42]	0	99	0	97		Not estimable
g. Gannon 1991[41]	0	25	0	26		Not estimable
Subtotal (95% CI)		**585**		**530**	**100.0%**	**10.94 (1.08 to 110.90)**
Total events	0		3			
Heterogeneity: $\chi^2 = 0.15$, df = 1 ($p = 0.70$); $I^2 = 0\%$						
Test for overall effect: $z = 2.03$ ($p = 0.04$)						
1.18.2 Excessive bleeding (both)						
a. Dickersin 2007[92]	0	110	0	118		Not estimable
b. Zupi 2003[78]	4	89	5	92	45.2%	1.22 (0.21 to 7.06)
c. Crosignani 1997[45]	0	41	0	44		Not estimable
d. O'Connor 1997[44]	0	116	0	56		Not estimable
e. Pinion 1994[43]	6	105	5	97	54.8%	0.90 (0.18 to 4.42)
f. Dwyer 1993[42]	0	99	0	97		Not estimable
g. Gannon 1991[41]	0	25	0	26		Not estimable
Subtotal (95% CI)		**585**		**530**	**100.0%**	**1.03 (0.42 to 2.53)**
Total events	10		10			
Heterogeneity: $\chi^2 = 0.11$, df = 1 ($p = 0.74$); $I^2 = 0\%$						
Test for overall effect: $z = 0.07$ ($p = 0.95$)						
1.18.3 Injury to surrounding organs (hysterectomy)						
a. Dickersin 2007[92]	0	110	2	118	66.5%	6.96 (0.18 to 268.76)
b. Zupi 2003[78]	0	89	0	92		Not estimable
c. Crosignani 1997[45]	0	41	0	44		Not estimable
d. O'Connor 1997[44]	0	116	0	56		Not estimable
e. Pinion 1994[43]	0	105	0	97		Not estimable
f. Dwyer 1993[42]	0	99	1	97	33.5%	7.54 (0.04 to 1303.16)
g. Gannon 1991[41]	0	25	0	26		Not estimable
Subtotal (95% CI)		**585**		**530**	**100.0%**	**7.15 (0.74 to 69.06)**
Total events	0		3			
Heterogeneity: $\chi^2 = 0.00$, df = 1 ($p = 0.97$); $I^2 = 0\%$						
Test for overall effect: $z = 1.70$ ($p = 0.09$)						
1.18.4 Uterine perforation (ablation)						
a. Dickersin 2007[92]	3	110	0	118	28.3%	0.12 (0.01 to 2.46)
b. Zupi 2003[78]	0	89	0	92		Not estimable
c. Crosignani 1997[45]	0	41	0	44		Not estimable
d. O'Connor 1997[44]	3	116	0	56	24.8%	0.22 (0.01 to 5.43)
e. Pinion 1994[43]	1	105	0	97	9.5%	0.15 (0.00 to 25.33)
f. Dwyer 1993[42]	4	99	0	97	37.5%	0.13 (0.01 to 1.80)
g. Gannon 1991[41]	0	25	0	26		Not estimable
Subtotal (95% CI)		**585**		**530**	**100.0%**	**0.15 (0.04 to 0.50)**
Total events	11		0			
Heterogeneity: $\chi^2 = 0.14$, df = 3 ($p = 0.99$); $I^2 = 0\%$						
Test for overall effect: $z = 3.08$ ($p = 0.002$)						
1.18.5 Fluid overload (ablation)						
a. Dickersin 2007[92]	1	110	0	118	5.0%	0.13 (0.00 to 21.80)
b. Zupi 2003[78]	5	89	0	92	24.7%	0.12 (0.01 to 1.28)
c. Crosignani 1997[45]	0	41	0	44		Not estimable
d. O'Connor 1997[44]	3	116	0	56	13.1%	0.22 (0.01 to 5.43)
e. Pinion 1994[43]	12	105	0	97	57.2%	0.13 (0.03 to 0.60)
f. Dwyer 1993[42]	0	99	0	97		Not estimable
g. Gannon 1991[41]	0	25	0	26		Not estimable
Subtotal (95% CI)		**585**		**530**	**100.0%**	**0.14 (0.06 to 0.33)**
Total events	21		0			
Heterogeneity: $\chi^2 = 0.17$, df = 3 ($p = 0.98$); $I^2 = 0\%$						
Test for overall effect: $z = 4.40$ ($p < 0.0001$)						

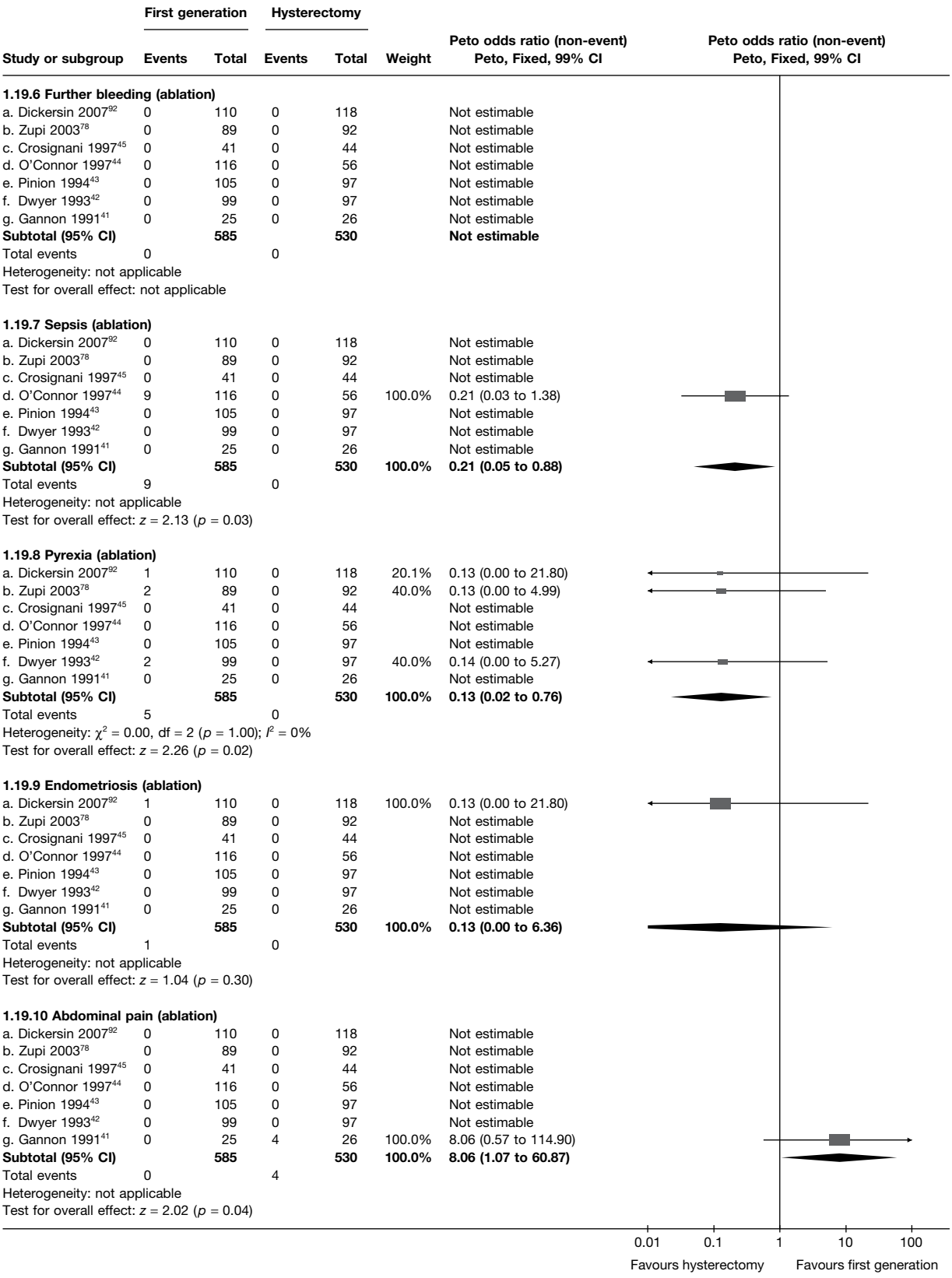


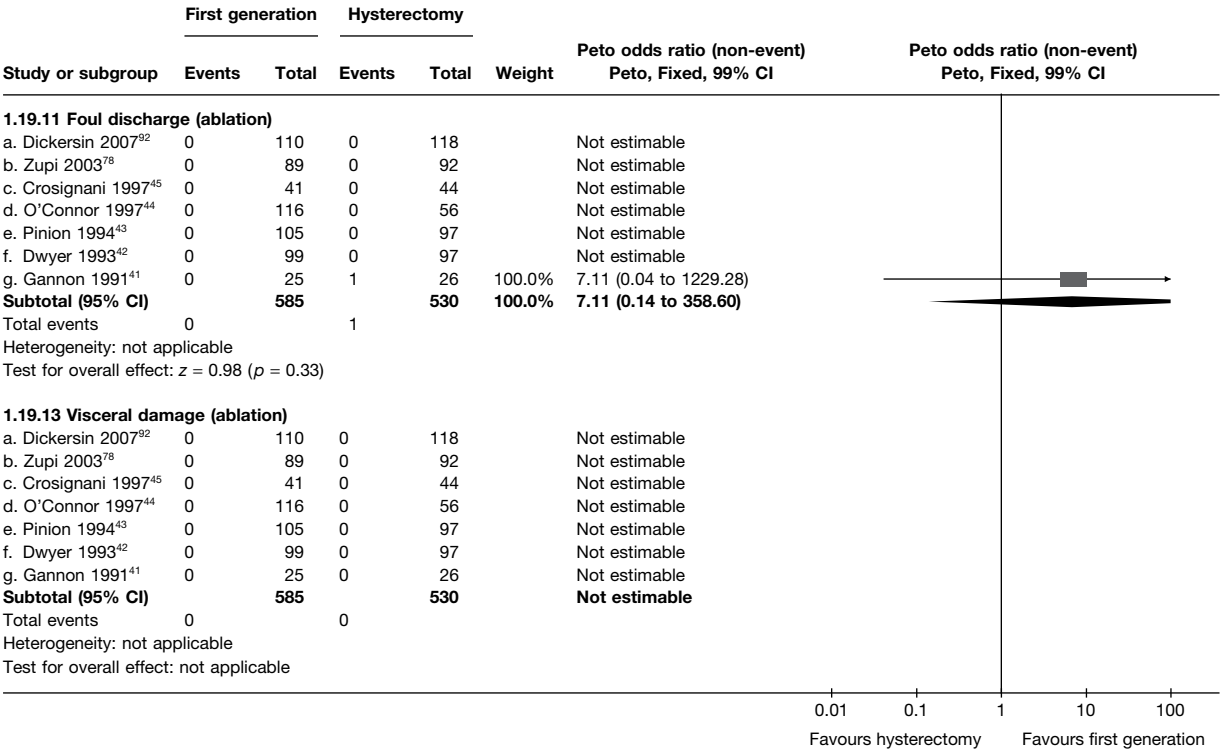

First generation
Hysterectomy
Study or subgroup
Events
Total
Events
Total
Weight
Peto odds ratio (non-event) Peto, Fixed, 99% CI
Peto odds ratio (non-event) Peto, Fixed, 99% CI
1.18.6 Extra visceral damage (ablation)
a. Dickersin 2007[92] 0 110 0 118 Not estimable
b. Zupi 2003[78] 0 89 0 92 Not estimable
c. Crosignani 1997[45] 0 41 0 44 Not estimable
d. O'Connor 1997[44] 0 116 0 56 Not estimable
e. Pinion 1994[43] 1 105 0 97 100.0% 0.15 (0.00 to 25.33)
f. Dwyer 1993[42] 0 99 0 97 Not estimable
g. Gannon 1991[41] 0 25 0 26 Not estimable
Subtotal (95% CI) 585 530 100.0% 0.15 (0.00 to 7.38)
Total events 1 0
Heterogeneity: not applicable
Test for overall effect: z = 0.96 (p = 0.34)
1.18.7 Cervical laceration (ablation)
a. Dickersin 2007[92] 1 110 0 118 26.7% 0.13 (0.00 to 21.80)
b. Zupi 2003[78] 1 89 0 92 26.7% 0.13 (0.00 to 22.62)
c. Crosignani 1997[45] 0 41 0 44 Not estimable
d. O'Connor 1997[44] 2 116 0 56 46.6% 0.23 (0.00 to 11.10)
e. Pinion 1994[43] 0 105 0 97 Not estimable
f. Dwyer 1993[42] 0 99 0 97 Not estimable
g. Gannon 1991[41] 0 25 0 26 Not estimable
Subtotal (95% CI) 585 530 100.0% 0.17 (0.02 to 1.26)
Total events 4 0
Heterogeneity: χ² = 0.07, df = 2 (p = 0.96); I² = 0%
Test for overall effect: z = 1.73 (p = 0.08)
1.18.8 Procedure abandoned (ablation)
a. Dickersin 2007[92] 0 110 0 118 Not estimable
b. Zupi 2003[78] 0 89 0 92 Not estimable
c. Crosignani 1997[45] 0 41 0 44 Not estimable
d. O'Connor 1997[44] 1 116 0 56 46.8% 0.23 (0.00 to 55.38)
e. Pinion 1994[43] 0 105 0 97 Not estimable
f. Dwyer 1993[42] 1 99 0 97 53.2% 0.14 (0.00 to 23.86)
g. Gannon 1991[41] 0 25 0 26 Not estimable
Subtotal (95% CI) 585 530 100.0% 0.17 (0.01 to 3.04)
Total events 2 0
Heterogeneity: χ² = 0.03, df = 1 (p = 0.87); I² = 0%
Test for overall effect: z = 1.20 (p = 0.23)
1.18.9 Procedure converted to hysterectomy (ablation)
a. Dickersin 2007[92] 10 123 0 118 71.5% 0.13 (0.02 to 0.69)
b. Zupi 2003[78] 3 89 0 92 22.0% 0.13 (0.01 to 2.55)
c. Crosignani 1997[45] 0 41 0 44 Not estimable
d. O'Connor 1997[44] 1 116 0 56 6.5% 0.23 (0.00 to 55.38)
e. Pinion 1994[43] 0 105 0 97 Not estimable
f. Dwyer 1993[42] 0 99 0 97 Not estimable
g. Gannon 1991[41] 0 25 0 26 Not estimable
Subtotal (95% CI) 598 530 100.0% 0.13 (0.05 to 0.39)
Total events 14 0
Heterogeneity: χ² = 0.06, df = 2 (p = 0.97); I² = 0%
Test for overall effect: z = 3.68 (p = 0.0002)
0.01
0.1
1
10
100
Favours hysterectomy
Favours first generation

Number of patients with adverse events – postoperatively (within 1 month)

Study or subgroup	First generation Events	Total	Hysterectomy Events	Total	Weight	Peto odds ratio (non-event) Peto, Fixed, 99% CI
1.19.1 Urinary tract infection (both)						
a. Dickersin 2007[92]	2	110	6	118	16.4%	2.61 (0.41 to 16.64)
b. Zupi 2003[78]	1	89	1	92	4.2%	0.97 (0.03 to 37.34)
c. Crosignani 1997[45]	0	41	0	44		Not estimable
d. O'Connor 1997[44]	0	116	0	56		Not estimable
e. Pinion 1994[43]	6	105	22	97	51.3%	4.11 (1.44 to 11.72)
f. Dwyer 1993[42]	0	99	12	97	24.0%	8.51 (1.84 to 39.35)
g. Gannon 1991[41]	0	25	2	26	4.1%	7.40 (0.19 to 293.55)
Subtotal (95% CI)		**585**		**530**	**100.0%**	**4.38 (2.48 to 7.75)**
Total events	9		43			
Heterogeneity: $\chi^2 = 3.06$, df = 4 ($p = 0.55$); $I^2 = 0\%$						
Test for overall effect: $z = 5.08$ ($p < 0.00001$)						
1.19.2 Deep vein thrombosis (both)						
a. Dickersin 2007[92]	0	110	2	118	100.0%	6.96 (0.18 to 268.76)
b. Zupi 2003[78]	0	89	0	92		Not estimable
c. Crosignani 1997[45]	0	41	0	44		Not estimable
d. O'Connor 1997[44]	0	116	0	56		Not estimable
e. Pinion 1994[43]	0	105	0	97		Not estimable
f. Dwyer 1993[42]	0	99	0	97		Not estimable
g. Gannon 1991[41]	0	25	0	26		Not estimable
Subtotal (95% CI)		**585**		**530**	**100.0%**	**6.96 (0.43 to 112.21)**
Total events	0		2			
Heterogeneity: not applicable						
Test for overall effect: $z = 1.37$ ($p = 0.17$)						
1.19.3 Haematometra (both)						
a. Dickersin 2007[92]	0	110	0	118		Not estimable
b. Zupi 2003[78]	0	89	0	92		Not estimable
c. Crosignani 1997[45]	0	41	1	44	4.2%	6.90 (0.04 to 1195.94)
d. O'Connor 1997[44]	0	116	0	56		Not estimable
e. Pinion 1994[43]	0	105	11	97	43.8%	8.95 (1.82 to 44.09)
f. Dwyer 1993[42]	1	99	8	97	36.2%	5.17 (0.90 to 29.88)
g. Gannon 1991[41]	0	25	4	26	15.8%	8.06 (0.57 to 114.90)
Subtotal (95% CI)		**585**		**530**	**100.0%**	**7.14 (3.20 to 15.94)**
Total events	1		24			
Heterogeneity: $\chi^2 = 0.37$, df = 3 ($p = 0.95$); $I^2 = 0\%$						
Test for overall effect: $z = 4.80$ ($p < 0.00001$)						
1.19.4 Excessive bleeding (hysterectomy)						
a. Dickersin 2007[92]	0	110	0	118		Not estimable
b. Zupi 2003[78]	0	89	0	92		Not estimable
c. Crosignani 1997[45]	0	41	0	44		Not estimable
d. O'Connor 1997[44]	0	116	3	56	30.8%	22.37 (0.92 to 544.67)
e. Pinion 1994[43]	0	105	0	97		Not estimable
f. Dwyer 1993[42]	0	99	6	97	69.2%	7.95 (0.94 to 66.99)
g. Gannon 1991[41]	0	25	0	26		Not estimable
Subtotal (95% CI)		**585**		**530**	**100.0%**	**10.94 (2.84 to 42.14)**
Total events	0		9			
Heterogeneity: $\chi^2 = 0.48$, df = 1 ($p = 0.49$); $I^2 = 0\%$						
Test for overall effect: $z = 3.48$ ($p = 0.0005$)						
1.19.5 Embolism (hysterectomy)						
a. Dickersin 2007[92]	0	110	2	118	100.0%	6.96 (0.18 to 268.76)
b. Zupi 2003[78]	0	89	0	92		Not estimable
c. Crosignani 1997[45]	0	41	0	44		Not estimable
d. O'Connor 1997[44]	0	116	0	56		Not estimable
e. Pinion 1994[43]	0	105	0	97		Not estimable
f. Dwyer 1993[42]	0	99	0	97		Not estimable
g. Gannon 1991[41]	0	25	0	26		Not estimable
Subtotal (95% CI)		**585**		**530**	**100.0%**	**6.96 (0.43 to 112.21)**
Total events	0		2			
Heterogeneity: not applicable						
Test for overall effect: $z = 1.37$ ($p = 0.17$)						




First generation
Hysterectomy
Study or subgroup Events Total Events Total Weight Peto odds ratio (non-event) Peto, Fixed, 99% CI
Peto odds ratio (non-event) Peto, Fixed, 99% CI
1.19.6 Further bleeding (ablation)
a. Dickersin 2007[92] 0 110 0 118 Not estimable
b. Zupi 2003[78] 0 89 0 92 Not estimable
c. Crosignani 1997[45] 0 41 0 44 Not estimable
d. O'Connor 1997[44] 0 116 0 56 Not estimable
e. Pinion 1994[43] 0 105 0 97 Not estimable
f. Dwyer 1993[42] 0 99 0 97 Not estimable
g. Gannon 1991[41] 0 25 0 26 Not estimable
Subtotal (95% CI) 585 530 Not estimable
Total events 0 0
Heterogeneity: not applicable
Test for overall effect: not applicable
1.19.7 Sepsis (ablation)
a. Dickersin 2007[92] 0 110 0 118 Not estimable
b. Zupi 2003[78] 0 89 0 92 Not estimable
c. Crosignani 1997[45] 0 41 0 44 Not estimable
d. O'Connor 1997[44] 9 116 0 56 100.0% 0.21 (0.03 to 1.38)
e. Pinion 1994[43] 0 105 0 97 Not estimable
f. Dwyer 1993[42] 0 99 0 97 Not estimable
g. Gannon 1991[41] 0 25 0 26 Not estimable
Subtotal (95% CI) 585 530 100.0% 0.21 (0.05 to 0.88)
Total events 9 0
Heterogeneity: not applicable
Test for overall effect: z = 2.13 (p = 0.03)
1.19.8 Pyrexia (ablation)
a. Dickersin 2007[92] 1 110 0 118 20.1% 0.13 (0.00 to 21.80)
b. Zupi 2003[78] 2 89 0 92 40.0% 0.13 (0.00 to 4.99)
c. Crosignani 1997[45] 0 41 0 44 Not estimable
d. O'Connor 1997[44] 0 116 0 56 Not estimable
e. Pinion 1994[43] 0 105 0 97 Not estimable
f. Dwyer 1993[42] 2 99 0 97 40.0% 0.14 (0.00 to 5.27)
g. Gannon 1991[41] 0 25 0 26 Not estimable
Subtotal (95% CI) 585 530 100.0% 0.13 (0.02 to 0.76)
Total events 5 0
Heterogeneity: χ² = 0.00, df = 2 (p = 1.00); I² = 0%
Test for overall effect: z = 2.26 (p = 0.02)
1.19.9 Endometriosis (ablation)
a. Dickersin 2007[92] 1 110 0 118 100.0% 0.13 (0.00 to 21.80)
b. Zupi 2003[78] 0 89 0 92 Not estimable
c. Crosignani 1997[45] 0 41 0 44 Not estimable
d. O'Connor 1997[44] 0 116 0 56 Not estimable
e. Pinion 1994[43] 0 105 0 97 Not estimable
f. Dwyer 1993[42] 0 99 0 97 Not estimable
g. Gannon 1991[41] 0 25 0 26 Not estimable
Subtotal (95% CI) 585 530 100.0% 0.13 (0.00 to 6.36)
Total events 1 0
Heterogeneity: not applicable
Test for overall effect: z = 1.04 (p = 0.30)
1.19.10 Abdominal pain (ablation)
a. Dickersin 2007[92] 0 110 0 118 Not estimable
b. Zupi 2003[78] 0 89 0 92 Not estimable
c. Crosignani 1997[45] 0 41 0 44 Not estimable
d. O'Connor 1997[44] 0 116 0 56 Not estimable
e. Pinion 1994[43] 0 105 0 97 Not estimable
f. Dwyer 1993[42] 0 99 0 97 Not estimable
g. Gannon 1991[41] 0 25 4 26 100.0% 8.06 (0.57 to 114.90)
Subtotal (95% CI) 585 530 100.0% 8.06 (1.07 to 60.87)
Total events 0 4
Heterogeneity: not applicable
Test for overall effect: z = 2.02 (p = 0.04)
0.01
0.1
1
10
100
Favours hysterectomy
Favours first generation



Appendix 5

Pooled results for hysterectomy versus Mirena

Appendix 5.1 Quality of life – clinical outcome

	Time point	Trials (no.)	WMD (95% CI)[a]	OR (95% CI)[b]	*p*-value	Hetero (*p*)/I^2 (%)
SF-36 general health (absolute)	6 months	1 (211)	–3.2 (–8.8 to 2.4)	–	0.3	–
SF-36 physical function (absolute)		1 (211)	–2.7 (–7.2 to 1.8)	–	0.2	–
SF-36 role physical (absolute)		1 (209)	–10.7 (–19.6 to –1.8)	–	0.02	–
SF-36 role emotional (absolute)		1 (208)	–8.3 (–17.5 to 0.9)	–	0.08	–
SF-36 mental health (absolute)		1 (211)	–5.3 (–10.1 to –0.5)	–	0.03	–
SF-36 social function (absolute)		1 (212)	–6.1 (–11.2 to –1.0)	–	0.02	–
SF-36 vitality (absolute)		1 (211)	–7.8 (–13.8 to –1.8)	–	0.01	–
SF-36 pain (absolute)		1 (212)	–5.7 (–11.8 to 0.4)	–	0.07	–
SF-36 general health (absolute)	12 months	1 (214)	–2.4 (–8.1 to 3.3)	–	0.4	–
SF-36 physical function (absolute)		1 (213)	–3.2 (–7.8 to 1.4)	–	0.2	–
SF-36 role physical (absolute)		1 (210)	–4.4 (–12.4 to 3.6)	–	0.3	–
SF-36 role emotional (absolute)		1 (208)	–9.4 (–18.3 to –0.50)	–	0.04	–
SF-36 mental health (absolute)		1 (214)	–3.9 (–8.6 to 0.8)	–	0.1	–
SF-36 social function (absolute)		1 (213)	–4.0 (–8.7 to 0.7)	–	0.09	–
SF-36 vitality (absolute)		1 (214)	–4.7 (–10.4 to 1.0)	–	0.1	–
SF-36 pain (absolute)		1 (213)	–6.5 (–12.4 to –0.6)	–	0.03	–
SF-36 general health (absolute)	5 years	1 (224)	–2.4 (–7.9 to 3.1)	–	0.4	–
SF-36 physical function (absolute)		1 (221)	–0.6 (–5.8 to 4.6)	–	0.8	–
SF-36 role physical (absolute)		1 (222)	–0.6 (–10.3 to 9.1)	–	0.9	–
SF-36 role emotional (absolute)		1 (225)	–1.0 (–10.2 to 8.2)	–	0.8	–
SF-36 mental health (absolute)		1 (225)	–2.9 (–7.2 to 1.4)	–	0.2	–
SF-36 social function (absolute)		1 (226)	–3.1 (–8.8 to 2.6)	–	0.3	–
SF-36 vitality (absolute)		1 (224)	–1.5 (–7.0 to 4.0)	–	0.6	–
SF-36 pain (absolute)		1 (226)	0.5 (–6.5 to 7.5)	–	0.9	–
SF-36 general health (change)	6 months	1 (209)	–3.8 (–8.2 to 0.6)	–	0.09	–
SF-36 physical function (change)		1 (207)	–2.7 (–6.6 to 1.2)	–	0.2	–
SF-36 role physical (change)		1 (206)	–8.3 (–19.3 to 2.7)	–	0.1	–
SF-36 role emotional (change)		1 (202)	–2.9 (–14.6 to 8.8)	–	0.6	–
SF-36 mental health (change)		1 (207)	–2.9 (–8.0 to –2.2)	–	0.3	–
SF-36 social function (change)		1 (210)	–2.2 (–7.5 to 3.1)	–	0.4	–
SF-36 vitality (change)		1 (208)	–5.9 (–12.0 to 0.2)	–	0.06	–
SF-36 pain (change)		1 (208)	–6.8 (–13.6 to 0.01)	–	0.05	–
SF-36 general health (change)	12 months	1 (212)	–0.6 (–4.9 to 3.7)	–	0.8	–
SF-36 physical function (change)		1 (209)	–2.3 (–6.6 to 2.0)	–	0.3	–
SF-36 role physical (change)		1 (208)	–0.5 (–11.6 to 10.6)	–	0.9	–
SF-36 role emotional (change)		1 (206)	–4.1 (–15.7 to 7.5)	–	0.5	–
SF-36 mental health (change)		1 (210)	–0.3 (–5.2 to 4.6)	–	0.9	–

continued

	Time point	Trials (no.)	WMD (95% CI)[a]	OR (95% CI)[b]	*p*-value	Hetero (*p*)/I^2 (%)
SF-36 social function (change)		1 (212)	–0.5 (–5.7 to 4.7)	–	0.9	–
SF-36 vitality (change)		1 (211)	–2.3 (–7.9 to 3.3)	–	0.4	–
SF-36 pain (change)		1 (210)	–9.6 (–16.6 to –2.7)	–	0.007	–
SF-36 general health (change)	5 years	1 (222)	–1.3 (–6.2 to 3.6)	–	0.6	–
SF-36 physical function (change)		1 (216)	–0.5 (–5.5 to 4.5)	–	0.9	–
SF-36 role physical (change)		1 (219)	–1.1 (–12.3 to 10.1)	–	0.9	–
SF-36 role emotional (change)		1 (221)	4.1 (–7.7 to 15.9)	–	0.5	–
SF-36 mental health (change)		1 (220)	0.5 (–4.4 to 5.4)	–	0.8	–
SF-36 social function (change)		1 (224)	–0.1 (–5.8 to 5.6)	–	1.0	–
SF-36 vitality (change)		1 (221)	–0.4 (–6.2 to 5.4)	–	0.9	–
SF-36 pain (change)		1 (222)	–0.6 (–8.0 to 6.8)	–	0.9	–
EQ-5D (absolute)	6 months	1 (214)	–0.04 (–0.09 to 0.01)	–	0.1	–
	12 months	1 (213)	–0.02 (–0.06 to 0.02)	–	0.4	–
	5 years	1 (224)	–0.03 (–0.08 to 0.02)	–	0.3	–
EQ-5D (change)	6 months	1 (210)	–0.01 (–0.06 to 0.04)	–	0.7	–
	12 months	1 (209)	–0.00 (–0.05 to 0.05)	–	1.0	–
	5 years	1 (220)	–0.01 (–0.07 to 0.05)	–	0.7	–
		Trials	**Frequency**			
Discontinued Mirena	6 months	1	22/119 (18%)			
	12 months	1	37/119 (31%)			
	5 years	1	60/119 (50%)			
Hysterectomy after Mirena	6 months	1	9/119 (8%)			
	12 months	1	24/119 (20%)			

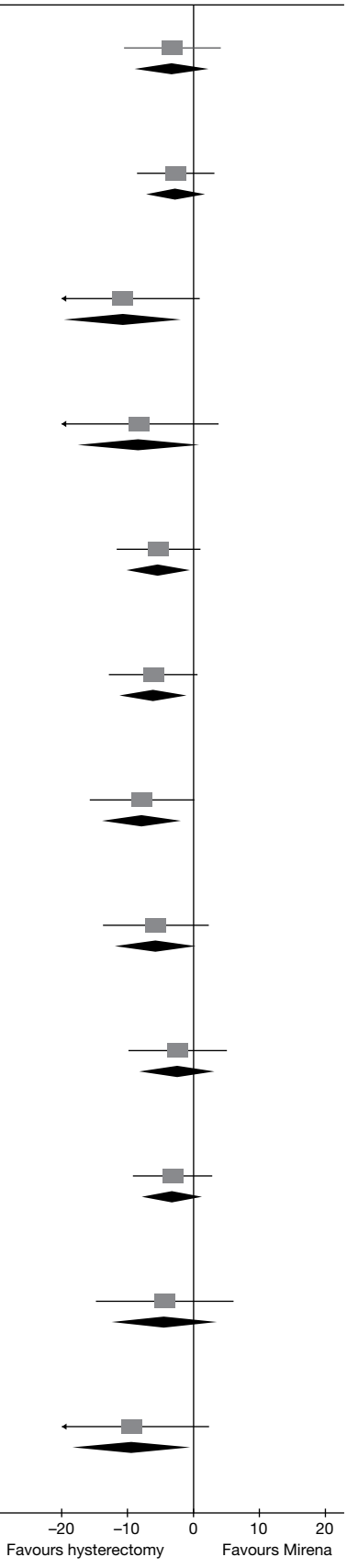
	Trials	Frequency (hysterectomy: max. 117; Mirena: max. 119)	OR (95% CI)[a]	*p*-value	Hetero (*p*)/I^2 (%)
Periprocedure complications (hysterectomy)					
Anaesthesia problems	1	0	–	–	–
Excessive bleeding	1	0	–	-–	–
Injury surrounding organs	1	5	–	–	–
Further complications (hysterectomy, <1 month)					
Urinary tract infection	1	0	–	–	–
Deep-vein thrombosis	1	0	–	–	–
Excessive bleeding	1	0	–	–	–
Embolism	1	0	–	–	–
Complication post-insertion (Mirena)					
Uterine perforation	1	0	–	–	–
Infection	1	5	–	–	–
Expelled/migrated	1	0	–	–	–
Cervical laceration	1	0	–	–	–
Failed to insert	1	2	–	–	–
Removed (before 3 months)	1	10	–	–	–

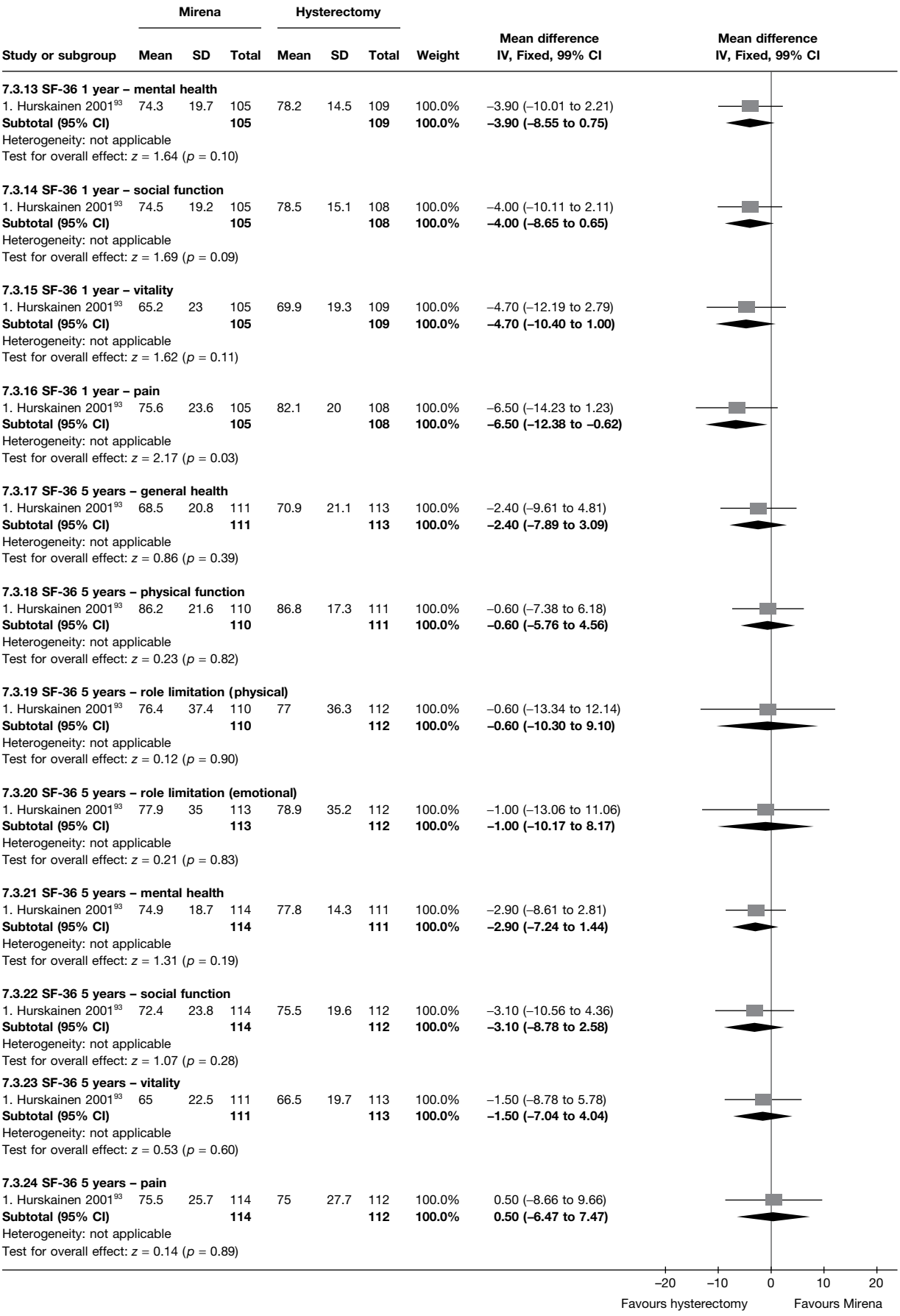
a < 0 favours hysterectomy, > 0 favours Mirena.
b < 1 favours hysterectomy, > 1 favours Mirena.

Appendix 5.2 Hysterectomy versus Mirena

SF-36 scores (absolute values)

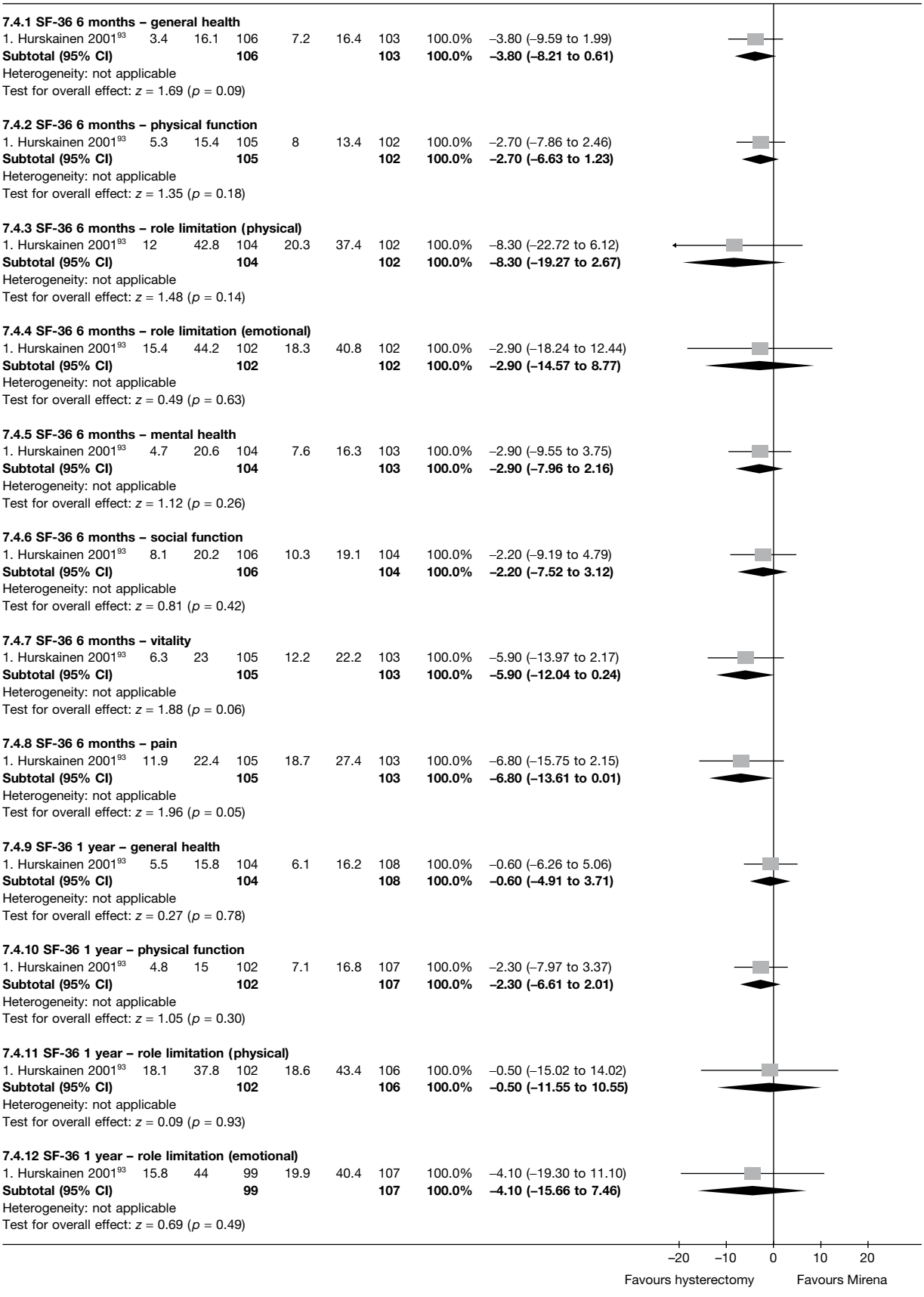
Study or subgroup	Mirena Mean	Mirena SD	Mirena Total	Hysterectomy Mean	Hysterectomy SD	Hysterectomy Total	Weight	Mean difference IV, Fixed, 99% CI
7.3.1 SF-36 6 months – general health								
1. Hurskainen 2001[93]	70.3	20.3	107	73.5	21	104	100.0%	–3.20 (–10.53 to 4.13)
Subtotal (95% CI)			**107**			**104**	**100.0%**	**–3.20 (–8.78 to 2.38)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.12$ ($p = 0.26$)								
7.3.2 SF-36 6 months – physical function								
1. Hurskainen 2001[93]	88.9	18.3	107	91.6	14.7	104	100.0%	–2.70 (–8.58 to 3.18)
Subtotal (95% CI)			**107**			**104**	**100.0%**	**–2.70 (–7.17 to 1.77)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.18$ ($p = 0.24$)								
7.3.3 SF-36 6 months – role limitation (physical)								
1. Hurskainen 2001[93]	76.7	36.7	106	87.4	28.2	103	100.0%	–10.70 (–22.34 to 0.94)
Subtotal (95% CI)			**106**			**103**	**100.0%**	**–10.70 (–19.56 to –1.84)**
Heterogeneity: not applicable								
Test for overall effect: $z = 2.37$ ($p = 0.02$)								
7.3.4 SF-36 6 months – role limitation (emotional)								
1. Hurskainen 2001[93]	77.5	37.4	105	85.8	30.1	103	100.0%	–8.30 (–20.41 to 3.81)
Subtotal (95% CI)			**105**			**103**	**100.0%**	**–8.30 (–17.52 to 0.92)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.76$ ($p = 0.08$)								
7.3.5 SF-36 6 months – mental health								
1. Hurskainen 2001[93]	72.4	19.5	107	77.7	16.2	104	100.0%	–5.30 (–11.65 to 1.05)
Subtotal (95% CI)			**107**			**104**	**100.0%**	**–5.30 (–10.13 to –0.47)**
Heterogeneity: not applicable								
Test for overall effect: $z = 2.15$ ($p = 0.03$)								
7.3.6 SF-36 6 months – social function								
1. Hurskainen 2001[93]	71.5	21.3	108	77.6	16.5	104	100.0%	–6.10 (–12.83 to 0.63)
Subtotal (95% CI)			**108**			**104**	**100.0%**	**–6.10 (–11.22 to –0.98)**
Heterogeneity: not applicable								
Test for overall effect: $z = 2.34$ ($p = 0.02$)								
7.3.7 SF-36 6 months – vitality								
1. Hurskainen 2001[93]	61.9	24.1	108	69.7	20.5	103	100.0%	–7.80 (–15.72 to 0.12)
Subtotal (95% CI)			**108**			**103**	**100.0%**	**–7.80 (–13.83 to –1.77)**
Heterogeneity: not applicable								
Test for overall effect: $z = 2.54$ ($p = 0.01$)								
7.3.8 SF-36 6 months – pain								
1. Hurskainen 2001[93]	74.3	22.4	108	80	23	104	100.0%	–5.70 (–13.74 to 2.34)
Subtotal (95% CI)			**108**			**104**	**100.0%**	**–5.70 (–11.81 to 0.41)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.83$ ($p = 0.07$)								
7.3.9 SF-36 1 year – general health								
1. Hurskainen 2001[93]	70.7	19.8	105	73.1	22.6	109	100.0%	–2.40 (–9.87 to 5.07)
Subtotal (95% CI)			**105**			**109**	**100.0%**	**–2.40 (–8.09 to 3.29)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.83$ ($p = 0.41$)								
7.3.10 SF-36 1 year – physical function								
1. Hurskainen 2001[93]	88.6	19.4	104	91.8	14.2	109	100.0%	–3.20 (–9.22 to 2.82)
Subtotal (95% CI)			**104**			**109**	**100.0%**	**–3.20 (–7.78 to 1.38)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.37$ ($p = 0.17$)								
7.3.11 SF-36 1 year – role limitation (physical)								
1. Hurskainen 2001[93]	83	30.8	103	87.4	27.8	107	100.0%	–4.40 (–14.84 to 6.04)
Subtotal (95% CI)			**103**			**107**	**100.0%**	**–4.40 (–12.35 to 3.55)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.09$ ($p = 0.28$)								
7.3.12 SF-36 1 year – role limitation (emotional)								
1. Hurskainen 2001[93]	77	35.7	100	86.4	29.2	108	100.0%	–9.40 (–21.10 to 2.30)
Subtotal (95% CI)			**100**			**108**	**100.0%**	**–9.40 (–18.30 to –0.50)**
Heterogeneity: not applicable								
Test for overall effect: $z = 2.07$ ($p = 0.04$)								

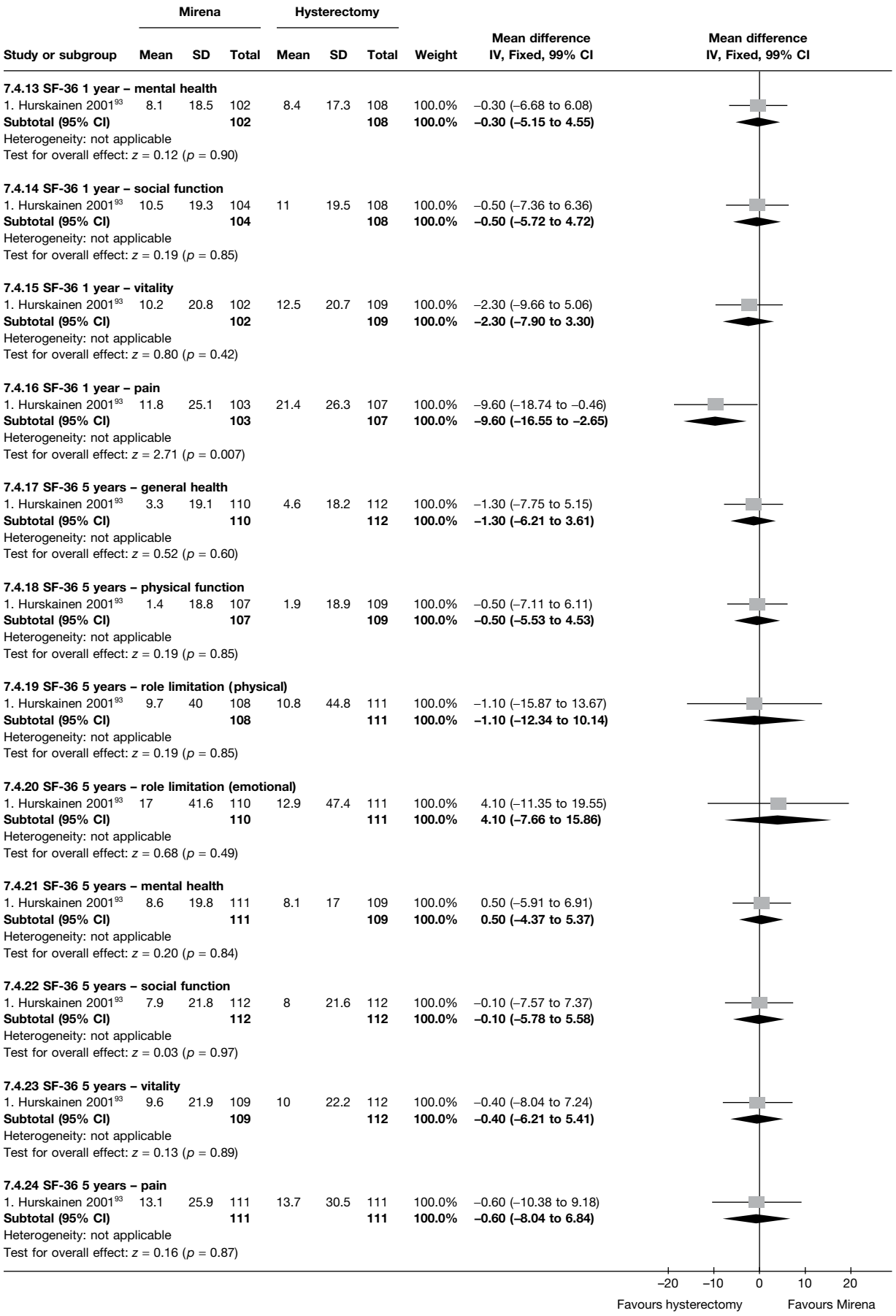




SF-36 scores (change from baseline)

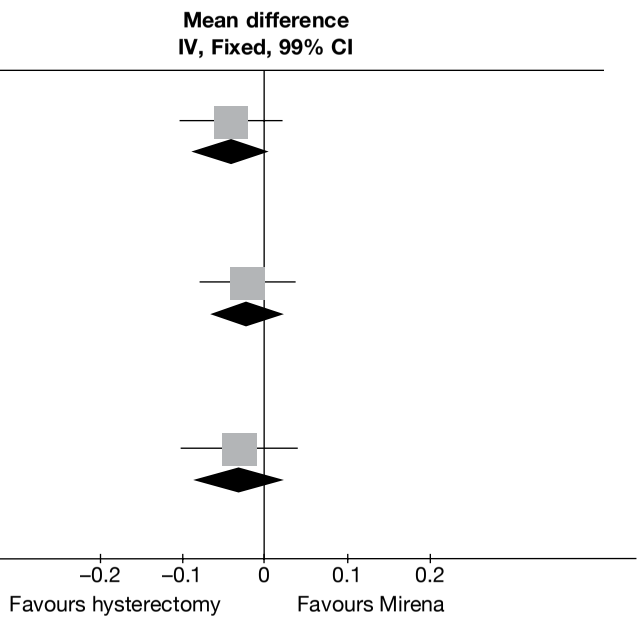
Study or subgroup	Mirena Mean	Mirena SD	Mirena Total	Hysterectomy Mean	Hysterectomy SD	Hysterectomy Total	Weight	Mean difference IV, Fixed, 99% CI
7.4.1 SF-36 6 months – general health								
1. Hurskainen 2001[93]	3.4	16.1	106	7.2	16.4	103	100.0%	–3.80 (–9.59 to 1.99)
Subtotal (95% CI)			**106**			**103**	**100.0%**	**–3.80 (–8.21 to 0.61)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.69$ ($p = 0.09$)								
7.4.2 SF-36 6 months – physical function								
1. Hurskainen 2001[93]	5.3	15.4	105	8	13.4	102	100.0%	–2.70 (–7.86 to 2.46)
Subtotal (95% CI)			**105**			**102**	**100.0%**	**–2.70 (–6.63 to 1.23)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.35$ ($p = 0.18$)								
7.4.3 SF-36 6 months – role limitation (physical)								
1. Hurskainen 2001[93]	12	42.8	104	20.3	37.4	102	100.0%	–8.30 (–22.72 to 6.12)
Subtotal (95% CI)			**104**			**102**	**100.0%**	**–8.30 (–19.27 to 2.67)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.48$ ($p = 0.14$)								
7.4.4 SF-36 6 months – role limitation (emotional)								
1. Hurskainen 2001[93]	15.4	44.2	102	18.3	40.8	102	100.0%	–2.90 (–18.24 to 12.44)
Subtotal (95% CI)			**102**			**102**	**100.0%**	**–2.90 (–14.57 to 8.77)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.49$ ($p = 0.63$)								
7.4.5 SF-36 6 months – mental health								
1. Hurskainen 2001[93]	4.7	20.6	104	7.6	16.3	103	100.0%	–2.90 (–9.55 to 3.75)
Subtotal (95% CI)			**104**			**103**	**100.0%**	**–2.90 (–7.96 to 2.16)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.12$ ($p = 0.26$)								
7.4.6 SF-36 6 months – social function								
1. Hurskainen 2001[93]	8.1	20.2	106	10.3	19.1	104	100.0%	–2.20 (–9.19 to 4.79)
Subtotal (95% CI)			**106**			**104**	**100.0%**	**–2.20 (–7.52 to 3.12)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.81$ ($p = 0.42$)								
7.4.7 SF-36 6 months – vitality								
1. Hurskainen 2001[93]	6.3	23	105	12.2	22.2	103	100.0%	–5.90 (–13.97 to 2.17)
Subtotal (95% CI)			**105**			**103**	**100.0%**	**–5.90 (–12.04 to 0.24)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.88$ ($p = 0.06$)								
7.4.8 SF-36 6 months – pain								
1. Hurskainen 2001[93]	11.9	22.4	105	18.7	27.4	103	100.0%	–6.80 (–15.75 to 2.15)
Subtotal (95% CI)			**105**			**103**	**100.0%**	**–6.80 (–13.61 to 0.01)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.96$ ($p = 0.05$)								
7.4.9 SF-36 1 year – general health								
1. Hurskainen 2001[93]	5.5	15.8	104	6.1	16.2	108	100.0%	–0.60 (–6.26 to 5.06)
Subtotal (95% CI)			**104**			**108**	**100.0%**	**–0.60 (–4.91 to 3.71)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.27$ ($p = 0.78$)								
7.4.10 SF-36 1 year – physical function								
1. Hurskainen 2001[93]	4.8	15	102	7.1	16.8	107	100.0%	–2.30 (–7.97 to 3.37)
Subtotal (95% CI)			**102**			**107**	**100.0%**	**–2.30 (–6.61 to 2.01)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.05$ ($p = 0.30$)								
7.4.11 SF-36 1 year – role limitation (physical)								
1. Hurskainen 2001[93]	18.1	37.8	102	18.6	43.4	106	100.0%	–0.50 (–15.02 to 14.02)
Subtotal (95% CI)			**102**			**106**	**100.0%**	**–0.50 (–11.55 to 10.55)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.09$ ($p = 0.93$)								
7.4.12 SF-36 1 year – role limitation (emotional)								
1. Hurskainen 2001[93]	15.8	44	99	19.9	40.4	107	100.0%	–4.10 (–19.30 to 11.10)
Subtotal (95% CI)			**99**			**107**	**100.0%**	**–4.10 (–15.66 to 7.46)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.69$ ($p = 0.49$)								





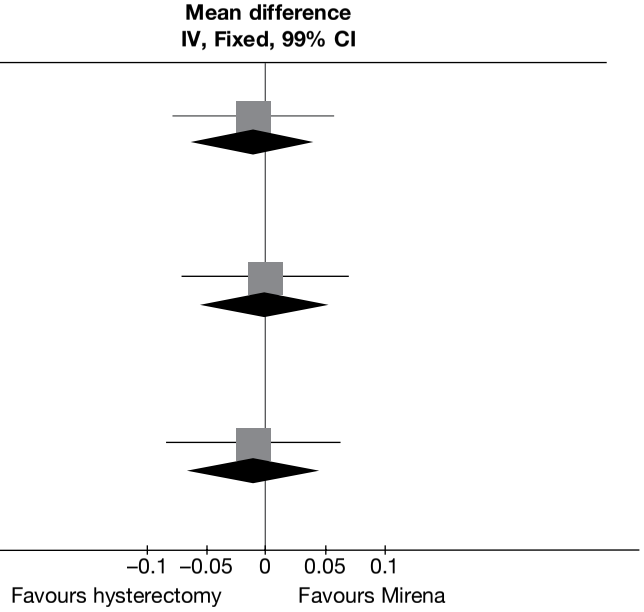
EQ-5D scores (absolute values)

	Mirena			Hysterectomy				Mean difference	Mean difference
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 99% CI**	**IV, Fixed, 99% CI**
7.5.2 6 months									
1. Hurskainen 2001[93]	0.85	0.21	111	0.89	0.14	103	100.0%	–0.04 (–0.10 to 0.02)	
Subtotal (95% CI)			**111**			**103**	**100.0%**	**–0.04 (–0.09 to 0.01)**	
Heterogeneity: not applicable									
Test for overall effect: $z = 1.65$ ($p = 0.10$)									
7.5.3 1 year									
1. Hurskainen 2001[93]	0.88	0.16	104	0.9	0.17	109	100.0%	–0.02 (–0.08 to 0.04)	
Subtotal (95% CI)			**104**			**109**	**100.0%**	**–0.02 (–0.06 to 0.02)**	
Heterogeneity: not applicable									
Test for overall effect: $z = 0.88$ ($p = 0.38$)									
7.5.4 5 years									
1. Hurskainen 2001[93]	0.84	0.23	112	0.87	0.18	112	100.0%	–0.03 (–0.10 to 0.04)	
Subtotal (95% CI)			**112**			**112**	**100.0%**	**–0.03 (–0.08 to 0.02)**	
Heterogeneity: not applicable									
Test for overall effect: $z = 1.09$ ($p = 0.28$)									



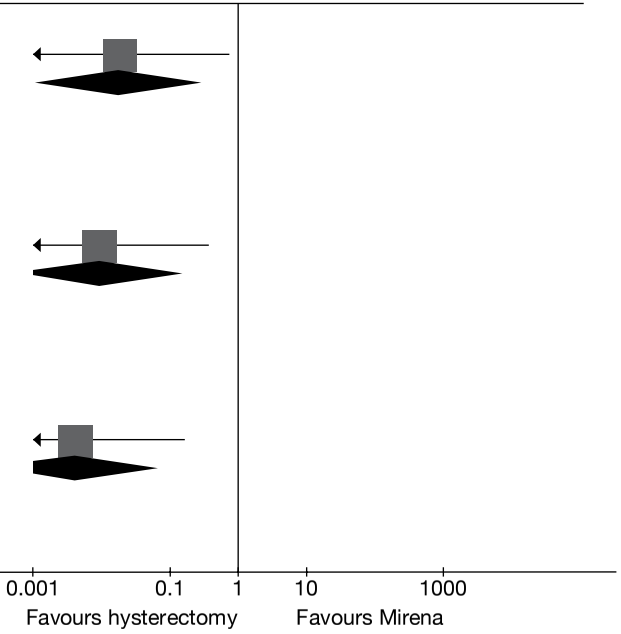
EQ-5D scores (change from baseline)

Study or subgroup	Mirena Mean	Mirena SD	Mirena Total	Hysterectomy Mean	Hysterectomy SD	Hysterectomy Total	Weight	Mean difference IV, Fixed, 99% CI
7.6.1 6 months								
1. Hurskainen 2001[93]	0.07	0.21	108	0.08	0.17	102	100.0%	–0.01 (–0.08 to 0.06)
Subtotal (95% CI)			**108**			**102**	**100.0%**	**–0.01 (–0.06 to 0.04)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.38$ ($p = 0.70$)								
7.6.2 1 year								
1. Hurskainen 2001[93]	0.07	0.16	101	0.07	0.23	108	100.0%	0.00 (–0.07 to 0.07)
Subtotal (95% CI)			**101**			**108**	**100.0%**	**0.00 (–0.05 to 0.05)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.00$ ($p = 1.00$)								
7.6.3 5 years								
1. Hurskainen 2001[93]	0.04	0.19	109	0.05	0.23	111	100.0%	–0.01 (–0.08 to 0.06)
Subtotal (95% CI)			**109**			**111**	**100.0%**	**–0.01 (–0.07 to 0.05)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.35$ ($p = 0.72$)								



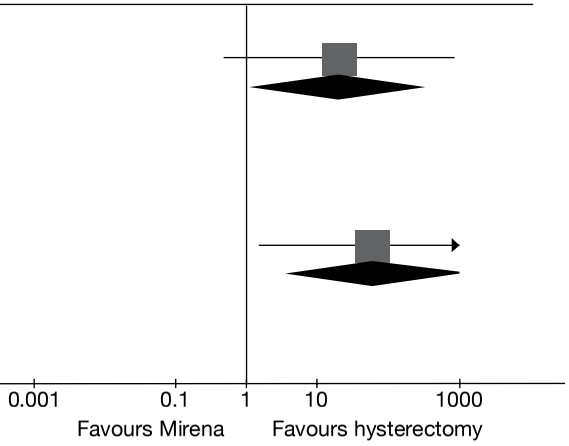
Proportion discontinuing Mirena

Study or subgroup	Mirena Events	Mirena Total	Hysterectomy Events	Hysterectomy Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
7.10.1 6 months						
1. Hurskainen 2001[93]	0	119	22	117	100.0%	0.02 (0.00 to 0.75)
Subtotal (95% CI)		**119**		**117**	**100.0%**	**0.02 (0.00 to 0.31)**
Total events	0		22			
Heterogeneity: not applicable						
Test for overall effect: $z = 2.78$ ($p = 0.005$)						
7.10.2 1 year						
1. Hurskainen 2001[93]	0	119	37	117	100.0%	0.01 (0.00 to 0.37)
Subtotal (95% CI)		**119**		**117**	**100.0%**	**0.01 (0.00 to 0.15)**
Total events	0		37			
Heterogeneity: not applicable						
Test for overall effect: $z = 3.26$ ($p = 0.001$)						
7.10.3 5 years						
1. Hurskainen 2001[93]	0	119	60	117	100.0%	0.00 (0.00 to 0.17)
Subtotal (95% CI)		**119**		**117**	**100.0%**	**0.00 (0.00 to 0.07)**
Total events	0		60			
Heterogeneity: not applicable						
Test for overall effect: $z = 3.83$ ($p = 0.0001$)						



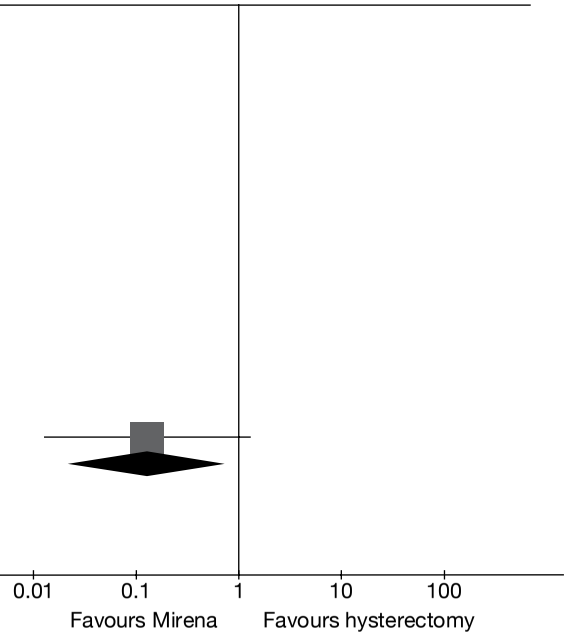
Proportion requiring hysterectomy after Mirena

Study or subgroup	Mirena Events	Mirena Total	Hysterectomy Events	Hysterectomy Total	Weight	Odds ratio M–H, Fixed, 99% CI
7.13.1 6 months						
1. Hurskainen 2001[93]	9	119	0	117	100.0%	20.20 (0.47 to 861.63)
Subtotal (95% CI)		**119**		**117**	**100.0%**	**20.20 (1.16 to 351.25)**
Total events	9		0			
Heterogeneity: not applicable						
Test for overall effect: $z = 2.06$ ($p = 0.04$)						
7.13.3 1 year						
1. Hurskainen 2001[93]	24	119	0	117	100.0%	60.29 (1.50 to 2430.77)
Subtotal (95% CI)		**119**		**117**	**100.0%**	**60.29 (3.62 to 1004.33)**
Total events	24		0			
Heterogeneity: not applicable						
Test for overall effect: $z = 2.86$ ($p = 0.004$)						



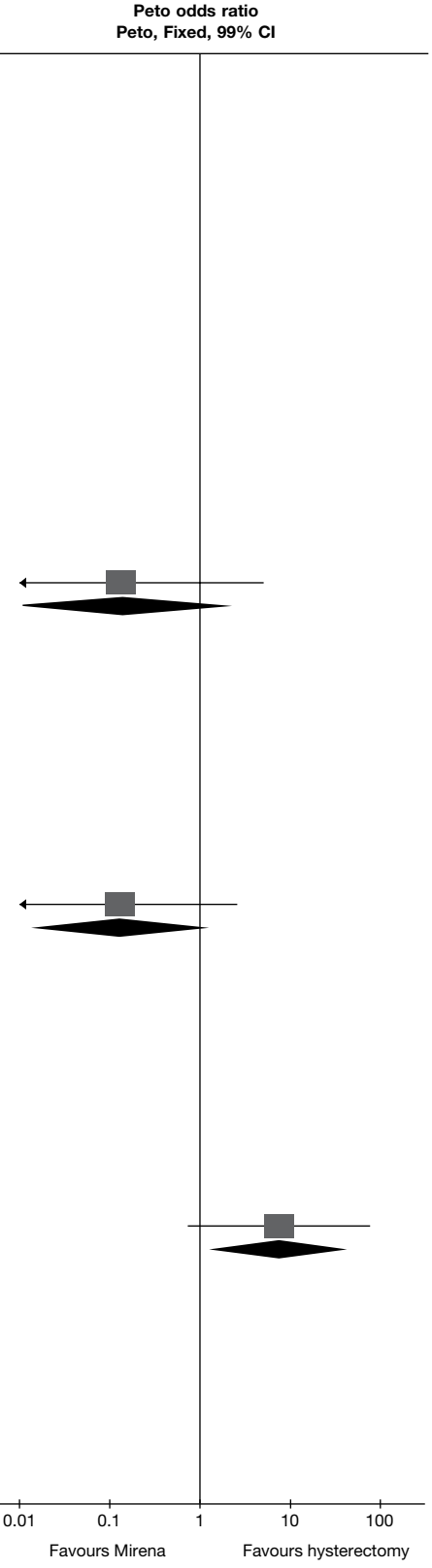
Number of patients with adverse events – periprocedure

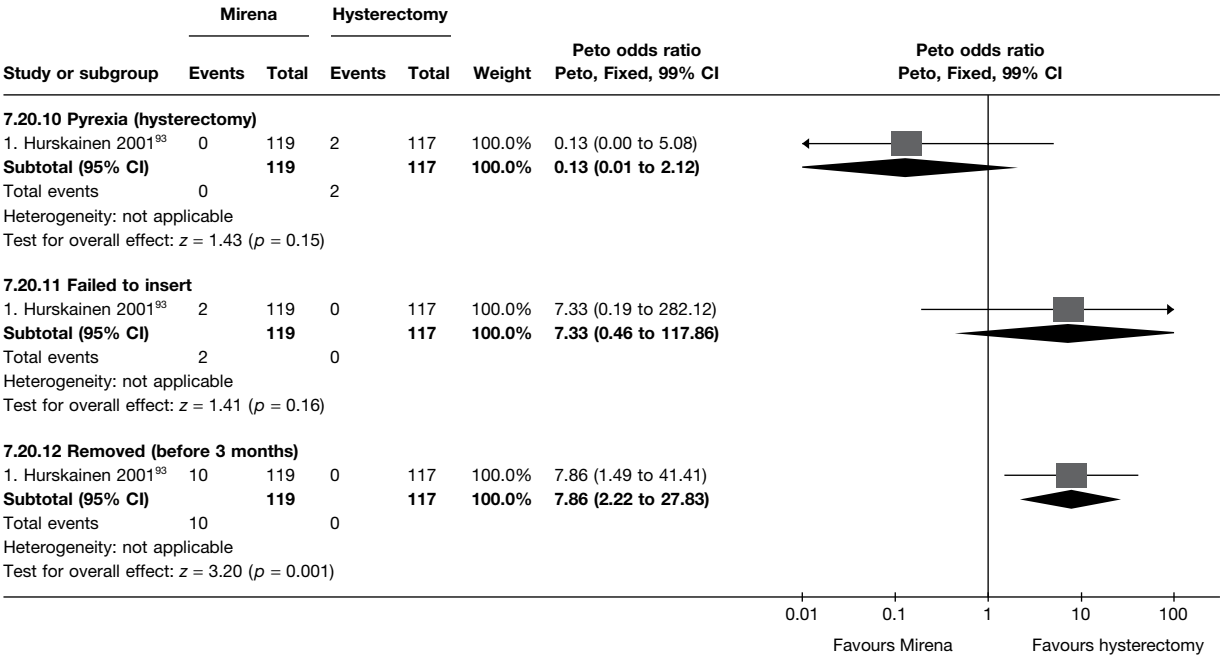
Study or subgroup	Mirena Events	Mirena Total	Hysterectomy Events	Hysterectomy Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
7.14.1 Anaesthesia problems (hysterectomy)						
1. Hurskainen 2001[93]	0	119	0	117		Not estimable
Subtotal (95% CI)		**119**		**117**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
7.14.2 Excessive bleeding (hysterectomy)						
1. Hurskainen 2001[93]	0	119	0	117		Not estimable
Subtotal (95% CI)		**119**		**117**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
7.14.3 Injury to surrounding organs (hysterectomy)						
1. Hurskainen 2001[93]	0	119	5	117	100.0%	0.13 (0.01 to 1.31)
Subtotal (95% CI)		**119**		**117**	**100.0%**	**0.13 (0.02 to 0.75)**
Total events	0		5			
Heterogeneity: not applicable						
Test for overall effect: $z = 2.27$ ($p = 0.02$)						



Number of patients with adverse events – postoperatively (within 1 month)

Study or subgroup	Mirena Events	Mirena Total	Hysterectomy Events	Hysterectomy Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
7.20.1 Urinary tract infection (both)						
1. Hurskainen 2001[93]	0	119	0	117		Not estimable
Subtotal (95% CI)		**119**		**117**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
7.20.2 Deep vein thrombosis (hysterectomy)						
1. Hurskainen 2001[93]	0	119	0	117		Not estimable
Subtotal (95% CI)		**119**		**117**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
7.20.3 Excessive bleeding (hysterectomy)						
1. Hurskainen 2001[93]	0	119	0	117		Not estimable
Subtotal (95% CI)		**119**		**117**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
7.20.4 Further bleeding (hysterectomy)						
1. Hurskainen 2001[93]	0	119	2	117	100.0%	0.13 (0.00 to 5.08)
Subtotal (95% CI)		**119**		**117**	**100.0%**	**0.13 (0.01 to 2.12)**
Total events	0		2			
Heterogeneity: not applicable						
Test for overall effect: $z = 1.43$ ($p = 0.15$)						
7.20.5 Embolism (hysterectomy)						
1. Hurskainen 2001[93]	0	119	0	117		Not estimable
Subtotal (95% CI)		**119**		**117**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
7.20.6 Abdominal pain (hysterectomy)						
1. Hurskainen 2001[93]	0	119	3	117	100.0%	0.13 (0.01 to 2.59)
Subtotal (95% CI)		**119**		**117**	**100.0%**	**0.13 (0.01 to 1.27)**
Total events	0		3			
Heterogeneity: not applicable						
Test for overall effect: $z = 1.75$ ($p = 0.08$)						
7.20.7 Uterine perforation (Mirena)						
1. Hurskainen 2001[93]	0	119	0	117		Not estimable
Subtotal (95% CI)		**119**		**117**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
7.20.8 Infection (Mirena)						
1. Hurskainen 2001[93]	5	119	0	117	100.0%	7.52 (0.74 to 76.81)
Subtotal (95% CI)		**119**		**117**	**100.0%**	**7.52 (1.28 to 44.07)**
Total events	5		0			
Heterogeneity: not applicable						
Test for overall effect: $z = 2.24$ ($p = 0.03$)						
7.20.9 Migrated coil (Mirena)						
1. Hurskainen 2001[93]	0	119	0	117		Not estimable
Subtotal (95% CI)		**119**		**117**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						




Mirena
Hysterectomy
Study or subgroup
Events
Total
Events
Total
Weight
Peto odds ratio
Peto, Fixed, 99% CI
Peto odds ratio
Peto, Fixed, 99% CI
7.20.10 Pyrexia (hysterectomy)
1. Hurskainen 2001[93] 0 119 2 117 100.0% 0.13 (0.00 to 5.08)
Subtotal (95% CI) 119 117 100.0% 0.13 (0.01 to 2.12)
Total events 0 2
Heterogeneity: not applicable
Test for overall effect: z = 1.43 (p = 0.15)
7.20.11 Failed to insert
1. Hurskainen 2001[93] 2 119 0 117 100.0% 7.33 (0.19 to 282.12)
Subtotal (95% CI) 119 117 100.0% 7.33 (0.46 to 117.86)
Total events 2 0
Heterogeneity: not applicable
Test for overall effect: z = 1.41 (p = 0.16)
7.20.12 Removed (before 3 months)
1. Hurskainen 2001[93] 10 119 0 117 100.0% 7.86 (1.49 to 41.41)
Subtotal (95% CI) 119 117 100.0% 7.86 (2.22 to 27.83)
Total events 10 0
Heterogeneity: not applicable
Test for overall effect: z = 3.20 (p = 0.001)
0.01
0.1
1
10
100
Favours Mirena
Favours hysterectomy

Appendix 6

Pooled results for first- versus second-generation endometrial ablation

Appendix 6.1 Baseline characteristics, quality of life and clinical outcome

	Time point	Trials (no.)	WMD (95% CI)[a]	OR (95% CI)[b]	*p*-value	Hetero (*p*)/I^2 (%)
Proportion amenorrhoea	6 months	5 (736)	–	1.16 (0.81 to 1.66)	0.4	0.09/50
	12 months	13 (2180)	–	1.12 (0.93 to 1.35)	0.3	<0.0001/74
	2 years	2 (370)	–	0.64 (0.41 to 0.99)	0.04	0.2/36
	3 years	1 (111)	–	0.24 (0.11 to 0.50)	0.0002	–
	5 years	1 (236)	–	1.19 (0.70 to 2.05)	0.5	–
	10 years	1 (189)	–	1.56 (0.69 to 3.51)	0.3	–
Proportion with heavy bleeding	6 months	5 (736)	–	1.33 (0.92 to 1.93)	0.1	0.5/0
	12 months	13 (2180)	–	0.97 (0.74 to 1.28)	0.9	0.8/0
	2 years	2 (370)	–	0.54 (0.30 to 0.97)	0.04	0.8/0
	3 years	1 (111)	–	0.58 (0.14 to 2.41)	0.5	–
	5 years	1 (266)	–	1.05 (0.30 to 3.70)	0.9	–
Bleeding score (change)	6 months	6 (1001)	–2 (–49 to 45)	–	0.9	0.2/33
	12 months	9 (1778)	–10 (–37 to 17)	–	0.5	0.009/61
	2 years	1 121)	6 (–122 to 134)	–	0.9	–
Proportion dysmenorrhoea	6 months	4 (562)	–	0.95 (0.64 to 1.41)	0.8	0.4/0
	12 months	8 (1548)	–	0.84 (0.67 to 1.07)	0.2	0.5/0
	2 years	2 (475)	–	0.95 (0.62 to 1.46)	0.8	0.3/0
	3 years	1 (212)	–	0.91 (0.47 to 1.76)	0.8	–
	5 years	1 (266)	–	1.05 (0.48 to 2.30)	0.9	–
Duration surgery (minutes)		11 (1911)	–15 (–15 to –14)	–	<0.0001	<0.0001/96
General anaesthesia		8 (1597)	–	0.16 (0.12 to 0.20)	<0.0001	<0.0001/86
Surgery pain score (0–10)		5 (342)	0.05 (–0.17 to 0.27)	–	0.7	<0.0001/89
Return to work (days)		2 (116)	–1.4 (–2.0 to –0.7)	–	<0.0001	0.3/10
Return normal activities (days)		5 (901)	–0.48 (–0.75 to –0.20)	–	0.0008	0.04/59
Proportion dyspareunia	6 months	2 (106)	–	1.09 (0.27 to 4.41)	0.9	–
	12 months	3 (330)	–	0.89 (0.46 to 1.73)	0.7	0.2/32
	2 years	1 (247)	–	0.95 (0.46 to 1.96)	0.9	–
	5 years	1 (218)	–	0.40 (0.18 to 0.93)	0.03	–
SF-36 general health (absolute)	6 months	1 (265)	0.6 (–3.5 to 4.7)	–	0.8	–
SF-36 physical function (absolute)		1 (267)	3.0 (–0.6 to 6.6)	–	0.1	–
SF-36 role physical (absolute)		1 (273)	3.3 (–3.1 to 9.7)	–	0.3	–
SF-36 role emotional (absolute)		1 (271)	3.3 (–4.0 to 10.6)	–	0.4	–
SF-36 mental health (absolute)		1 (269)	0.8 (–3.5 to 5.1)	–	0.7	–
SF-36 social function (absolute)		1 (257)	0.6 (–2.8 to 4.0)	–	0.7	–
SF-36 vitality (absolute)		1 (269)	0.6 (–4.6 to 5.8)	–	0.8	–
SF-36 pain (absolute)		1 (269)	0.7 (–4.6 to 6.0)	–	0.8	–

continued

	Time point	Trials (no.)	WMD (95% CI)[a]	OR (95% CI)[b]	*p*-value	Hetero (*p*)/I^2 (%)
SF-36 general health (absolute)	12 months	2 (522)	–1.5 (–4.5 to 1.4)	–	0.3	0.7/0
SF-36 physical function (absolute)		2 (519)	–0.4 (–3.6 to 2.7)	–	0.8	0.9/0
SF-36 role physical (absolute)		2 (512)	–5.8 (–11.0 to –0.6)	–	0.03	0.7/0
SF-36 role emotional (absolute)		2 (521)	–2.2 (–7.5 to 3.2)	–	0.4	0.7/0
SF-36 mental health (absolute)		2 (521)	–1.5 (–4.8 to 1.9)	–	0.4	0.8/0
SF-36 social function (absolute)		2 (512)	–1.0 (–3.9 to 1.9)	–	0.5	0.3/25
SF-36 vitality (absolute)		2 (521)	–3.1 (–7.0 to 0.9)	–	0.1	0.7/0
SF-36 pain (absolute)		2 (522)	0.7 (–3.6 to 4.9)	–	0.8	0.9/0
SF-36 general health (absolute)	2 years	1 (249)	0.3 (–5.9 to 6.5)	–	0.9	–
SF-36 physical function (absolute)		1 (249)	–2.4 (–8.1 to 3.3)	–	0.4	–
SF-36 role physical (absolute)		1 (249)	–3.9 (–13.9 to 6.1)	–	0.5	–
SF-36 role emotional (absolute)		1 (249)	–5.6 (–15.4 to 4.2)	–	0.3	–
SF-36 mental health (absolute)		1 (249)	–1.3 (–6.5 to 3.9)	–	0.6	–
SF-36 social function (absolute)		1 (249)	–3.2 (–9.2 to 2.8)	–	0.3	–
SF-36 vitality (absolute)		1 (249)	0.4 (–5.5 to 6.3)	–	0.9	–
SF-36 pain (absolute)		1 (249)	–2.0 (–9.1 to 5.1)	–	0.6	–
SF-36 general health (absolute)	5 years	1 (235)	2.8 (–3.6 to 9.2)	–	0.4	–
SF-36 physical function (absolute)		1 (232)	–2.2 (–8.7 to 4.3)	–	0.5	–
SF-36 role physical (absolute)		1 (232)	1.3 (–8.8 to 11.4)	–	0.8	–
SF-36 role emotional (absolute)		1 (234)	2.7 (–6.4 to 11.8)	–	0.6	–
SF-36 mental health (absolute)		1 (235)	0.3 (–4.9 to 5.5)	–	0.9	–
SF-36 social function (absolute)		1 (235)	1.6 (–4.7 to 7.9)	–	0.6	–
SF-36 vitality (absolute)		1 (234)	0.0 (–6.3 to 6.3)	–	1.0	–
SF-36 pain (absolute)		1 (235)	2.6 (–4.6 to 9.8)	–	0.5	–
SF-36 general health (change)	6 months	1 (259)	–1.3 (–5.5 to 2.9)	–	0.5	–
SF-36 physical function (change)		1 (259)	3.0 (–2.6 to 8.6)	–	0.3	–
SF-36 role physical (change)		1 (264)	7.6 (–4.2 to 19.4)	–	0.2	–
SF-36 role emotional (change)		1 (264)	3.9 (–6.5 to 14.3)	–	0.5	–
SF-36 mental health (change)		1 (261)	–1.2 (–6.0 to 3.6)	–	0.6	–
SF-36 social function (change)		1 (230)	1.4 (–4.1 to 6.9)	–	0.6	–
SF-36 vitality (change)		1 (261)	2.8 (–3.1 to 8.7)	–	0.4	–
SF-36 pain (change)		1 (261)	4.5 (–3.3 to 12.3)	–	0.3	–
SF-36 general health (change)	12 months	2 (515)	–3.4 (–6.3 to –0.6)	–	0.02	0.6/0
SF-36 physical function (change)		2 (504)	1.0 (–2.6 to 4.6)	–	0.6	0.6/0
SF-36 role physical (change)		2 (512)	–7.0 (–15.2 to 1.2)	–	0.09	0.04/75
SF-36 role emotional (change)		2 (513)	–0.8 (–9.0 to 7.6)	–	0.9	0.3/0
SF-36 mental health (change)		2 (512)	–2.0 (–5.6 to 1.6)	–	0.3	0.6/0
SF-36 social function (change)		2 (478)	–2.1 (–6.1 to 1.9)	–	0.3	0.2/31
SF-36 vitality (change)		2 (511)	–1.2 (–5.3 to 2.9)	–	0.6	0.9/0
SF-36 pain (change)		2 (512)	–2.1 (–7.7 to 3.6)	–	0.5	0.03/78
SF-36 general health (change)	2 years	1 (249)	–1.9 (–7.4 to 3.6)	–	0.5	–
SF-36 physical function (change)		1 (244)	–1.3 (–6.5 to 3.9)	–	0.6	–
SF-36 role physical (change)		1 (249)	–12.3 (–24.5 to –0.1)	–	0.05	–
SF-36 role emotional (change)		1 (249)	–8.2 (–17.1 to 0.7)	–	0.07	–
SF-36 mental health (change)		1 (248)	–1.8 (–6.9 to 3.3)	–	0.5	–
SF-36 social function (change)		1 (248)	–4.2 (–10.6 to 2.2)	–	0.2	–
SF-36 vitality (change)		1 (248)	0.5 (–5.4 to 6.4)	–	0.9	–
SF-36 pain (change)		1 (249)	–10.8 (–18.6 to –3.0)	–	0.007	–

	Time point	Trials (no.)	WMD (95% CI)[a]	OR (95% CI)[b]	*p*-value	Hetero (p)/I^2 (%)
SF-36 general health (change)	5 years	1 (235)	1.3 (–4.4 to 7.0)	–	0.7	–
SF-36 physical function (change)		1 (228)	–0.7 (–6.5 to 5.1)	–	0.8	–
SF-36 role physical (change)		1 (232)	–6.5 (–19.0 to 6.0)	–	0.3	–
SF-36 role emotional (change)		1 (234)	1.7 (–9.7 to 13.1)	–	0.8	–
SF-36 mental health (change)		1 (234)	0.4 (–5.5 to 6.3)	–	0.9	–
SF-36 social function (change)		1 (234)	0.8 (–6.0 to 7.6)	–	0.8	–
SF-36 vitality (change)		1 (233)	1.9 (–5.0 to 8.8)	–	0.6	–
SF-36 pain (change)		1 (235)	–4.1 (–12.7 to 4.5)	–	0.4	–
SF-36 general health (change)	10 years	1 (189)	1.9 (–4.5 to 8.3)	–	0.6	–
SF-36 physical function (change)		1 (189)	1.4 (–6.0 to 8.8)	–	0.7	–
SF-36 role physical (change)		1 (189)	–4.1 (–18.4 to 10.2)	–	0.6	–
SF-36 role emotional (change)		1 (189)	–7.6 (–21.4 to 6.2)	–	0.3	–
SF-36 mental health (change)		1 (189)	0.7 (–5.9 to 7.3)	–	0.8	–
SF-36 social function (change)		1 (189)	–0.2 (–8.2 to 7.8)	–	1.0	–
SF-36 vitality (change)		1 (189)	2.4 (–5.6 to 10.4)	–	0.6	–
SF-36 pain (change)		1 (189)	0.7 (–9.6 to 11.0)	–	0.9	–
EQ-5D (absolute)	6 months	1 (68)	0.00 (–0.12 to 0.12)	–	1.0	–
	12 months	1 (61)	–0.03 (–0.15 to 0.09)	–	0.6	–
EQ-5D (change)	6 months	1 (66)	0.13 (–0.01 to 0.27)	–	0.08	–
	12 months	1 (60)	0.08 (–0.06 to 0.22)	–	0.3	–
Repeat EA	12 months	6 (1469)	–	0.71 (0.17 to 2.94)	0.6	0.4/0
	2 years	3 (677)	–	0.76 (0.16 to 3.63)	0.7	0.3/0
	3 years	1 (275)	–	5.11 (0.24 to 107)	0.3	–
	5 years	1 (263)	–	0.20 (0.01 to 4.30)	0.3	–
	10 years	1 (263)	–	0.34 (0.04 to 3.32)	0.4	
Hysterectomy after EA	6 months	1 (63)	–	0.56 (0.11 to 2.75)	0.5	–
	12 months	11 (2265)	–	0.77 (0.47 to 1.24)	0.3	1.0/0
	2 years	4 (939)	–	0.68 (0.41 to 1.13)	0.1	0.4/0
	3 years	1 (275)	–	0.48 (0.19 to 1.22)	0.1	–
	5 years	1 (266)	–	0.58 (0.31 to 1.06)	0.08	–
	10 years	1 (263)	–	0.52 (0.29 to 0.94)	0.03	–
		Trials	**Frequency**			
Repeat EA (overall)	12 months	6	8/1469 (< 1%)			
	2 years	3	7/677 (1%)			
	3 years	1	2/275 (1%)			
	5 years	1	2/263 (1%)			
	10 years	1	4/263 (2%)			
Hysterectomy after EA (overall)	6 months	1	7/63 (11%)			
	12 months	11	74/2265 (3%)			
	2 years	4	71/939 (8%)			
	3 years	1	21/275 (8%)			
	5 years	1	55/266 (21%)			
	10 years	1	60/263 (23%)			

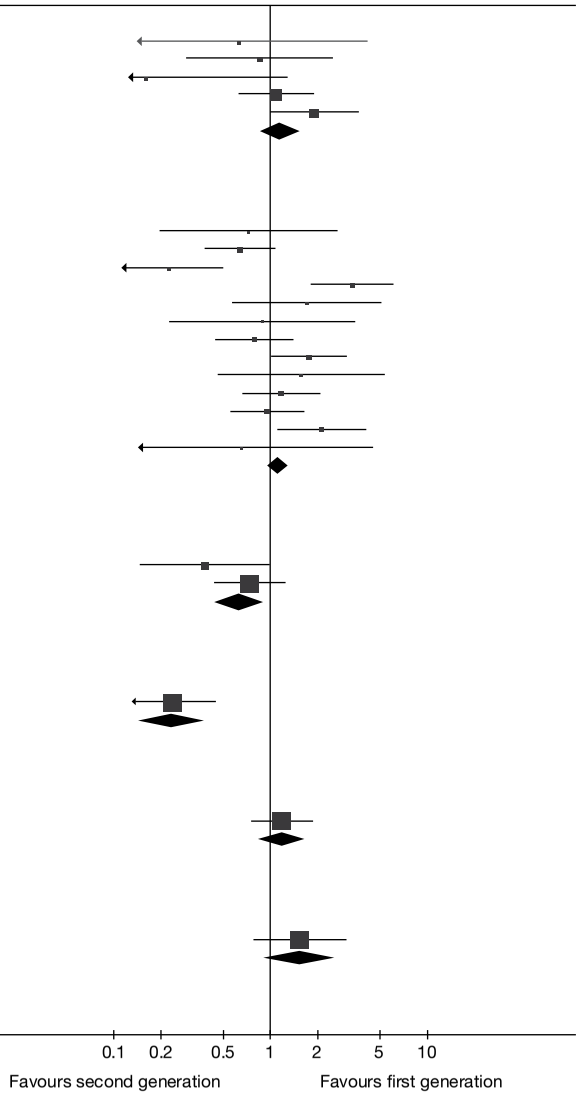
	Trials	Frequency (first-generation: max. 1017; second-generation: max. 1467)	OR (95% CI)[a]	*p*-value	Hetero (*p*)/I^2 (%)
Periprocedure complications					
Anaesthesia problems	14	0; 2	4.40 (0.23 to 85.1)	0.3	1.0/0
Excessive bleeding	14	8; 0	0.14 (0.03 to 0.55)	0.005	1.0/0
Uterine perforation	14	12; 3	0.20 (0.07 to 0.57)	0.003	0.3/12
Fluid overload	14	14; 0	0.12 (0.04 to 0.36)	0.0001	1.0/0
Visceral damage	14	0; 2	4.40 (0.23 to 85.8)	0.3	–
Cervical laceration	14	15; 2	0.12 (0.05 to 0.33)	<0.0001	0.9/0
Procedure abandoned	14	7; 16	1.58 (0.67 to 3.72)	0.3	0.3/14
Converted to hysterectomy	14	3; 1	0.38 (0.05 to 2.73)	0.3	0.3/1
Further complications (< 1 month)					
Urinary tract infection	14	12; 19	0.90 (0.42 to 1.90)	0.8	0.6/0
Deep-vein thrombosis	14	0; 0	–	–	–
Further bleeding	14	3; 5	1.17 (0.28 to 4.92)	0.8	0.07/57
Sepsis	14	0; 0	–	–	–
Pyrexia	14	1; 3	1.88 (0.25 to 14.2)	0.5	0.4/0
Endometriosis	14	9; 19	1.47 (0.68 to 3.18)	0.3	0.4/10
Haematomata	14	11; 5	0.26 (0.09 to 0.72)	0.01	0.8/0
Abdominal pain	14	34; 31	0.43 (0.26 to 0.74)	0.002	0.009/67
Foul discharge	14	1; 1	0.56 (0.03 to 9.94)	0.7	0.2/47
Visceral damage	14	0; 0	–	–	–

a <0 favours second-generation EA, >0 favours first-generation EA.
b <1 favours second-generation EA, >1 favours first-generation EA.

Appendix 6.2 First- versus second-generation endometrial ablation

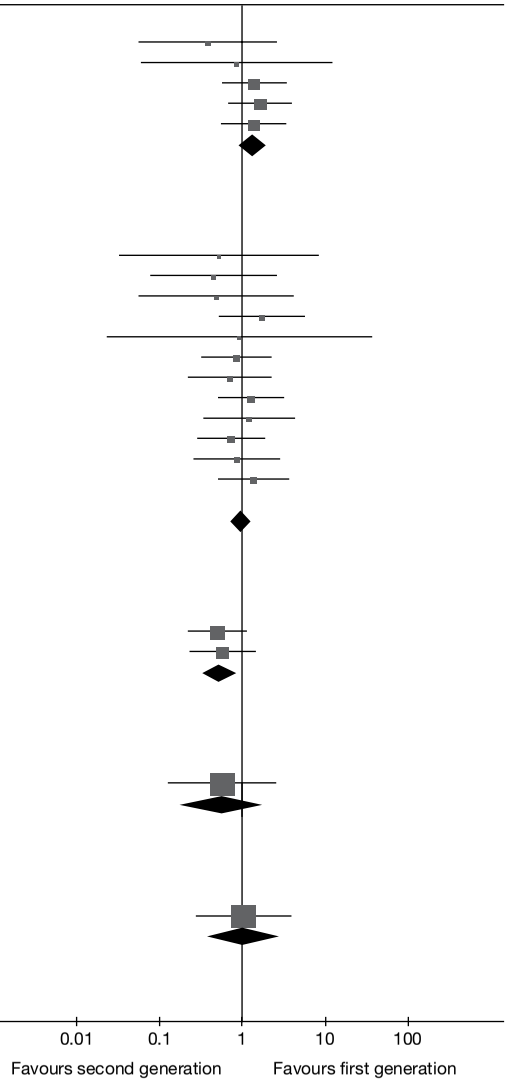
Proportion with amenorrhoea

Study or subgroup	First generation Events	First generation Total	Second generation Events	Second generation Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.1.1 <12 months						
a. Brun 2006[103]	2	15	4	20	4.2%	0.63 (0.06 to 6.33)
e. Hawe 2003[94]	12	27	15	31	12.2%	0.86 (0.22 to 3.29)
f. van Zon-Rabelink 2001[95]	0	58	5	76	4.0%	0.16 (0.02 to 1.72)
i. Corson 2001[79]	35	84	71	180	46.2%	1.10 (0.55 to 2.19)
m. Meyer 1998[53]	31	120	19	125	33.4%	1.92 (0.85 to 4.34)
Subtotal (95% CI)		**304**		**432**	**100.0%**	**1.16 (0.81 to 1.66)**
Total events	80		114			
Heterogeneity: $\chi^2 = 7.97$, df = 4 ($p = 0.09$); $I^2 = 50\%$						
Test for overall effect: $z = 0.82$ ($p = 0.41$)						
1.1.2 12 months						
a. Brun 2006[103]	5	17	11	30	2.3%	0.73 (0.14 to 3.73)
b. Cooper 2004[99]	49	96	120	194	14.2%	0.64 (0.33 to 1.23)
c. Perino 2004[100]	14	55	35	56	6.3%	0.23 (0.08 to 0.60)
d. Duleba 2003[98]	40	72	43	156	10.4%	3.33 (1.56 to 7.13)
e. Hawe 2003[94]	13	28	10	30	3.2%	1.71 (0.43 to 6.75)
f. van Zon-Rabelink 2001[95]	4	55	6	74	2.1%	0.89 (0.16 to 4.92)
g. Cooper 2002[96]	29	82	63	154	11.6%	0.79 (0.39 to 1.63)
i. Corson 2001[79]	43	83	65	172	12.5%	1.77 (0.88 to 3.55)
j. Soysal 2001[96]	8	48	5	45	2.6%	1.58 (0.34 to 7.32)
k. Corson 2000[101]	39	112	38	122	11.8%	1.18 (0.58 to 2.41)
l. Cooper 1999[54]	48	124	46	116	13.0%	0.96 (0.49 to 1.90)
m. Meyer 1998[53]	32	117	18	122	9.0%	2.13 (0.94 to 4.84)
n. Romer 1998[97]	3	10	4	10	1.1%	0.66 (0.06 to 6.93)
Subtotal (95% CI)		**899**		**1281**	**100.0%**	**1.12 (0.93 to 1.35)**
Total events	327		464			
Heterogeneity: $\chi^2 = 46.99$, df = 12 ($p < 0.00001$); $I^2 = 74\%$						
Test for overall effect: $z = 1.16$ ($p = 0.25$)						
1.1.3 2 years						
f. van Zon-Rabelink 2001[95]	6	55	17	66	23.2%	0.38 (0.12 to 1.27)
l. Cooper 1999[54]	52	128	58	121	76.8%	0.74 (0.39 to 1.44)
Subtotal (95% CI)		**183**		**187**	**100.0%**	**0.64 (0.41 to 0.99)**
Total events	58		75			
Heterogeneity: $\chi^2 = 1.56$, df = 1 ($p = 0.21$); $I^2 = 36\%$						
Test for overall effect: $z = 2.01$ ($p = 0.04$)						
1.1.4 3 years						
c. Perino 2004[100]	13	55	33	56	100.0%	0.24 (0.09 to 0.64)
Subtotal (95% CI)		**55**		**56**	**100.0%**	**0.24 (0.11 to 0.50)**
Total events	13		33			
Heterogeneity: not applicable						
Test for overall effect: $z = 3.76$ ($p = 0.0002$)						
1.1.5 5 years						
l. Cooper 1999[54]	82	119	76	117	100.0%	1.19 (0.59 to 2.43)
Subtotal (95% CI)		**119**		**117**	**100.0%**	**1.19 (0.70 to 2.05)**
Total events	82		76			
Heterogeneity: not applicable						
Test for overall effect: $z = 0.64$ ($p = 0.52$)						
1.1.6 10 years						
l. Cooper 1999[54]	84	95	78	94	100.0%	1.56 (0.53 to 4.53)
Subtotal (95% CI)		**95**		**94**	**100.0%**	**1.56 (0.69 to 3.51)**
Total events	84		78			
Heterogeneity: not applicable						
Test for overall effect: $z = 1.07$ ($p = 0.29$)						



Proportion with menorrhagia

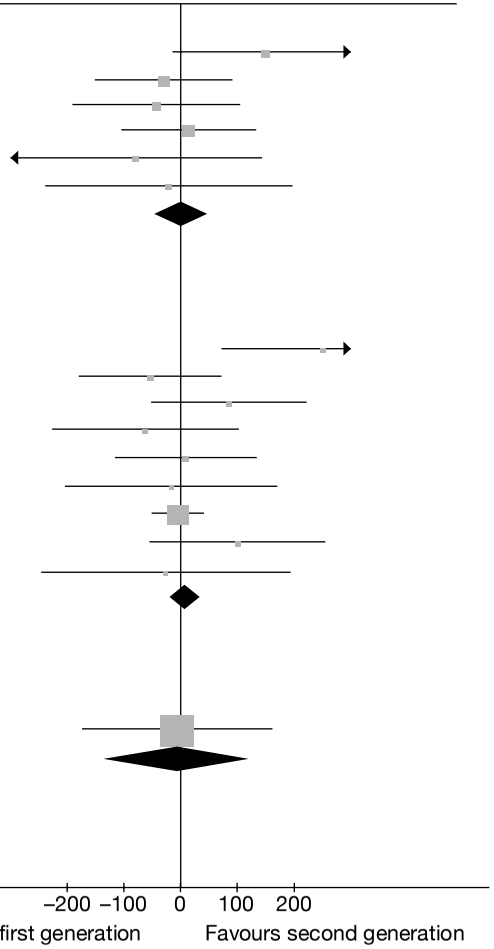
Study or subgroup	First generation Events	First generation Total	Second generation Events	Second generation Total	Weight	Peto odds ratio (non-event) Peto, Fixed, 99% CI
1.2.2 <12 months						
a. Brun 2006[103]	6	15	4	20	6.5%	0.39 (0.06 to 2.63)
e. Hawe 2003[94]	2	27	2	31	3.4%	0.86 (0.06 to 12.26)
f. van Zon-Rabelink 2001[95]	24	58	38	76	29.9%	1.41 (0.58 to 3.46)
i. Corson 2001[79]	11	84	37	180	31.0%	1.65 (0.68 to 3.98)
m. Meyer 1998[53]	16	120	22	125	29.2%	1.38 (0.56 to 3.43)
Subtotal (95% CI)		**304**		**432**	**100.0%**	**1.33 (0.92 to 1.93)**
Total events	59		103			
Heterogeneity: $\chi^2 = 3.37$, df = 4 ($p = 0.50$); $I^2 = 0\%$						
Test for overall effect: $z = 1.50$ ($p = 0.13$)						
1.2.3 12 months						
a. Brun 2006[103]	2	17	2	30	1.7%	0.53 (0.03 to 8.43)
b. Cooper 2004[99]	5	96	5	194	4.1%	0.45 (0.08 to 2.64)
c. Perino 2004[100]	4	55	2	56	2.7%	0.49 (0.06 to 4.20)
d. Duleba 2003[98]	5	72	19	156	8.9%	1.74 (0.53 to 5.74)
e. Hawe 2003[94]	1	28	1	30	0.9%	0.93 (0.02 to 36.86)
f. van Zon-Rabelink 2001[95]	19	55	23	74	13.4%	0.86 (0.32 to 2.27)
g. Cooper 2002[56]	10	82	14	154	9.4%	0.71 (0.22 to 2.28)
i. Corson 2001[79]	12	83	31	172	15.1%	1.29 (0.51 to 3.23)
j. Soysal 2001[96]	10	48	11	45	7.8%	1.23 (0.34 to 4.37)
k. Corson 2000[101]	19	112	16	122	14.3%	0.74 (0.29 to 1.90)
l. Cooper 1999[54]	11	124	9	116	8.8%	0.87 (0.26 to 2.88)
m. Meyer 1998[53]	13	117	18	122	12.9%	1.38 (0.51 to 3.71)
n. Romer 1998[97]	0	10	0	10		Not estimable
Subtotal (95% CI)		**899**		**1281**	**100.0%**	**0.97 (0.74 to 1.28)**
Total events	111		151			
Heterogeneity: $\chi^2 = 6.70$, df = 11 ($p = 0.82$); $I^2 = 0\%$						
Test for overall effect: $z = 0.19$ ($p = 0.85$)						
1.2.4 2 years						
f. van Zon-Rabelink 2001[95]	19	55	16	76	55.6%	0.50 (0.18 to 1.41)
l. Cooper 1999[54]	14	128	8	121	44.4%	0.59 (0.19 to 1.85)
Subtotal (95% CI)		**183**		**197**	**100.0%**	**0.54 (0.30 to 0.97)**
Total events	33		24			
Heterogeneity: $\chi^2 = 0.06$, df = 1 ($p = 0.80$); $I^2 = 0\%$						
Test for overall effect: $z = 2.08$ ($p = 0.04$)						
1.2.5 3 years						
c. Perino 2004[100]	5	55	3	56	100.0%	0.58 (0.09 to 3.78)
Subtotal (95% CI)		**55**		**56**	**100.0%**	**0.58 (0.14 to 2.41)**
Total events	5		3			
Heterogeneity: not applicable						
Test for overall effect: $z = 0.76$ ($p = 0.45$)						
1.2.6 5 years						
l. Cooper 1999[54]	5	136	5	130	100.0%	1.05 (0.20 to 5.50)
Subtotal (95% CI)		**136**		**130**	**100.0%**	**1.05 (0.30 to 3.70)**
Total events	5		5			
Heterogeneity: not applicable						
Test for overall effect: $z = 0.07$ ($p = 0.94$)						



DOI: 10.3310/hta15190

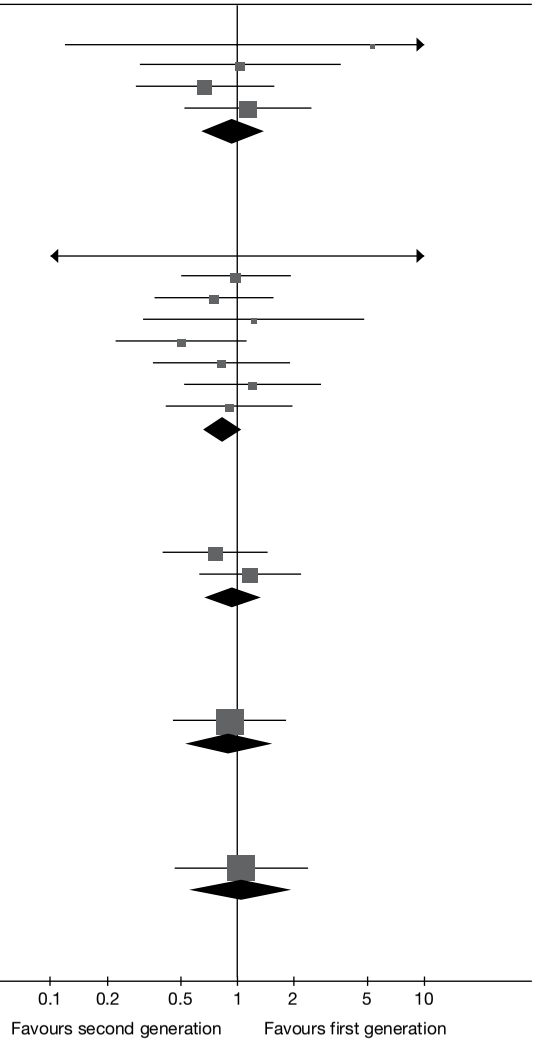
Bleeding/pictorial blood loss assessment chart scores (change from baseline)

Study or subgroup	First generation Mean	SD	Total	Second generation Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.8.2 <12 months								
a. Brun 2006[103]	–201	146	15	–351	231	20	14.0%	150.00 (–14.71 to 314.71)
e. Hawe 2003[94]	–361	211	27	–331	133	31	25.9%	–30.00 (–151.35 to 91.35)
f. van Zon-Rabelink 2001[95]	–379	331	58	–336	326	76	17.5%	–43.00 (–190.69 to 104.69)
g. Cooper 2002[56]	–520	356	90	–534	356	175	26.9%	14.00 (–104.95 to 132.95)
i. Corson 2001[79]	–544	641	84	–464	687	180	7.6%	–80.00 (–303.27 to 143.27)
m. Meyer 1998[53]	–519	592	120	–498	728	125	8.0%	–21.00 (–238.96 to 196.96)
Subtotal (95% CI)			**394**			**607**	**100.0%**	**1.76 (–45.21 to 48.74)**
Heterogeneity: $\chi^2 = 7.47$, df = 5 ($p = 0.19$); $I^2 = 33\%$								
Test for overall effect: $z = 0.07$ ($p = 0.94$)								
1.8.3 12 months								
a. Brun 2006[103]	–221	164	17	–471	305	29	3.9%	250.00 (71.73 to 428.27)
b. Cooper 2004[99]	–498	433	106	–444	359	209	7.8%	–54.00 (–179.81 to 71.81)
d. Duleba 2003[98]	–441	348	72	–526	423	156	6.6%	85.00 (–52.00 to 222.00)
f. van Zon-Rabelink 2001[95]	–406	367	55	–344	349	74	4.5%	–62.00 (–226.83 to 102.83)
g. Cooper 2002[56]	–526	353	82	–535	356	154	7.9%	9.00 (–115.67 to 133.67)
i. Corson 2001[79]	–511	458	83	–494	689	172	3.5%	–17.00 (–204.30 to 170.30)
j. Soysal 2001[96]	–347	88	54	–342	86	42	58.1%	–5.00 (–51.04 to 41.04)
k. Corson 2000[101]	–417	285	112	–517	594	122	5.1%	100.00 (–54.92 to 254.92)
m. Meyer 1998[53]	–536	585	117	–510	730	122	2.5%	–26.00 (–245.97 to 193.97)
Subtotal (95% CI)			**698**			**1080**	**100.0%**	**9.94 (–16.77 to 36.65)**
Heterogeneity: $\chi^2 = 20.26$, df = 8 ($p = 0.009$); $I^2 = 61\%$								
Test for overall effect: $z = 0.73$ ($p = 0.47$)								
1.8.4 2 years								
f. van Zon-Rabelink 2001[95]	–404	392	55	–398	310	66	100.0%	–6.00 (–173.92 to 161.92)
Subtotal (95% CI)			**55**			**66**	**100.0%**	**–6.00 (–133.77 to 121.77)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.09$ ($p = 0.93$)								



Proportion with dysmenorrhoea

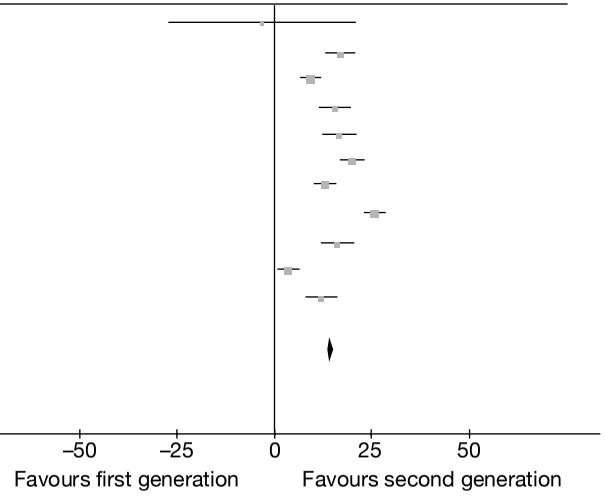
Study or subgroup	First generation Events	First generation Total	Second generation Events	Second generation Total	Weight	Peto odds ratio (non-event) Peto, Fixed, 99% CI
1.9.2 <12 months						
a. Brun 2006[103]	0	15	2	24	1.9%	5.31 (0.12 to 235.80)
e. Hawe 2003[94]	17	32	20	37	17.5%	1.04 (0.30 to 3.58)
i. Corson 2001[79]	22	64	38	145	36.9%	0.67 (0.29 to 1.58)
m. Meyer 1998[53]	26	120	30	125	43.8%	1.14 (0.52 to 2.49)
Subtotal (95% CI)		**231**		**331**	**100.0%**	**0.95 (0.64 to 1.41)**
Total events	65		90			
Heterogeneity: $\chi^2 = 2.85$, df = 3 ($p = 0.41$); $I^2 = 0\%$						
Test for overall effect: $z = 0.26$ ($p = 0.80$)						
1.9.4 12 months						
a. Brun 2006[103]	1	13	0	22	0.3%	0.07 (0.00 to 13.99)
b. Cooper 2004[99]	33	97	66	196	20.7%	0.98 (0.50 to 1.93)
d. Duleba 2003[98]	37	72	69	156	17.5%	0.75 (0.36 to 1.56)
e. Hawe 2003[94]	15	28	17	29	5.1%	1.22 (0.31 to 4.78)
g. Cooper 2002[56]	28	82	32	153	14.5%	0.50 (0.22 to 1.12)
i. Corson 2001[79]	20	67	40	154	13.2%	0.82 (0.35 to 1.92)
l. Cooper 1999[54]	22	124	24	116	13.3%	1.21 (0.52 to 2.81)
m. Meyer 1998[53]	29	117	28	122	15.5%	0.90 (0.41 to 1.97)
Subtotal (95% CI)		**600**		**948**	**100.0%**	**0.84 (0.67 to 1.07)**
Total events	185		276			
Heterogeneity: $\chi^2 = 6.54$, df = 7 ($p = 0.48$); $I^2 = 0\%$						
Test for overall effect: $z = 1.41$ ($p = 0.16$)						
1.9.5 2 years						
l. Cooper 1999[54]	29	128	22	121	48.5%	0.76 (0.34 to 1.71)
m. Meyer 1998[53]	26	108	32	118	51.5%	1.17 (0.54 to 2.57)
Subtotal (95% CI)		**236**		**239**	**100.0%**	**0.95 (0.62 to 1.46)**
Total events	55		54			
Heterogeneity: $\chi^2 = 0.98$, df = 1 ($p = 0.32$); $I^2 = 0\%$						
Test for overall effect: $z = 0.23$ ($p = 0.82$)						
1.9.6 3 years						
m. Meyer 1998[53]	22	102	22	110	100.0%	0.91 (0.38 to 2.17)
Subtotal (95% CI)		**102**		**110**	**100.0%**	**0.91 (0.47 to 1.76)**
Total events	22		22			
Heterogeneity: not applicable						
Test for overall effect: $z = 0.28$ ($p = 0.78$)						
1.9.7 5 years						
l. Cooper 1999[54]	14	136	14	130	100.0%	1.05 (0.38 to 2.94)
Subtotal (95% CI)		**136**		**130**	**100.0%**	**1.05 (0.48 to 2.30)**
Total events	14		14			
Heterogeneity: not applicable						
Test for overall effect: $z = 0.13$ ($p = 0.90$)						



Duration of surgery (minutes)

Study or subgroup	First generation Mean	SD	Total	Second generation Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
a. Brun 2006[103]	52.8	27.4	16	56	31.2	24	0.2%	–3.20 (–27.29 to 20.89)
b. Cooper 2004[99]	20.3	15.6	106	3.4	1	209	7.3%	16.90 (12.99 to 20.81)
c. Perino 2004[100]	16.4	7.8	55	7.1	0.1	56	15.2%	9.30 (6.59 to 12.01)
e. Hawe 2003[94]	30.5	9.2	33	15	0.1	37	6.6%	15.50 (11.37 to 19.63)
f. van Zon-Rabelink 2001[95]	34.7	12.7	62	18	4.9	77	5.8%	16.70 (12.30 to 21.10)
g. Cooper 2002[56]	24.2	11.4	90	4.2	3.5	175	11.1%	20.00 (16.83 to 23.17)
h. Pellicano 2002[102]	37	6	42	24	4	40	13.4%	13.00 (10.11 to 15.89)
j. Soysal 2001[96]	37.3	7.5	48	11.5	0.8	45	14.2%	25.80 (22.99 to 28.61)
k. Corson 2000[101]	39.3	16.6	126	23.1	9.5	150	6.0%	16.20 (11.90 to 20.50)
l. Cooper 1999[54]	15	7.2	134	11.4	10.5	129	13.6%	3.60 (0.73 to 6.47)
m. Meyer 1998[53]	39.2	14.8	129	27.1	10.4	128	6.6%	12.10 (7.99 to 16.21)
Total (95% CI)			**841**			**1070**	**100.0%**	**14.52 (13.72 to 15.33)**

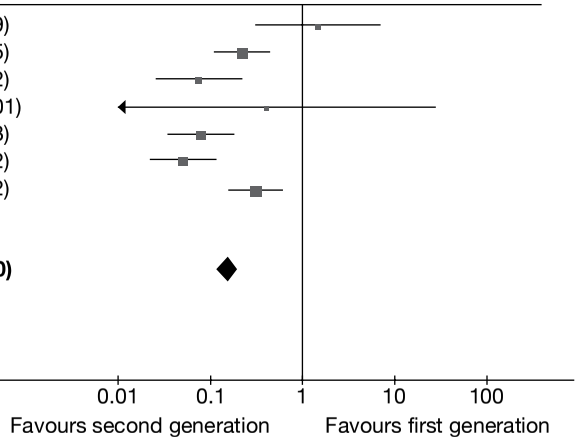
Heterogeneity: $\chi^2 = 261.00$, df = 10 ($p < 0.00001$); $I^2 = 96\%$
Test for overall effect: $z = 35.39$ ($p < 0.00001$)



Use of general anaesthesia

Study or subgroup	First generation Events	First generation Total	Second generation Events	Second generation Total	Weight	Peto odds ratio (non-event) Peto, Fixed, 99% CI
a. Brun 2006[103]	12	20	20	29	4.9%	1.48 (0.31 to 7.09)
b. Cooper 2004[99]	83	106	93	209	24.2%	0.22 (0.11 to 0.45)
d. Duleba 2003[98]	79	86	89	193	10.3%	0.08 (0.03 to 0.22)
f. van Zon-Rabelink 2001[95]	62	62	76	77	0.7%	0.41 (0.01 to 28.01)
g. Cooper 2002[56]	74	90	47	175	17.2%	0.08 (0.03 to 0.18)
k. Corson 2000[101]	98	123	22	132	17.3%	0.05 (0.02 to 0.12)
m. Meyer 1998[53]	107	138	71	137	25.5%	0.31 (0.16 to 0.62)
n. Romer 1998[97]	10	10	10	10		Not estimable
Total (95% CI)		**635**		**962**	**100.0%**	**0.16 (0.12 to 0.20)**
Total events	525		428			

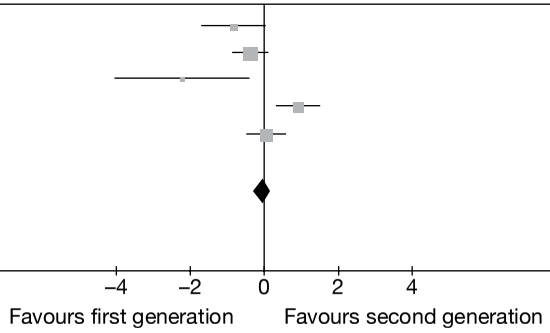
Heterogeneity: $\chi^2 = 41.70$, df = 6 ($p < 0.00001$); $I^2 = 86\%$
Test for overall effect: $z = 13.85$ ($p < 0.00001$)



Surgery pain score

Study or subgroup	First generation Mean	First generation SD	First generation Total	Second generation Mean	Second generation SD	Second generation Total	Weight	Standard mean difference IV, Fixed, 99% CI
a. Brun 2006[103]	2.3	2.9	15	4.7	2.8	26	11.0%	−0.83 (−1.70 to 0.04)
c. Perino 2004[100]	3.7	1.4	55	4.4	2.2	56	34.2%	−0.38 (−0.87 to 0.12)
e. Hawe 2003[94]	2.4	1.2	6	6.5	2	9	2.5%	−2.22 (−4.05 to −0.40)
h. Pellicano 2002[102]	3.8	0.6	42	3.2	0.7	40	23.2%	0.91 (0.31 to 1.51)
j. Soysal 2001[96]	3.2	2.1	48	3.1	1.7	45	29.1%	0.05 (−0.48 to 0.59)
Total (95% CI)			**166**			**176**	**100.0%**	**−0.05 (−0.27 to 0.17)**

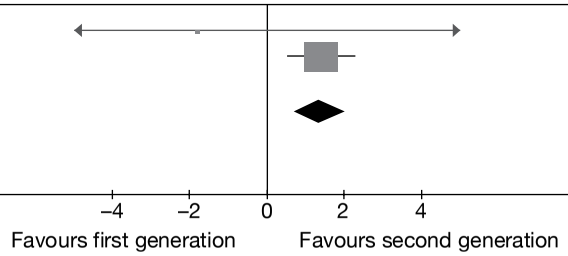
Heterogeneity: $\chi^2 = 35.01$, df = 4 ($p < 0.00001$); $I^2 = 89\%$
Test for overall effect: $z = 0.43$ ($p = 0.66$)



Time to return to work (days)

	First generation			Second generation				Mean difference	Mean difference
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 99% CI**	**IV, Fixed, 99% CI**
a. Brun 2006[103]	6.4	8.2	13	8.2	9.1	21	1.3%	–1.80 (–9.58 to 5.98)	
h. Pellicano 2002[102]	6.3	2.1	42	4.9	0.7	40	98.7%	1.40 (0.52 to 2.28)	
Total (95% CI)			**55**			**61**	**100.0%**	**1.36 (0.69 to 2.03)**	

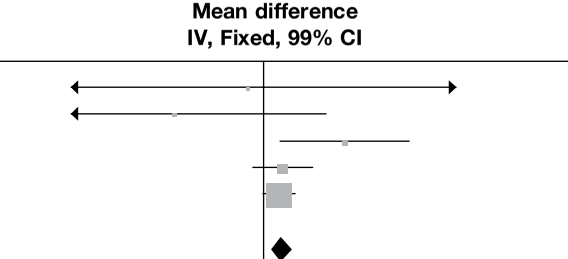
Heterogeneity: $\chi^2 = 1.11$, df = 1 ($p = 0.29$); $I^2 = 10\%$
Test for overall effect: $z = 4.00$ ($p < 0.0001$)



Time to return to normal activities (days)

	First generation			Second generation				Mean difference	Mean difference
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 99% CI**	**IV, Fixed, 99% CI**
a. Brun 2006[103]	5.6	8.1	16	6	5.6	23	0.4%	–0.40 (–6.42 to 5.62)	
b. Cooper 2004[99]	11	8.5	92	13.3	17.1	189	0.9%	–2.30 (–6.23 to 1.63)	
h. Pellicano 2002[102]	6.2	3.3	42	4.1	2.6	40	4.7%	2.10 (0.41 to 3.79)	
k. Corson 2000[101]	2.9	3	126	2.4	1.8	150	21.6%	0.50 (–0.29 to 1.29)	
m. Meyer 1998[53]	1.8	1.5	109	1.4	0.9	114	72.5%	0.40 (–0.03 to 0.83)	
Total (95% CI)			**385**			**516**	**100.0%**	**0.48 (0.20 to 0.75)**	

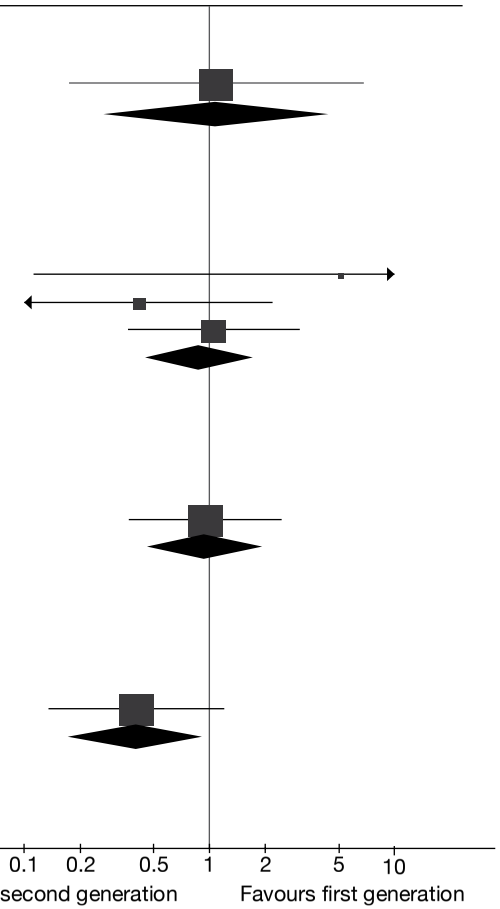
Heterogeneity: $\chi^2 = 9.82$, df = 4 ($p = 0.04$); $I^2 = 59\%$
Test for overall effect: $z = 3.35$ ($p = 0.0008$)



DOI: 10.3310/hta15190

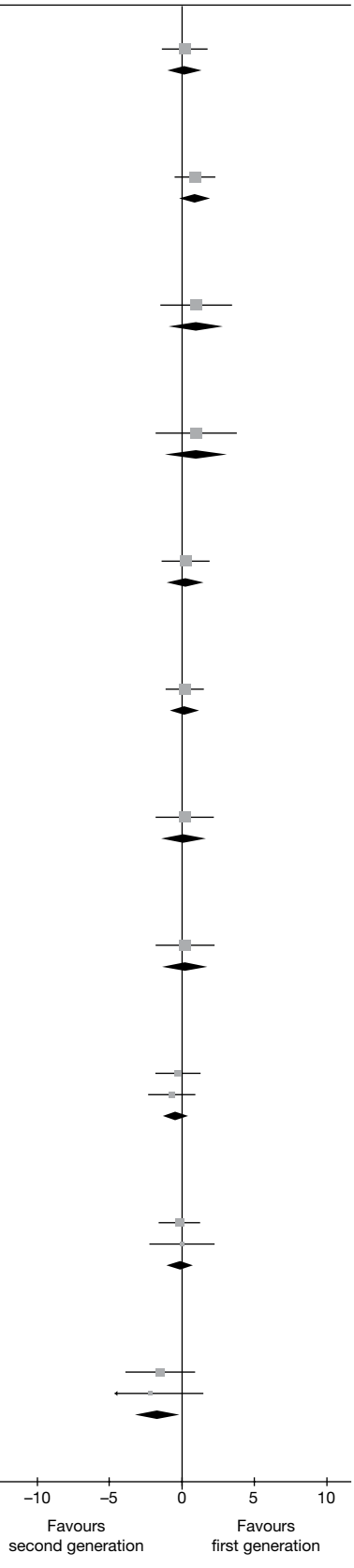
Proportion with dyspareunia

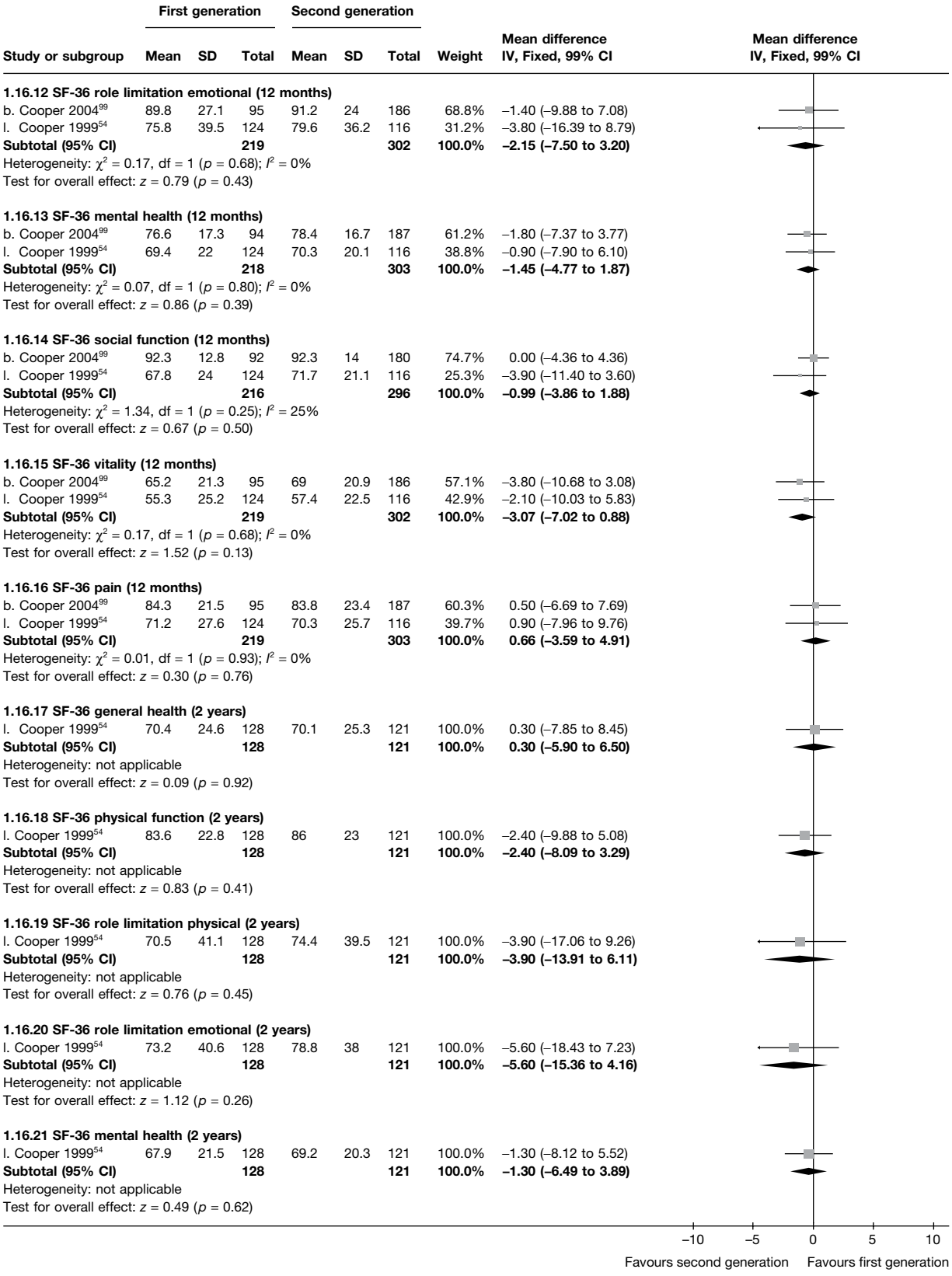
Study or subgroup	First generation Events	First generation Total	Second generation Events	Second generation Total	Weight	Peto odds ratio (non-event) Peto, Fixed, 99% CI
1.15.2 <12 months						
a. Brun 2006[103]	0	15	0	22		Not estimable
e. Hawe 2003[94]	4	32	5	37	100.0%	1.09 (0.17 to 6.83)
Subtotal (95% CI)		**47**		**59**	**100.0%**	**1.09 (0.27 to 4.41)**
Total events	4		5			
Heterogeneity: not applicable						
Test for overall effect: $z = 0.12$ ($p = 0.90$)						
1.15.3 12 months						
a. Brun 2006[103]	0	13	2	22	5.2%	5.15 (0.11 to 236.36)
e. Hawe 2003[94]	8	28	4	29	27.8%	0.42 (0.08 to 2.19)
I. Cooper 1999[54]	13	122	13	116	67.0%	1.06 (0.36 to 3.08)
Subtotal (95% CI)		**163**		**167**	**100.0%**	**0.89 (0.46 to 1.73)**
Total events	21		19			
Heterogeneity: $\chi^2 = 2.95$, df = 2 ($p = 0.23$); $I^2 = 32\%$						
Test for overall effect: $z = 0.35$ ($p = 0.73$)						
1.15.4 2 years						
I. Cooper 1999[54]	18	128	16	119	100.0%	0.95 (0.37 to 2.46)
Subtotal (95% CI)		**128**		**119**	**100.0%**	**0.95 (0.46 to 1.96)**
Total events	18		16			
Heterogeneity: not applicable						
Test for overall effect: $z = 0.14$ ($p = 0.89$)						
1.15.5 5 years						
I. Cooper 1999[54]	18	113	7	105	100.0%	0.40 (0.14 to 1.20)
Subtotal (95% CI)		**113**		**105**	**100.0%**	**0.40 (0.18 to 0.93)**
Total events	18		7			
Heterogeneity: not applicable						
Test for overall effect: $z = 2.14$ ($p = 0.03$)						



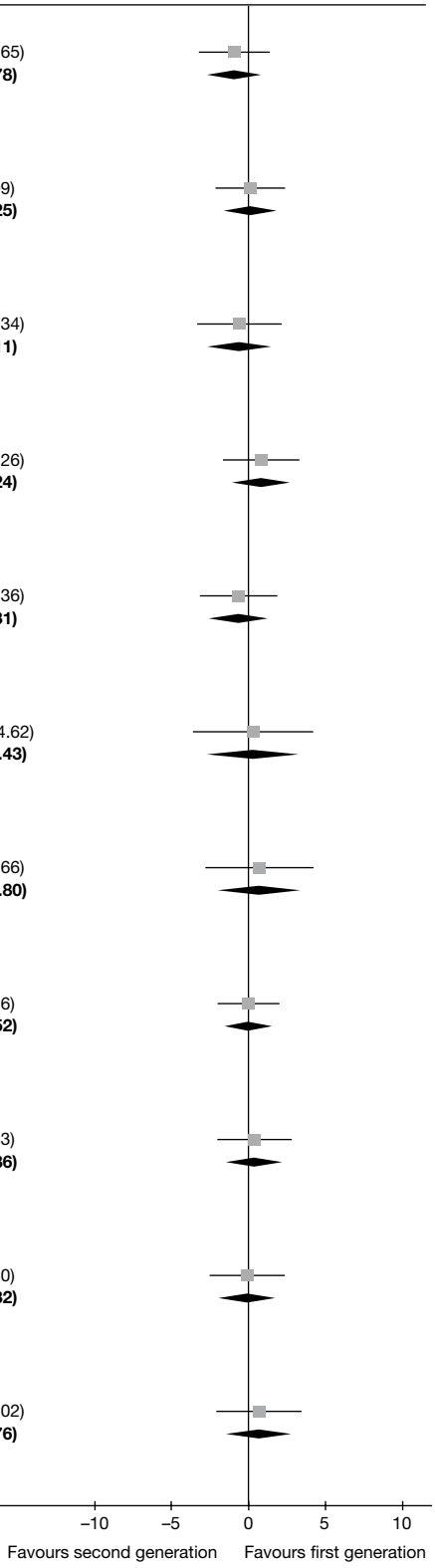
SF-36 scores (absolute values)

Study or subgroup	First generation Mean	SD	Total	Second generation Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.16.1 SF-36 general health (6 months)								
b. Cooper 2004[99]	82.1	15.3	87	81.5	17.3	178	100.0%	0.60 (–4.79 to 5.99)
Subtotal (95% CI)			**87**			**178**	**100.0%**	**0.60 (–3.50 to 4.70)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.29$ ($p = 0.77$)								
1.16.2 SF-36 physical function (6 months)								
b. Cooper 2004[99]	93.7	11.7	87	90.7	18.3	180	100.0%	3.00 (–1.77 to 7.77)
Subtotal (95% CI)			**87**			**180**	**100.0%**	**3.00 (–0.63 to 6.63)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.62$ ($p = 0.11$)								
1.16.3 SF-36 role limitation physical (6 months)								
b. Cooper 2004[99]	91.8	23.6	91	88.5	28.9	182	100.0%	3.30 (–5.13 to 11.73)
Subtotal (95% CI)			**91**			**182**	**100.0%**	**3.30 (–3.11 to 9.71)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.01$ ($p = 0.31$)								
1.16.4 SF-36 role limitation emotional (6 months)								
b. Cooper 2004[99]	89.4	27.6	91	86.1	31.1	180	100.0%	3.30 (–6.25 to 12.85)
Subtotal (95% CI)			**91**			**180**	**100.0%**	**3.30 (–3.97 to 10.57)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.89$ ($p = 0.37$)								
1.16.5 SF-36 mental health (6 months)								
b. Cooper 2004[99]	77.9	16.6	90	77.1	17.7	179	100.0%	0.80 (–4.85 to 6.45)
Subtotal (95% CI)			**90**			**179**	**100.0%**	**0.80 (–3.50 to 5.10)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.36$ ($p = 0.72$)								
1.16.6 SF-36 social function (6 months)								
b. Cooper 2004[99]	92.5	12.1	89	91.9	15.3	168	100.0%	0.60 (–3.89 to 5.09)
Subtotal (95% CI)			**89**			**168**	**100.0%**	**0.60 (–2.82 to 4.02)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.34$ ($p = 0.73$)								
1.16.7 SF-36 vitality (6 months)								
b. Cooper 2004[99]	66.5	19.7	92	65.9	22.5	177	100.0%	0.60 (–6.25 to 7.45)
Subtotal (95% CI)			**92**			**177**	**100.0%**	**0.60 (–4.61 to 5.81)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.23$ ($p = 0.82$)								
1.16.8 SF-36 pain (6 months)								
b. Cooper 2004[99]	83.8	19.7	92	83.1	23.1	177	100.0%	0.70 (–6.23 to 7.63)
Subtotal (95% CI)			**92**			**177**	**100.0%**	**0.70 (–4.57 to 5.97)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.26$ ($p = 0.79$)								
1.16.9 SF-36 general health (12 months)								
b. Cooper 2004[99]	81.5	16.3	95	82.5	16.7	187	51.9%	–1.00 (–6.33 to 4.33)
l.Cooper 1999[54]	65.2	16.4	124	67.3	16.9	116	48.1%	–2.10 (–7.64 to 3.44)
Subtotal (95% CI)			**219**			**303**	**100.0%**	**–1.53 (–4.45 to 1.40)**
Heterogeneity: $\chi^2 = 0.14$, df = 1 ($p = 0.71$); $I^2 = 0\%$								
Test for overall effect: $z = 1.02$ ($p = 0.31$)								
1.16.10 SF-36 physical function (12 months)								
b. Cooper 2004[99]	92.3	15.3	94	92.9	14.4	185	71.2%	–0.60 (–5.49 to 4.29)
l. Cooper 1999[54]	84.9	22.5	124	84.9	23.7	116	28.8%	0.00 (–7.70 to 7.70)
Subtotal (95% CI)			**218**			**301**	**100.0%**	**–0.43 (–3.57 to 2.72)**
Heterogeneity: $\chi^2 = 0.03$, df = 1 ($p = 0.87$); $I^2 = 0\%$								
Test for overall effect: $z = 0.27$ ($p = 0.79$)								
1.16.11 SF-36 role limitation physical (12 months)								
b. Cooper 2004[99]	88.9	27.2	95	94	20.9	187	70.0%	–5.10 (–13.30 to 3.10)
l. Cooper 1999[54]	73.8	40.3	124	81.3	34.9	116	30.0%	–7.50 (–20.01 to 5.01)
Subtotal (95% CI)			**219**			**303**	**100.0%**	**–5.82 (–11.04 to –0.60)**
Heterogeneity: $\chi^2 = 0.17$, df = 1 ($p = 0.68$); $I^2 = 0\%$								
Test for overall effect: $z = 2.19$ ($p = 0.03$)								



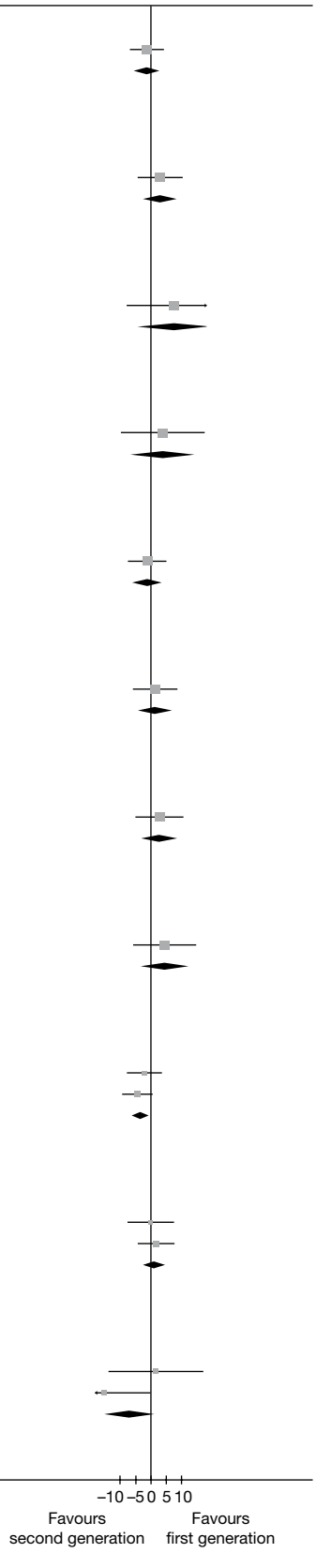


Study or subgroup	First generation Mean	SD	Total	Second generation Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.16.22 SF-36 social function (2 years)								
I. Cooper 1999[54]	66.8	25.1	128	70	23	121	100.0%	–3.20 (–11.05 to 4.65)
Subtotal (95% CI)			**128**			**121**	**100.0%**	**–3.20 (–9.18 to 2.78)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.05$ ($p = 0.29$)								
1.16.23 SF-36 vitality (2 years)								
I. Cooper 1999[54]	55.5	24	128	55.1	23.1	121	100.0%	0.40 (–7.29 to 8.09)
Subtotal (95% CI)			**128**			**121**	**100.0%**	**0.40 (–5.45 to 6.25)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.13$ ($p = 0.89$)								
1.16.24 SF-36 pain (2 years)								
I. Cooper 1999[54]	67.1	28.6	128	69.1	28.6	121	100.0%	–2.00 (–11.34 to 7.34)
Subtotal (95% CI)			**128**			**121**	**100.0%**	**–2.00 (–9.11 to 5.11)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.55$ ($p = 0.58$)								
1.16.25 SF-36 general health (5 years)								
I. Cooper 1999[54]	70.3	22.6	118	67.5	27.5	117	100.0%	2.80 (–5.66 to 11.26)
Subtotal (95% CI)			**118**			**117**	**100.0%**	**2.80 (–3.64 to 9.24)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.85$ ($p = 0.39$)								
1.16.26 SF-36 physical function (5 years)								
I. Cooper 1999[54]	81.3	24.9	118	83.5	25.7	114	100.0%	–2.20 (–10.76 to 6.36)
Subtotal (95% CI)			**118**			**114**	**100.0%**	**–2.20 (–8.71 to 4.31)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.66$ ($p = 0.51$)								
1.16.27 SF-36 role limitation physical (5 years)								
I. Cooper 1999[54]	75.2	38.5	118	73.9	40.2	114	100.0%	1.30 (–12.02 to 14.62)
Subtotal (95% CI)			**118**			**114**	**100.0%**	**1.30 (–8.83 to 11.43)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.25$ ($p = 0.80$)								
1.16.28 SF-36 role limitation emotional (5 years)								
I. Cooper 1999[54]	81.1	34.4	118	78.4	36.6	116	100.0%	2.70 (–9.26 to 14.66)
Subtotal (95% CI)			**118**			**116**	**100.0%**	**2.70 (–6.40 to 11.80)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.58$ ($p = 0.56$)								
1.16.29 SF-36 mental health (5 years)								
I. Cooper 1999[54]	71.3	20.8	118	71	20	117	100.0%	0.30 (–6.56 to 7.16)
Subtotal (95% CI)			**118**			**117**	**100.0%**	**0.30 (–4.92 to 5.52)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.11$ ($p = 0.91$)								
1.16.30 SF-36 social function (5 years)								
I. Cooper 1999[54]	69.9	23.5	119	68.3	25.4	116	100.0%	1.60 (–6.63 to 9.83)
Subtotal (95% CI)			**119**			**116**	**100.0%**	**1.60 (–4.66 to 7.86)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.50$ ($p = 0.62$)								
1.16.31 SF-36 vitality (5 years)								
I. Cooper 1999[54]	56.5	24.8	118	56.5	24.5	116	100.0%	0.00 (–8.30 to 8.30)
Subtotal (95% CI)			**118**			**116**	**100.0%**	**0.00 (–6.32 to 6.32)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.00$ ($p = 1.00$)								
1.16.32 SF-36 pain (5 years)								
I. Cooper 1999[54]	72.2	27.4	119	69.6	28.6	116	100.0%	2.60 (–6.82 to 12.02)
Subtotal (95% CI)			**119**			**116**	**100.0%**	**2.60 (–4.56 to 9.76)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.71$ ($p = 0.48$)								

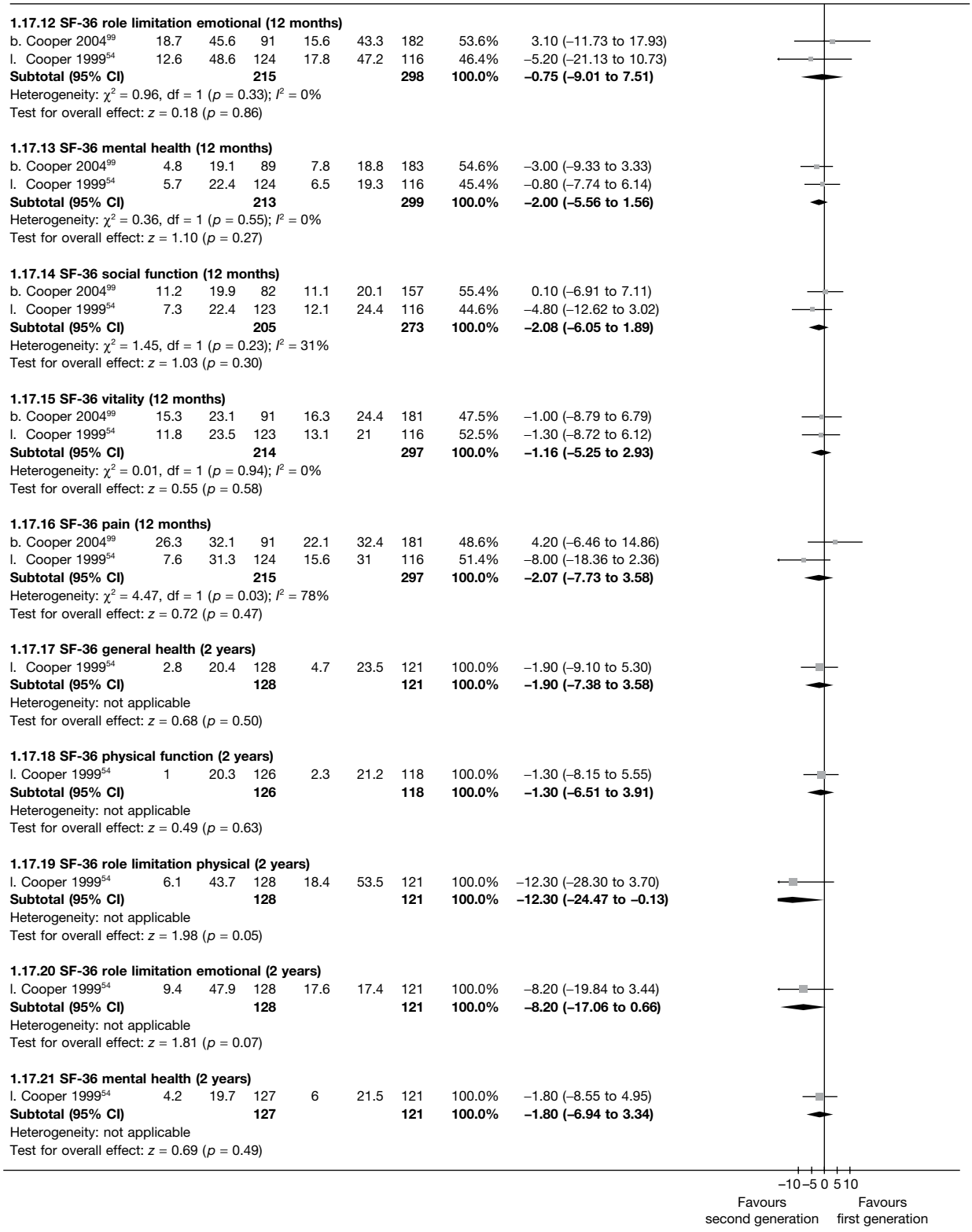


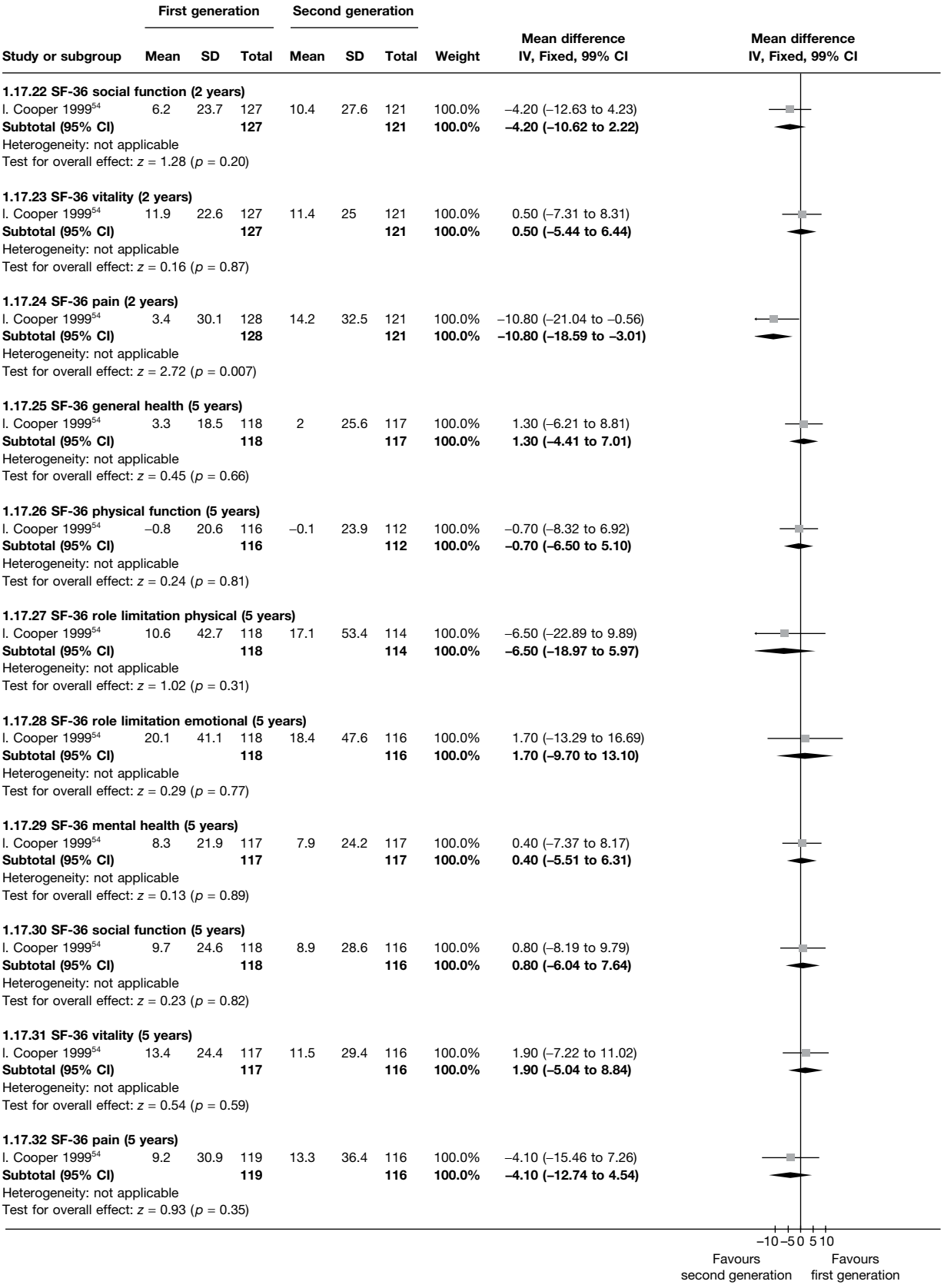
SF-36 scores (change from baseline)

Study or subgroup	First generation Mean	SD	Total	Second generation Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.17.1 SF-36 general health (6 months)								
b. Cooper 2004[99]	6.3	15.7	84	7.6	17.1	175	100.0%	–1.30 (–6.83 to 4.23)
Subtotal (95% CI)			**84**			**175**	**100.0%**	**–1.30 (–5.51 to 2.91)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.61$ ($p = 0.54$)								
1.17.2 SF-36 physical function (6 months)								
b. Cooper 2004[99]	8.9	19.8	84	5.9	24.6	175	100.0%	3.00 (–4.34 to 10.34)
Subtotal (95% CI)			**84**			**175**	**100.0%**	**3.00 (–2.59 to 8.59)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.05$ ($p = 0.29$)								
1.17.3 SF-36 role limitation physical (6 months)								
b. Cooper 2004[99]	28.5	45.4	86	20.9	46.9	178	100.0%	7.60 (–7.92 to 23.12)
Subtotal (95% CI)			**86**			**178**	**100.0%**	**7.60 (–4.21 to 19.41)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.26$ ($p = 0.21$)								
1.17.4 SF-36 role limitation emotional (6 months)								
b. Cooper 2004[99]	16.1	38.7	87	12.2	44	177	100.0%	3.90 (–9.77 to 17.57)
Subtotal (95% CI)			**87**			**177**	**100.0%**	**3.90 (–6.50 to 14.30)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.74$ ($p = 0.46$)								
1.17.5 SF-36 mental health (6 months)								
b. Cooper 2004[99]	5.3	18.5	86	6.5	18.6	175	100.0%	–1.20 (–7.49 to 5.09)
Subtotal (95% CI)			**86**			**175**	**100.0%**	**–1.20 (–5.98 to 3.58)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.49$ ($p = 0.62$)								
1.17.6 SF-36 social function (6 months)								
b. Cooper 2004[99]	11.8	20.7	81	10.4	19.7	149	100.0%	1.40 (–5.84 to 8.64)
Subtotal (95% CI)			**81**			**149**	**100.0%**	**1.40 (–4.11 to 6.91)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.50$ ($p = 0.62$)								
1.17.7 SF-36 vitality (6 months)								
b. Cooper 2004[99]	15.5	22.4	88	12.7	24.6	173	100.0%	2.80 (–5.01 to 10.61)
Subtotal (95% CI)			**88**			**173**	**100.0%**	**2.80 (–3.14 to 8.74)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.92$ ($p = 0.36$)								
1.17.8 SF-36 pain (6 months)								
b. Cooper 2004[99]	25.6	31.3	88	21.1	28.7	173	100.0%	4.50 (–5.77 to 14.77)
Subtotal (95% CI)			**88**			**173**	**100.0%**	**4.50 (–3.31 to 12.31)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.13$ ($p = 0.26$)								
1.17.9 SF-36 general health (12 months)								
b. Cooper 2004[99]	6.4	17.2	92	8.5	17.6	183	43.3%	–2.10 (–7.81 to 3.61)
l. Cooper 1999[54]	–2.2	14.3	124	2.2	15.6	116	56.7%	–4.40 (–9.39 to 0.59)
Subtotal (95% CI)			**216**			**299**	**100.0%**	**–3.40 (–6.26 to –0.55)**
Heterogeneity: $\chi^2 = 0.61$, df = 1 ($p = 0.43$); $I^2 = 0\%$								
Test for overall effect: $z = 2.34$ ($p = 0.02$)								
1.17.10 SF-36 physical function (12 months)								
b. Cooper 2004[99]	7.9	23.5	89	8	21	180	38.6%	–0.10 (–7.68 to 7.48)
l. Cooper 1999[54]	2.4	16.7	122	0.7	18.9	113	61.4%	1.70 (–4.31 to 7.71)
Subtotal (95% CI)			**211**			**293**	**100.0%**	**1.00 (–2.58 to 4.59)**
Heterogeneity: $\chi^2 = 0.23$, df = 1 ($p = 0.63$); $I^2 = 0\%$								
Test for overall effect: $z = 0.55$ ($p = 0.58$)								
1.17.11 SF-36 role limitation physical (12 months)								
b. Cooper 2004[99]	27.8	47.7	89	26.2	43.5	183	49.0%	1.60 (–13.83 to 17.03)
l. Cooper 1999[54]	9.9	41.9	124	25.2	48.5	116	51.0%	–15.30 (–30.42 to –0.18)
Subtotal (95% CI)			**213**			**299**	**100.0%**	**–7.03 (–15.24 to 1.19)**
Heterogeneity: $\chi^2 = 4.06$, df = 1 ($p = 0.04$); $I^2 = 75\%$								
Test for overall effect: $z = 1.68$ ($p = 0.09$)								

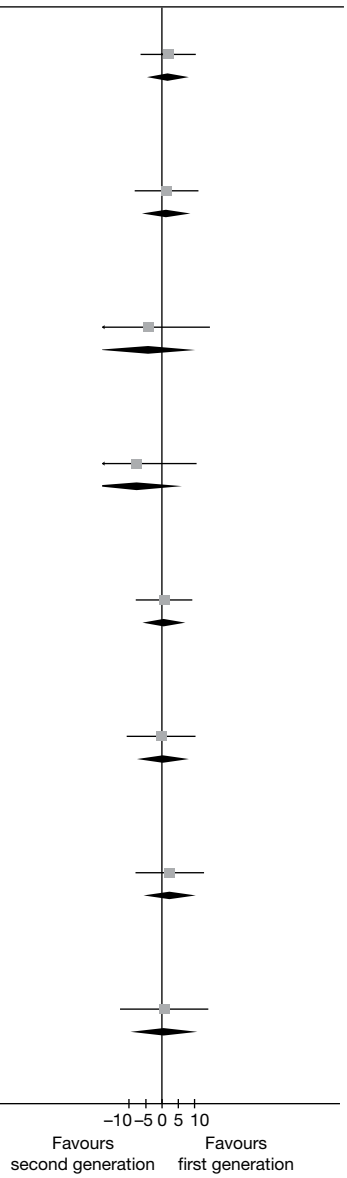


Study or subgroup	First generation Mean	SD	Total	Second generation Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.17.12 SF-36 role limitation emotional (12 months)								
b. Cooper 2004[99]	18.7	45.6	91	15.6	43.3	182	53.6%	3.10 (–11.73 to 17.93)
l. Cooper 1999[54]	12.6	48.6	124	17.8	47.2	116	46.4%	–5.20 (–21.13 to 10.73)
Subtotal (95% CI)			**215**			**298**	**100.0%**	**–0.75 (–9.01 to 7.51)**
Heterogeneity: $\chi^2 = 0.96$, df = 1 ($p = 0.33$); $I^2 = 0\%$								
Test for overall effect: $z = 0.18$ ($p = 0.86$)								
1.17.13 SF-36 mental health (12 months)								
b. Cooper 2004[99]	4.8	19.1	89	7.8	18.8	183	54.6%	–3.00 (–9.33 to 3.33)
l. Cooper 1999[54]	5.7	22.4	124	6.5	19.3	116	45.4%	–0.80 (–7.74 to 6.14)
Subtotal (95% CI)			**213**			**299**	**100.0%**	**–2.00 (–5.56 to 1.56)**
Heterogeneity: $\chi^2 = 0.36$, df = 1 ($p = 0.55$); $I^2 = 0\%$								
Test for overall effect: $z = 1.10$ ($p = 0.27$)								
1.17.14 SF-36 social function (12 months)								
b. Cooper 2004[99]	11.2	19.9	82	11.1	20.1	157	55.4%	0.10 (–6.91 to 7.11)
l. Cooper 1999[54]	7.3	22.4	123	12.1	24.4	116	44.6%	–4.80 (–12.62 to 3.02)
Subtotal (95% CI)			**205**			**273**	**100.0%**	**–2.08 (–6.05 to 1.89)**
Heterogeneity: $\chi^2 = 1.45$, df = 1 ($p = 0.23$); $I^2 = 31\%$								
Test for overall effect: $z = 1.03$ ($p = 0.30$)								
1.17.15 SF-36 vitality (12 months)								
b. Cooper 2004[99]	15.3	23.1	91	16.3	24.4	181	47.5%	–1.00 (–8.79 to 6.79)
l. Cooper 1999[54]	11.8	23.5	123	13.1	21	116	52.5%	–1.30 (–8.72 to 6.12)
Subtotal (95% CI)			**214**			**297**	**100.0%**	**–1.16 (–5.25 to 2.93)**
Heterogeneity: $\chi^2 = 0.01$, df = 1 ($p = 0.94$); $I^2 = 0\%$								
Test for overall effect: $z = 0.55$ ($p = 0.58$)								
1.17.16 SF-36 pain (12 months)								
b. Cooper 2004[99]	26.3	32.1	91	22.1	32.4	181	48.6%	4.20 (–6.46 to 14.86)
l. Cooper 1999[54]	7.6	31.3	124	15.6	31	116	51.4%	–8.00 (–18.36 to 2.36)
Subtotal (95% CI)			**215**			**297**	**100.0%**	**–2.07 (–7.73 to 3.58)**
Heterogeneity: $\chi^2 = 4.47$, df = 1 ($p = 0.03$); $I^2 = 78\%$								
Test for overall effect: $z = 0.72$ ($p = 0.47$)								
1.17.17 SF-36 general health (2 years)								
l. Cooper 1999[54]	2.8	20.4	128	4.7	23.5	121	100.0%	–1.90 (–9.10 to 5.30)
Subtotal (95% CI)			**128**			**121**	**100.0%**	**–1.90 (–7.38 to 3.58)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.68$ ($p = 0.50$)								
1.17.18 SF-36 physical function (2 years)								
l. Cooper 1999[54]	1	20.3	126	2.3	21.2	118	100.0%	–1.30 (–8.15 to 5.55)
Subtotal (95% CI)			**126**			**118**	**100.0%**	**–1.30 (–6.51 to 3.91)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.49$ ($p = 0.63$)								
1.17.19 SF-36 role limitation physical (2 years)								
l. Cooper 1999[54]	6.1	43.7	128	18.4	53.5	121	100.0%	–12.30 (–28.30 to 3.70)
Subtotal (95% CI)			**128**			**121**	**100.0%**	**–12.30 (–24.47 to –0.13)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.98$ ($p = 0.05$)								
1.17.20 SF-36 role limitation emotional (2 years)								
l. Cooper 1999[54]	9.4	47.9	128	17.6	17.4	121	100.0%	–8.20 (–19.84 to 3.44)
Subtotal (95% CI)			**128**			**121**	**100.0%**	**–8.20 (–17.06 to 0.66)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.81$ ($p = 0.07$)								
1.17.21 SF-36 mental health (2 years)								
l. Cooper 1999[54]	4.2	19.7	127	6	21.5	121	100.0%	–1.80 (–8.55 to 4.95)
Subtotal (95% CI)			**127**			**121**	**100.0%**	**–1.80 (–6.94 to 3.34)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.69$ ($p = 0.49$)								



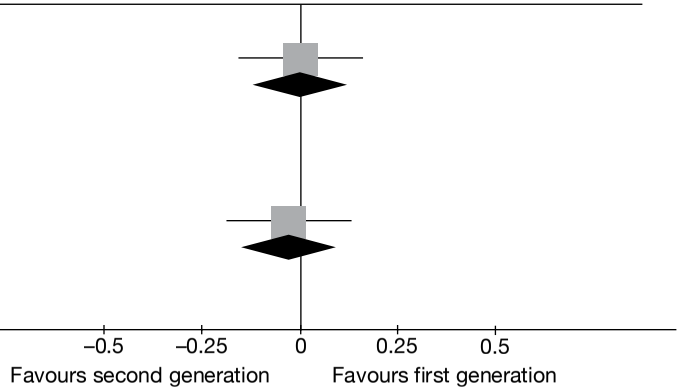


Study or subgroup	First generation Mean	SD	Total	Second generation Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.17.33 SF-36 general health (10 years)								
I. Cooper 1999[54]	2.8	22	95	0.9	23	94	100.0%	1.90 (–6.53 to 10.33)
Subtotal (95% CI)			**95**			**94**	**100.0%**	**1.90 (–4.52 to 8.32)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.58$ ($p = 0.56$)								
1.17.34 SF-36 physical function (10 years)								
I. Cooper 1999[54]	–3	25	95	–4.4	27	94	100.0%	1.40 (–8.35 to 11.15)
Subtotal (95% CI)			**95**			**94**	**100.0%**	**1.40 (–6.02 to 8.82)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.37$ ($p = 0.71$)								
1.17.35 SF-36 role limitation physical (10 years)								
I. Cooper 1999[54]	10.9	47	95	15	53	94	100.0%	–4.10 (–22.88 to 14.68)
Subtotal (95% CI)			**95**			**94**	**100.0%**	**–4.10 (–18.39 to 10.19)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.56$ ($p = 0.57$)								
1.17.36 SF-36 role limitation emotional (10 years)								
I. Cooper 1999[54]	13.5	47	95	21.1	50	94	100.0%	–7.60 (–25.79 to 10.59)
Subtotal (95% CI)			**95**			**94**	**100.0%**	**–7.60 (–21.44 to 6.24)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.08$ ($p = 0.28$)								
1.17.37 SF-36 mental health (10 years)								
I. Cooper 1999[54]	7.9	25	95	7.2	21	94	100.0%	0.70 (–7.95 to 9.35)
Subtotal (95% CI)			**95**			**94**	**100.0%**	**0.70 (–5.88 to 7.28)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.21$ ($p = 0.83$)								
1.17.38 SF-36 social function (10 years)								
I. Cooper 1999[54]	9.9	26	95	10.1	30	94	100.0%	–0.20 (–10.72 to 10.32)
Subtotal (95% CI)			**95**			**94**	**100.0%**	**–0.20 (–8.21 to 7.81)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.05$ ($p = 0.96$)								
1.17.39 SF-36 vitality (10 years)								
I. Cooper 1999[54]	15.3	27	95	12.9	29	94	100.0%	2.40 (–8.10 to 12.90)
Subtotal (95% CI)			**95**			**94**	**100.0%**	**2.40 (–5.59 to 10.39)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.59$ ($p = 0.56$)								
1.17.40 SF-36 pain (10 years)								
I. Cooper 1999[54]	12.3	35	95	11.6	37	94	100.0%	0.70 (–12.80 to 14.20)
Subtotal (95% CI)			**95**			**94**	**100.0%**	**0.70 (–9.57 to 10.97)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.13$ ($p = 0.89$)								



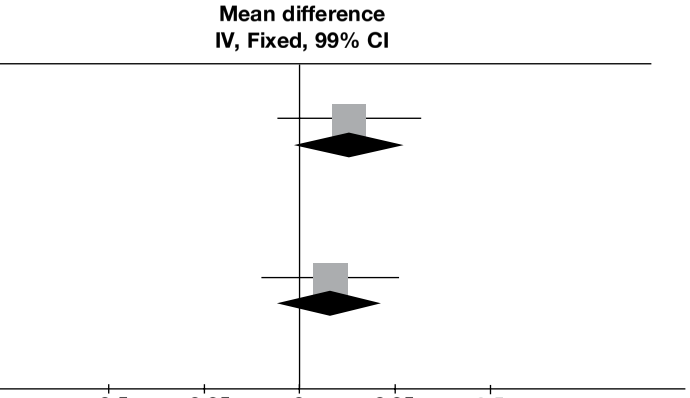
EQ-5D scores (absolute values)

Study or subgroup	First generation Mean	SD	Total	Second generation Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.18.2 6 months								
e. Hawe 2003[94]	0.8	0.24	31	0.8	0.27	37	100.0%	0.00 (–0.16 to 0.16)
Subtotal (95% CI)			**31**			**37**	**100.0%**	**0.00 (–0.12 to 0.12)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.00$ ($p = 1.00$)								
1.18.3 12 months								
e. Hawe 2003[94]	0.82	0.25	28	0.85	0.23	33	100.0%	–0.03 (–0.19 to 0.13)
Subtotal (95% CI)			**28**			**33**	**100.0%**	**–0.03 (–0.15 to 0.09)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.48$ ($p = 0.63$)								



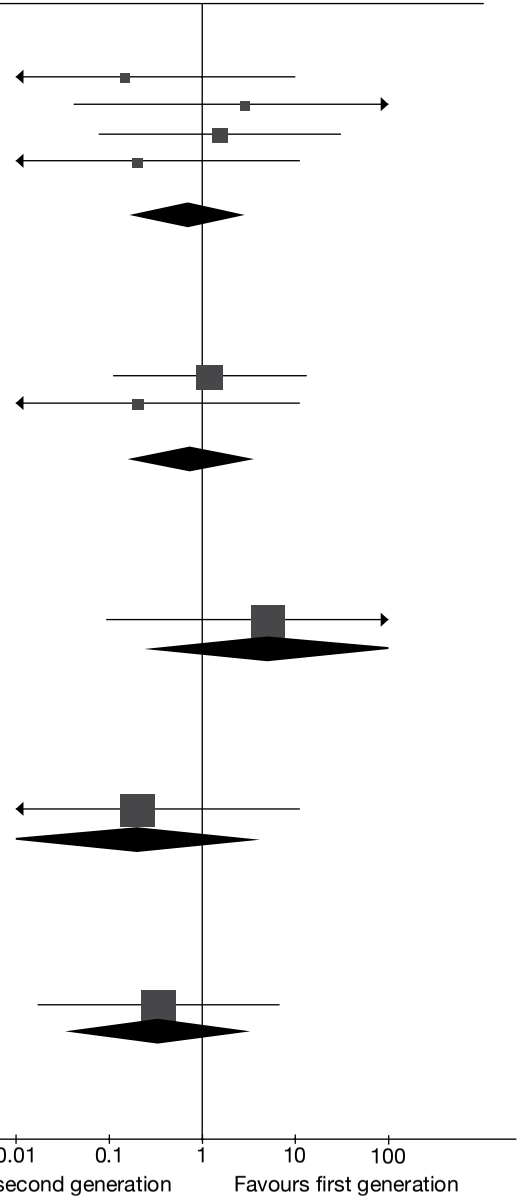
EQ-5D scores (change from baseline)

Study or subgroup	First generation Mean	SD	Total	Second generation Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.19.1 6 months								
e. Hawe 2003[94]	0.15	0.29	29	0.02	0.3	37	100.0%	0.13 (–0.06 to 0.32)
Subtotal (95% CI)			**29**			**37**	**100.0%**	**0.13 (–0.01 to 0.27)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.78$ ($p = 0.08$)								
1.19.2 12 months								
e. Hawe 2003[94]	0.13	0.22	27	0.05	0.32	33	100.0%	0.08 (–0.10 to 0.26)
Subtotal (95% CI)			**27**			**33**	**100.0%**	**0.08 (–0.06 to 0.22)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.14$ ($p = 0.25$)								



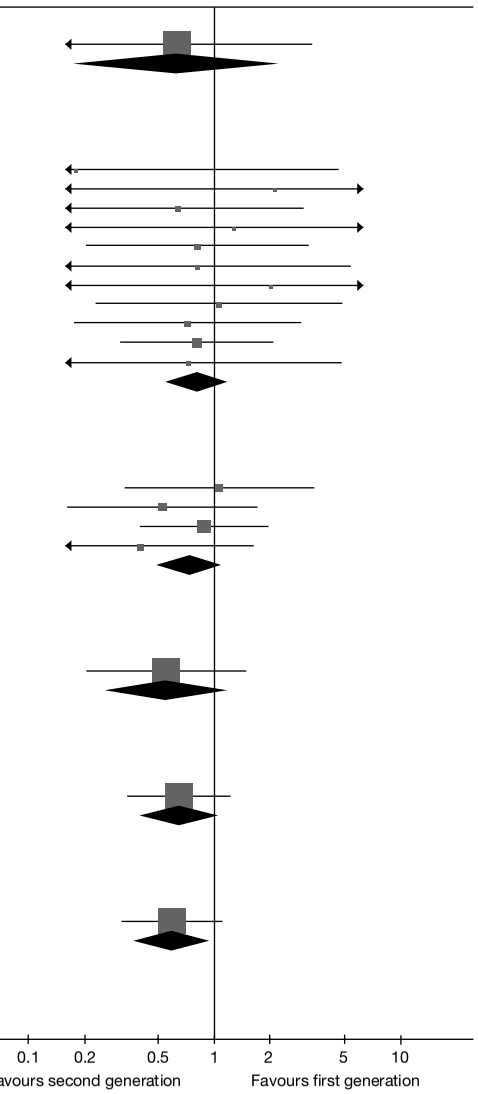
Proportion requiring repeat endometrial ablation

Study or subgroup	First generation Events	First generation Total	Second generation Events	Second generation Total	Weight	Peto odds ratio (non-event) Peto, Fixed, 99% CI
1.20.1 12 months						
b. Cooper 2004[99]	0	106	0	209		Not estimable
d. Duleba 2003[98]	1	86	0	193	19.7%	0.15 (0.00 to 10.02)
e. Hawe 2003[94]	0	35	1	37	19.4%	2.92 (0.04 to 204.54)
g. Cooper 2002[56]	1	90	3	175	39.1%	1.55 (0.08 to 30.97)
l. Cooper 1999[54]	2	134	0	129	21.9%	0.20 (0.00 to 11.21)
m. Meyer 1998[53]	0	138	0	137		Not estimable
Subtotal (95% CI)		**589**		**880**	**100.0%**	**0.71 (0.17 to 2.94)**
Total events	4		4			
Heterogeneity: $\chi^2 = 2.75$, df = 3 ($p = 0.43$); $I^2 = 0\%$						
Test for overall effect: $z = 0.47$ ($p = 0.64$)						
1.20.2 2 years						
f. van Zon-Rabelink 2001[95]	2	62	3	77	73.7%	1.22 (0.11 to 13.32)
l. Cooper 1999[54]	2	134	0	129	26.3%	0.20 (0.00 to 11.21)
m. Meyer 1998[53]	0	138	0	137		Not estimable
Subtotal (95% CI)		**334**		**343**	**100.0%**	**0.76 (0.16 to 3.63)**
Total events	4		3			
Heterogeneity: $\chi^2 = 0.97$, df = 1 ($p = 0.32$); $I^2 = 0\%$						
Test for overall effect: $z = 0.34$ ($p = 0.73$)						
1.20.3 3 years						
m. Meyer 1998[53]	0	138	2	137	100.0%	5.11 (0.09 to 279.75)
Subtotal (95% CI)		**138**		**137**	**100.0%**	**5.11 (0.24 to 107.44)**
Total events	0		2			
Heterogeneity: not applicable						
Test for overall effect: $z = 1.05$ ($p = 0.29$)						
1.20.4 5 years						
l. Cooper 1999[54]	2	134	0	129	100.0%	0.20 (0.00 to 11.21)
Subtotal (95% CI)		**134**		**129**	**100.0%**	**0.20 (0.01 to 4.30)**
Total events	2		0			
Heterogeneity: not applicable						
Test for overall effect: $z = 1.02$ ($p = 0.31$)						
1.20.5 10 years						
l. Cooper 1999[54]	3	134	1	129	100.0%	0.34 (0.02 to 6.79)
Subtotal (95% CI)		**134**		**129**	**100.0%**	**0.34 (0.04 to 3.32)**
Total events	3		1			
Heterogeneity: not applicable						
Test for overall effect: $z = 0.93$ ($p = 0.35$)						



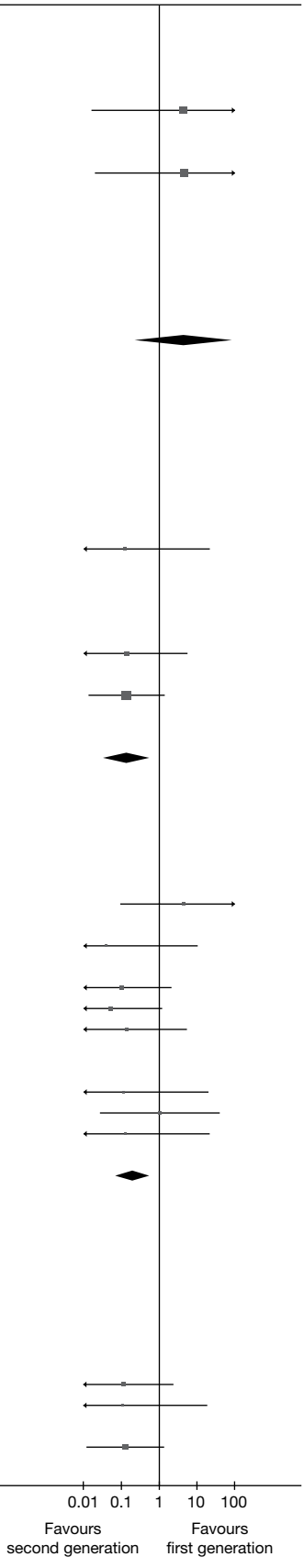
Proportion requiring hysterectomy after endometrial ablation

Study or subgroup	First generation Events	First generation Total	Second generation Events	Second generation Total	Weight	Peto odds ratio (non-event) Peto, Fixed, 99% CI
1.21.1 <12 months						
e. Hawe 2003[94]	4	28	3	35	100.0%	0.56 (0.07 to 4.53)
Subtotal (95% CI)		**28**		**35**	**100.0%**	**0.56 (0.11 to 2.75)**
Total events	4		3			
Heterogeneity: not applicable						
Test for overall effect: $z = 0.71$ ($p = 0.48$)						
1.21.2 12 months						
a. Brun 2006[103]	2	20	0	31	2.4%	0.12 (0.00 to 6.82)
b. Cooper 2004[99]	0	106	2	209	2.5%	2.57 (0.05 to 140.43)
c. Perino 2004[100]	5	55	3	56	10.5%	0.57 (0.08 to 3.97)
d. Duleba 2003[98]	0	86	1	193	2.2%	1.35 (0.02 to 91.67)
e. Hawe 2003[94]	6	30	5	31	13.4%	0.77 (0.14 to 4.30)
g. Cooper 2002[56]	2	90	3	175	7.1%	0.77 (0.07 to 8.25)
i. Corson 2001[79]	0	89	2	187	2.5%	2.41 (0.04 to 132.27)
j. Soysal 2001[96]	4	48	4	45	11.0%	1.07 (0.16 to 7.21)
k. Corson 2000[101]	5	126	4	150	12.9%	0.66 (0.11 to 3.84)
l. Cooper 1999[54]	12	134	9	129	28.4%	0.76 (0.23 to 2.49)
m. Meyer 1998[53]	3	138	2	137	7.1%	0.67 (0.06 to 7.15)
Subtotal (95% CI)		**922**		**1343**	**100.0%**	**0.77 (0.47 to 1.24)**
Total events	39		35			
Heterogeneity: $\chi^2 = 3.12$, df = 10 ($p = 0.98$); $I^2 = 0\%$						
Test for overall effect: $z = 1.08$ ($p = 0.28$)						
1.21.3 2 years						
f. van Zon-Rabelink 2001[95]	6	62	8	77	20.5%	1.08 (0.25 to 4.69)
k. Corson 2000[101]	9	126	5	150	20.3%	0.45 (0.10 to 1.95)
l. Cooper 1999[54]	17	128	14	121	44.7%	0.85 (0.32 to 2.31)
m. Meyer 1998[53]	9	138	3	137	14.5%	0.32 (0.06 to 1.84)
Subtotal (95% CI)		**454**		**485**	**100.0%**	**0.68 (0.41 to 1.13)**
Total events	41		30			
Heterogeneity: $\chi^2 = 2.77$, df = 3 ($p = 0.43$); $I^2 = 0\%$						
Test for overall effect: $z = 1.48$ ($p = 0.14$)						
1.21.4 3 years						
m. Meyer 1998[53]	14	138	7	137	100.0%	0.48 (0.14 to 1.64)
Subtotal (95% CI)		**138**		**137**	**100.0%**	**0.48 (0.19 to 1.22)**
Total events	14		7			
Heterogeneity: not applicable						
Test for overall effect: $z = 1.54$ ($p = 0.12$)						
1.21.5 5 years						
l. Cooper 1999[54]	34	136	21	130	100.0%	0.58 (0.26 to 1.28)
Subtotal (95% CI)		**136**		**130**	**100.0%**	**0.58 (0.31 to 1.06)**
Total events	34		21			
Heterogeneity: not applicable						
Test for overall effect: $z = 1.77$ ($p = 0.08$)						
1.21.6 10 years						
l. Cooper 1999[54]	38	134	22	129	100.0%	0.52 (0.24 to 1.13)
Subtotal (95% CI)		**134**		**129**	**100.0%**	**0.52 (0.29 to 0.94)**
Total events	38		22			
Heterogeneity: not applicable						
Test for overall effect: $z = 2.17$ ($p = 0.03$)						

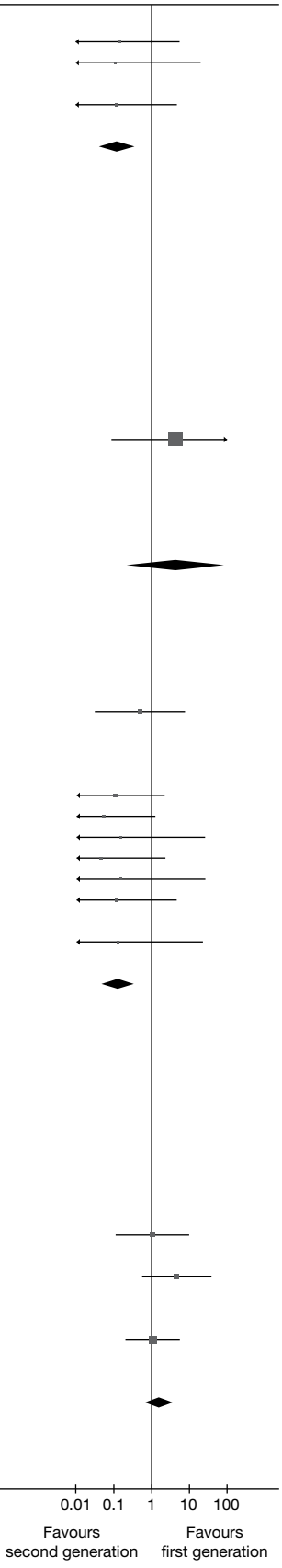


Number of patients with adverse events – periprocedure

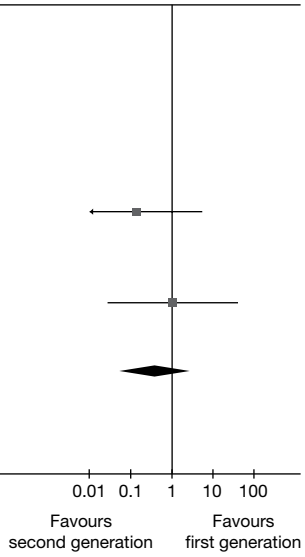
Study or subgroup	First generation Events	First generation Total	Second generation Events	Second generation Total	Weight	Peto odds ratio (non-event) Peto, Fixed, 99% CI
1.22.1 Anaesthesia problems						
a. Brun 2006[103]	0	20	0	31		Not estimable
b. Cooper 2004[99]	0	107	0	215		Not estimable
c. Perino 2004[100]	0	55	0	56		Not estimable
d. Duleba 2003[98]	0	86	1	193	48.7%	4.24 (0.02 to 1123.03)
e. Hawe 2003[94]	0	34	0	37		Not estimable
f. van Zon-Rabelink 2001[95]	0	62	0	77		Not estimable
g. Cooper 2002[56]	0	90	1	175	51.3%	4.55 (0.02 to 1046.63)
h. Pellicano 2002[102]	0	42	0	40		Not estimable
i. Corson 2001[79]	0	89	0	187		Not estimable
j. Soysal 2001[96]	0	48	0	45		Not estimable
k. Corson 2000[101]	0	123	0	144		Not estimable
l. Cooper 1999[54]	0	134	0	129		Not estimable
m. Meyer 1998[53]	0	117	0	128		Not estimable
n. Romer 1998[97]	0	10	0	10		Not estimable
Subtotal (95% CI)		**1017**		**1467**	**100.0%**	**4.40 (0.23 to 85.11)**
Total events	0		2			
Heterogeneity: $\chi^2 = 0.00$, df = 1 ($p = 0.98$); $I^2 = 0\%$						
Test for overall effect: $z = 0.98$ ($p = 0.33$)						
1.22.2 Excessive bleeding						
a. Brun 2006[103]	0	20	0	31		Not estimable
b. Cooper 2004[99]	0	107	0	215		Not estimable
c. Perino 2004[100]	0	55	0	56		Not estimable
d. Duleba 2003[98]	0	86	0	193		Not estimable
e. Hawe 2003[94]	1	34	0	37	12.6%	0.12 (0.00 to 21.50)
f. van Zon-Rabelink 2001[95]	0	62	0	77		Not estimable
g. Cooper 2002[56]	0	90	0	175		Not estimable
h. Pellicano 2002[102]	0	42	0	40		Not estimable
i. Corson 2001[79]	0	89	0	187		Not estimable
j. Soysal 2001[96]	2	48	0	45	25.0%	0.14 (0.00 to 5.51)
k. Corson 2000[101]	0	123	0	144		Not estimable
l. Cooper 1999[54]	5	134	0	129	62.3%	0.14 (0.01 to 1.39)
m. Meyer 1998[53]	0	117	0	128		Not estimable
n. Romer 1998[97]	0	10	0	10		Not estimable
Subtotal (95% CI)		**1017**		**1467**	**100.0%**	**0.14 (0.03 to 0.55)**
Total events	8		0			
Heterogeneity: $\chi^2 = 0.00$, df = 2 ($p = 1.00$); $I^2 = 0\%$						
Test for overall effect: $z = 2.81$ ($p = 0.005$)						
1.22.3 Uterine perforation						
a. Brun 2006[103]	0	20	0	31		Not estimable
b. Cooper 2004[99]	0	107	2	215	12.5%	4.49 (0.09 to 215.97)
c. Perino 2004[100]	0	55	0	56		Not estimable
d. Duleba 2003[98]	1	86	0	193	6.0%	0.04 (0.00 to 10.32)
e. Hawe 2003[94]	0	34	0	37		Not estimable
f. van Zon-Rabelink 2001[95]	3	62	0	77	20.6%	0.10 (0.01 to 2.09)
g. Cooper 2002[56]	3	90	0	175	18.8%	0.05 (0.00 to 1.20)
h. Pellicano 2002[102]	2	42	0	40	13.9%	0.14 (0.00 to 5.42)
i. Corson 2001[79]	0	89	0	187		Not estimable
j. Soysal 2001[96]	0	48	0	45		Not estimable
k. Corson 2000[101]	1	123	0	144	7.0%	0.11 (0.00 to 20.02)
l. Cooper 1999[54]	1	134	1	129	14.1%	1.04 (0.03 to 39.99)
m. Meyer 1998[53]	1	117	0	128	7.0%	0.12 (0.00 to 21.39)
n. Romer 1998[97]	0	10	0	10		Not estimable
Subtotal (95% CI)		**1017**		**1467**	**100.0%**	**0.20 (0.07 to 0.57)**
Total events	12		3			
Heterogeneity: $\chi^2 = 7.96$, df = 7 ($p = 0.34$); $I^2 = 12\%$						
Test for overall effect: $z = 3.02$ ($p = 0.003$)						
1.22.4 Fluid overload						
a. Brun 2006[103]	0	20	0	31		Not estimable
b. Cooper 2004[99]	0	107	0	215		Not estimable
c. Perino 2004[100]	0	55	0	56		Not estimable
d. Duleba 2003[98]	0	86	0	193		Not estimable
e. Hawe 2003[94]	3	34	0	37	21.4%	0.12 (0.01 to 2.39)
f. van Zon-Rabelink 2001[95]	1	62	0	77	7.3%	0.11 (0.00 to 18.92)
g. Cooper 2002[56]	0	90	0	175		Not estimable
h. Pellicano 2002[102]	5	42	0	40	34.9%	0.13 (0.01 to 1.36)



Study or subgroup	First generation Events	Total	Second generation Events	Total	Weight	Peto odds ratio (non-event) Peto, Fixed, 99% CI
i. Corson 2001[79]	0	89	0	187		Not estimable
j. Soysal 2001[96]	2	48	0	45	14.5%	0.14 (0.00 to 5.51)
k. Corson 2000[101]	1	123	0	144	7.3%	0.11 (0.00 to 20.02)
l. Cooper 1999[54]	0	134	0	129		Not estimable
m. Meyer 1998[53]	2	117	0	128	14.6%	0.12 (0.00 to 4.72)
n. Romer 1998[97]	0	10	0	10		Not estimable
Subtotal (95% CI)		**1017**		**1467**	**100.0%**	**0.12 (0.04 to 0.36)**
Total events	14		0			
Heterogeneity: $\chi^2 = 0.02$, df = 5 ($p = 1.00$); $I^2 = 0\%$						
Test for overall effect: $z = 3.85$ ($p = 0.0001$)						
1.22.5 Excessive visceral damage						
a. Brun 2006[103]	0	20	0	31		Not estimable
b. Cooper 2004[99]	0	107	0	215		Not estimable
c. Perino 2004[100]	0	55	0	56		Not estimable
d. Duleba 2003[98]	0	86	0	193		Not estimable
e. Hawe 2003[94]	0	34	0	37		Not estimable
f. van Zon-Rabelink 2001[95]	0	62	0	77		Not estimable
g. Cooper 2002[56]	0	90	0	175		Not estimable
h. Pellicano 2002[102]	0	42	0	40		Not estimable
i. Corson 2001[79]	0	89	2	187	100.0%	4.40 (0.09 to 218.13)
j. Soysal 2001[96]	0	48	0	45		Not estimable
k. Corson 2000[101]	0	123	0	144		Not estimable
l. Cooper 1999[54]	0	134	0	129		Not estimable
m. Meyer 1998[53]	0	117	0	128		Not estimable
n. Romer 1998[97]	0	10	0	10		Not estimable
Subtotal (95% CI)		**1017**		**1467**	**100.0%**	**4.40 (0.23 to 85.78)**
Total events	0		2			
Heterogeneity: not applicable						
Test for overall effect: $z = 0.98$ ($p = 0.33$)						
1.22.6 Cervical laceration						
a. Brun 2006[103]	0	20	0	31		Not estimable
b. Cooper 2004[99]	2	107	2	215	22.2%	0.47 (0.03 to 7.27)
c. Perino 2004[100]	0	55	0	56		Not estimable
d. Duleba 2003[98]	0	86	0	193		Not estimable
e. Hawe 2003[94]	0	34	0	37		Not estimable
f. van Zon-Rabelink 2001[95]	3	62	0	77	18.5%	0.10 (0.01 to 2.09)
g. Cooper 2002[56]	3	90	0	175	16.9%	0.05 (0.00 to 1.20)
h. Pellicano 2002[102]	1	42	0	40	6.3%	0.14 (0.00 to 24.55)
i. Corson 2001[79]	2	89	0	187	11.0%	0.04 (0.00 to 2.21)
j. Soysal 2001[96]	1	48	0	45	6.3%	0.14 (0.00 to 24.95)
k. Corson 2000[101]	2	123	0	144	12.5%	0.11 (0.00 to 4.40)
l. Cooper 1999[54]	0	134	0	129		Not estimable
m. Meyer 1998[53]	1	117	0	128		0.12 (0.00 to 21.39)
n. Romer 1998[97]	0	10	0	10		Not estimable
Subtotal (95% CI)		**1017**		**1467**		**0.12 (0.05 to 0.33)**
Total events	15		2			
Heterogeneity: $\chi^2 = 2.56$, df = 7 ($p = 0.92$); $I^2 = 0\%$						
Test for overall effect: $z = 4.14$ ($p < 0.0001$)						
1.22.7 Procedure abandoned						
a. Brun 2006[103]	0	20	0	31		Not estimable
b. Cooper 2004[99]	0	107	0	215		Not estimable
c. Perino 2004[100]	0	55	0	56		Not estimable
d. Duleba 2003[98]	0	86	0	193		Not estimable
e. Hawe 2003[94]	0	34	0	37		Not estimable
f. van Zon-Rabelink 2001[95]	0	62	0	77		Not estimable
g. Cooper 2002[56]	2	90	4	175	25.2%	1.03 (0.11 to 9.68)
h. Pellicano 2002[102]	0	42	0	40		Not estimable
i. Corson 2001[79]	0	89	7	187	28.6%	4.52 (0.55 to 37.14)
j. Soysal 2001[96]	0	48	0	45		Not estimable
k. Corson 2000[101]	0	123	0	144		Not estimable
l. Cooper 1999[54]	5	134	5	129	46.1%	1.04 (0.20 to 5.46)
m. Meyer 1998[53]	0	117	0	128		Not estimable
n. Romer 1998[97]	0	10	0	10		Not estimable
Subtotal (95% CI)		**1017**		**1467**	**100.0%**	**1.58 (0.67 to 3.72)**
Total events	7		16			
Heterogeneity: $\chi^2 = 2.32$, df = 2 ($p = 0.31$); $I^2 = 14\%$						
Test for overall effect: $z = 1.05$ ($p = 0.30$)						

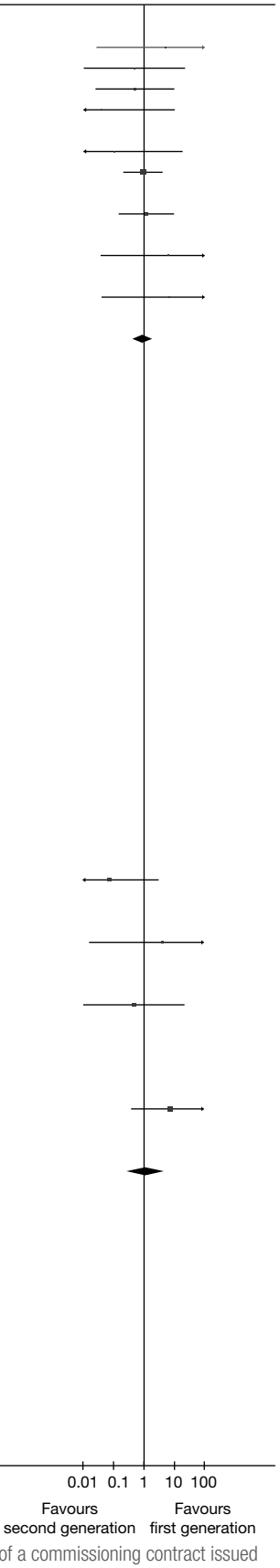


Study or subgroup	First generation Events	First generation Total	Second generation Events	Second generation Total	Weight	Peto odds ratio (non-event) Peto, Fixed, 99% CI
1.22.8 Procedure converted to hysterectomy						
a. Brun 2006[103]	0	20	0	31		Not estimable
b. Cooper 2004[99]	0	107	0	215		Not estimable
c. Perino 2004[100]	0	55	0	56		Not estimable
d. Duleba 2003[98]	0	86	0	193		Not estimable
e. Hawe 2003[94]	0	34	0	37		Not estimable
f. van Zon-Rabelink 2001[95]	0	62	0	77		Not estimable
g. Cooper 2002[56]	0	90	0	175		Not estimable
h. Pellicano 2002[102]	2	42	0	40	49.8%	0.14 (0.00 to 5.42)
i. Corson 2001[79]	0	89	0	187		Not estimable
j. Soysal 2001[96]	0	48	0	45		Not estimable
k. Corson 2000[101]	0	123	0	144		Not estimable
l. Cooper 1999[54]	1	134	1	129	50.2%	1.04 (0.03 to 39.99)
m. Meyer 1998[53]	0	117	0	128		Not estimable
n. Romer 1998[97]	0	10	0	10		Not estimable
Subtotal (95% CI)		**1017**		**1467**	**100.0%**	**0.38 (0.05 to 2.73)**
Total events	3		1			
Heterogeneity: $\chi^2 = 1.01$, df = 1 ($p = 0.32$); $I^2 = 1\%$						
Test for overall effect: $z = 0.96$ ($p = 0.34$)						

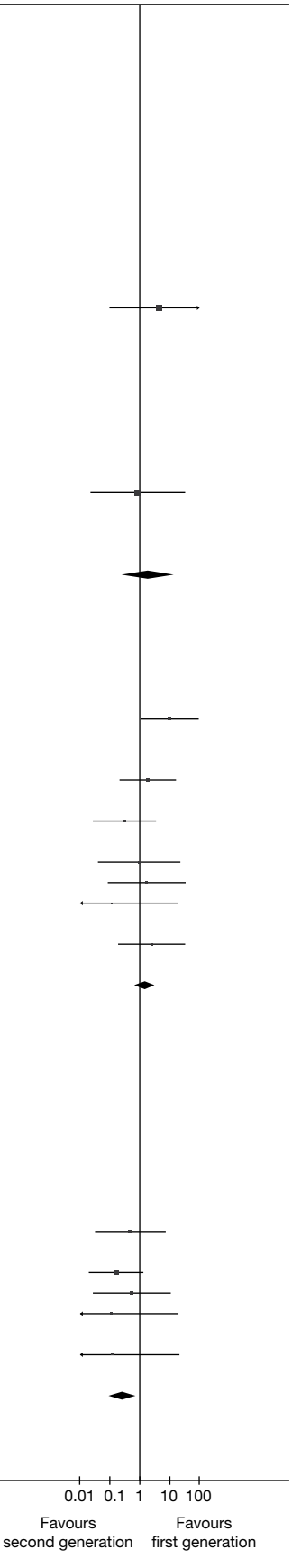


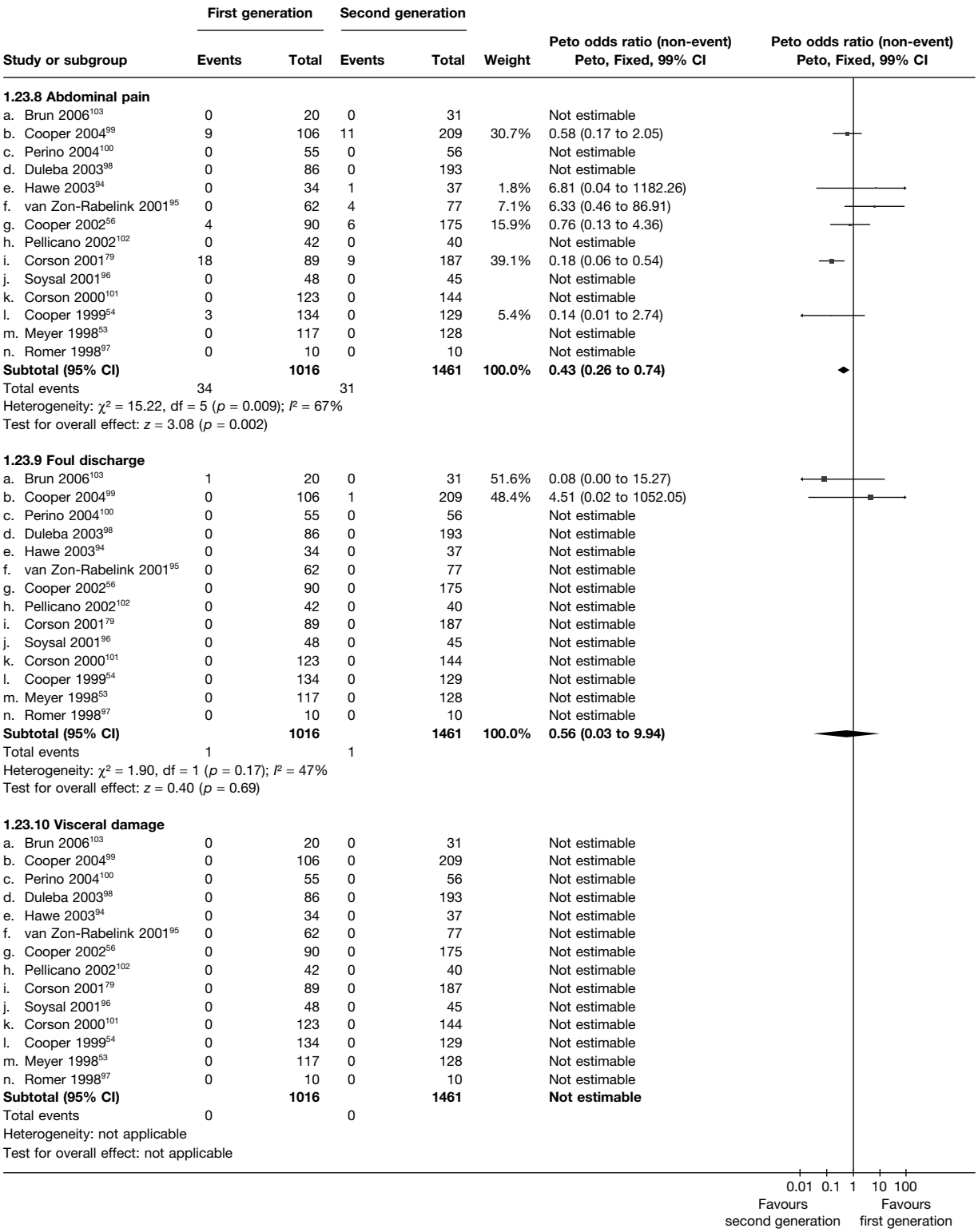
Number of patients with adverse events – postoperatively (within 1 month)

Study or subgroup	First generation Events	Total	Second generation Events	Total	Weight	Peto odds ratio (non-event) Peto, Fixed, 99% CI
1.23.1 Urinary tract infection						
a. Brun 2006[103]	0	20	1	31	3.5%	5.18 (0.03 to 1013.33)
b. Cooper 2004[99]	1	106	1	209	6.5%	0.48 (0.01 to 22.79)
c. Perino 2004[100]	2	55	1	56	10.7%	0.50 (0.02 to 10.02)
d. Duleba 2003[98]	1	86	0	193	3.1%	0.04 (0.00 to 10.32)
e. Hawe 2003[94]	0	34	0	37		Not estimable
f. van Zon-Rabelink 2001[95]	1	62	0	77	3.6%	0.11 (0.00 to 18.92)
g. Cooper 2002[56]	5	90	9	175	43.5%	0.92 (0.21 to 4.09)
h. Pellicano 2002[102]	0	42	0	40		Not estimable
i. Corson 2001[79]	2	89	5	187	21.8%	1.19 (0.14 to 9.76)
j. Soysal 2001[96]	0	48	0	45		Not estimable
k. Corson 2000[101]	0	123	1	144	3.6%	6.39 (0.04 to 1120.84)
l. Cooper 1999[54]	0	134	0	129		Not estimable
m. Meyer 1998[53]	0	117	1	128	3.6%	6.78 (0.04 to 1177.23)
n. Romer 1998[97]	0	10	0	10		Not estimable
Subtotal (95% CI)		**1016**		**1461**	**100.0%**	**0.90 (0.42 to 1.90)**
Total events	12		19			
Heterogeneity: $\chi^2 = 6.48$, df = 8 ($p = 0.59$); $I^2 = 0\%$						
Test for overall effect: $z = 0.28$ ($p = 0.78$)						
1.23.2 Deep-vein thrombosis						
a. Brun 2006[103]	0	20	0	31		Not estimable
b. Cooper 2004[99]	0	106	0	209		Not estimable
c. Perino 2004[100]	0	55	0	56		Not estimable
d. Duleba 2003[98]	0	86	0	193		Not estimable
e. Hawe 2003[94]	0	34	0	37		Not estimable
f. van Zon-Rabelink 2001[95]	0	62	0	77		Not estimable
g. Cooper 2002[56]	0	90	0	175		Not estimable
h. Pellicano 2002[102]	0	42	0	40		Not estimable
i. Corson 2001[79]	0	89	0	187		Not estimable
j. Soysal 2001[96]	0	48	0	45		Not estimable
k. Corson 2000[101]	0	123	0	144		Not estimable
l. Cooper 1999[54]	0	134	0	129		Not estimable
m. Meyer 1998[53]	0	117	0	128		Not estimable
n. Romer 1998[97]	0	10	0	10		Not estimable
Subtotal (95% CI)		**1016**		**1461**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
1.23.3 Further bleeding						
a. Brun 2006[103]	2	20	0	31	25.0%	0.07 (0.00 to 3.21)
b. Cooper 2004[99]	0	106	0	209		Not estimable
c. Perino 2004[100]	0	55	0	56		Not estimable
d. Duleba 2003[98]	0	86	1	193	11.4%	4.24 (0.02 to 1123.03)
e. Hawe 2003[94]	0	34	0	37		Not estimable
f. van Zon-Rabelink 2001[95]	0	62	0	77		Not estimable
g. Cooper 2002[56]	1	90	1	175	23.9%	0.49 (0.01 to 23.00)
h. Pellicano 2002[102]	0	42	0	40		Not estimable
i. Corson 2001[79]	0	89	0	187		Not estimable
j. Soysal 2001[96]	0	48	0	45		Not estimable
k. Corson 2000[101]	0	123	0	144		Not estimable
l. Cooper 1999[54]	0	134	3	129	39.8%	7.80 (0.39 to 154.58)
m. Meyer 1998[53]	0	117	0	128		Not estimable
n. Romer 1998[97]	0	10	0	10		Not estimable
Subtotal (95% CI)		**1016**		**1461**	**100.0%**	**1.17 (0.28 to 4.92)**
Total events	3		5			
Heterogeneity: $\chi^2 = 6.93$, df = 3 ($p = 0.07$); $I^2 = 57\%$						
Test for overall effect: $z = 0.22$ ($p = 0.83$)						
1.23.4 Sepsis						
a. Brun 2006[103]	0	20	0	31		Not estimable
b. Cooper 2004[99]	0	106	0	209		Not estimable
c. Perino 2004[100]	0	55	0	56		Not estimable
d. Duleba 2003[98]	0	86	0	193		Not estimable
e. Hawe 2003[94]	0	34	0	37		Not estimable
f. van Zon-Rabelink 2001[95]	0	62	0	77		Not estimable
g. Cooper 2002[56]	0	90	0	175		Not estimable



Study or subgroup	First generation Events	Total	Second generation Events	Total	Weight	Peto odds ratio (non-event) Peto, Fixed, 99% CI
h. Pellicano 2002[102]	0	42	0	40		Not estimable
i. Corson 2001[79]	0	89	0	187		Not estimable
j. Soysal 2001[96]	0	48	0	45		Not estimable
k. Corson 2000[101]	0	123	0	144		Not estimable
l. Cooper 1999[54]	0	134	0	129		Not estimable
m. Meyer 1998[53]	0	117	0	128		Not estimable
n. Romer 1998[97]	0	10	0	10		Not estimable
Subtotal (95% CI)		**1016**		**1461**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
1.23.5 Pyrexia						
a. Brun 2006[103]	0	20	0	31		Not estimable
b. Cooper 2004[99]	0	106	2	209	47.3%	4.54 (0.10 to 215.47)
c. Perino 2004[100]	0	55	0	56		Not estimable
d. Duleba 2003[98]	0	86	0	193		Not estimable
e. Hawe 2003[94]	0	34	0	37		Not estimable
f. van Zon-Rabelink 2001[95]	0	62	0	77		Not estimable
g. Cooper 2002[56]	0	90	0	175		Not estimable
h. Pellicano 2002[102]	0	42	0	40		Not estimable
i. Corson 2001[79]	0	89	0	187		Not estimable
j. Soysal 2001[96]	0	48	0	45		Not estimable
k. Corson 2000[101]	1	123	1	144	52.7%	0.85 (0.02 to 33.19)
l. Cooper 1999[54]	0	134	0	129		Not estimable
m. Meyer 1998[53]	0	117	0	128		Not estimable
n. Romer 1998[97]	0	10	0	10		Not estimable
Subtotal (95% CI)		**1016**		**1461**	**100.0%**	**1.88 (0.25 to 14.20)**
Total events	1		3			
Heterogeneity: $\chi^2 = 0.65$, df = 1 ($p = 0.42$); $I^2 = 0\%$						
Test for overall effect: $z = 0.61$ ($p = 0.54$)						
1.23.6 Endometritis						
a. Brun 2006[103]	0	20	0	31		Not estimable
b. Cooper 2004[99]	0	106	6	209	20.5%	4.63 (0.49 to 43.60)
c. Perino 2004[100]	0	55	0	56		Not estimable
d. Duleba 2003[98]	0	86	0	193		Not estimable
e. Hawe 2003[94]	2	34	4	37	21.6%	1.87 (0.21 to 16.65)
f. van Zon-Rabelink 2001[95]	0	62	0	77		Not estimable
g. Cooper 2002[56]	3	90	2	175	17.2%	0.31 (0.03 to 3.57)
h. Pellicano 2002[102]	0	42	0	40		Not estimable
i. Corson 2001[79]	1	89	2	187	10.1%	0.95 (0.04 to 23.18)
j. Soysal 2001[96]	1	48	2	45	11.4%	2.11 (0.10 to 42.82)
k. Corson 2000[101]	1	123	0	144	3.9%	0.11 (0.00 to 20.02)
l. Cooper 1999[54]	0	134	0	129		Not estimable
m. Meyer 1998[53]	1	117	3	128	15.3%	2.52 (0.19 to 33.71)
n. Romer 1998[97]	0	10	0	10		Not estimable
Subtotal (95% CI)		**1016**		**1461**	**100.0%**	**1.47 (0.68 to 3.18)**
Total events	9		19			
Heterogeneity: $\chi^2 = 6.64$, df = 6 ($p = 0.35$); $I^2 = 10\%$						
Test for overall effect: $z = 0.98$ ($p = 0.33$)						
1.23.7 Hematometra						
a. Brun 2006[103]	0	20	0	31		Not estimable
b. Cooper 2004[99]	0	106	0	209		Not estimable
c. Perino 2004[100]	0	55	0	56		Not estimable
d. Duleba 2003[98]	0	86	0	193		Not estimable
e. Hawe 2003[94]	0	34	0	37		Not estimable
f. van Zon-Rabelink 2001[95]	0	62	0	77		Not estimable
g. Cooper 2002[56]	2	90	2	175	24.5%	0.49 (0.03 to 7.48)
h. Pellicano 2002[102]	0	42	0	40		Not estimable
i. Corson 2001[79]	5	89	2	187	41.4%	0.16 (0.02 to 1.31)
j. Soysal 2001[96]	2	48	1	45	20.3%	0.54 (0.03 to 10.94)
k. Corson 2000[101]	1	123	0	144	6.9%	0.11 (0.00 to 20.02)
l. Cooper 1999[54]	0	134	0	129		Not estimable
m. Meyer 1998[53]	1	117	0	128	6.9%	0.12 (0.00 to 21.39)
n. Romer 1998[97]	0	10	0	10		Not estimable
Subtotal (95% CI)		**1016**		**1461**	**100.0%**	**0.26 (0.09 to 0.72)**
Total events	11		5			
Heterogeneity: $\chi^2 = 1.40$, df = 4 ($p = 0.84$); $I^2 = 0\%$						
Test for overall effect: $z = 2.58$ ($p = 0.010$)						





Appendix 7

Pooled results for Mirena versus first-generation endometrial ablation

Appendix 7.1 Clinical outcome and quality of life

	Time point	Trials (no.)	WMD (95% CI)[a]	OR (95%CI)[b]	*p*-value	Hetero (*p*)/I^2 (%)
Proportion amenorrhoea	12 months	3 (177)	–	0.84 (0.43 to 1.63)	0.6	0.3/11
	2 years	1 (44)	–	0.68 (0.19 to 2.45)	0.6	–
	3 years	1 (41)	–	0.68 (0.19 to 2.38)	0.5	–
Proportion with heavy bleeding	12 months	2 (125)	–	1.13 (0.33 to 3.86)	0.9	0.3/0
	3 years	1 (41)	–	1.84 (0.29 to 11.7)	0.5	–
Bleeding score (change)	6 months	1 (68)	–28 (–57 to 1.4)	–	0.06	–
	12 months	3 (168)	–39 (–66 to –12)	–	0.004	0.6/0
	2 years	1 (44)	41 (–189 to 271)	–	0.7	–
	3 years	1 (41)	37 (–202 to 276)	–	0.8	–
SF-36 general health (absolute)	12 months	1 (62)	–6.2 (–14.6 to 2.2)	–	0.2	–
SF-36 physical function (absolute)		1 (62)	–1.2 (–12.7 to 10.3)	–	0.8	–
SF-36 role physical (absolute)		1 (62)	–1.7 (–19.0 to 15.6)	–	0.9	–
SF-36 role emotional (absolute)		1 (62)	–11.1 (–29.1 to 6.9)	–	0.2	–
SF-36 mental health (absolute)		1 (62)	0.5 (–9.2 to 10.2)	–	0.9	–
SF-36 social function (absolute)		1 (62)	0.1 (–11.5 to 11.7)	–	1.0	–
SF-36 vitality (absolute)		1 (62)	1.5 (–7.3 to 10.3)	–	0.7	–
SF-36 pain (absolute)		1 (62)	–11.4 (–24.2 to 1.4)	–	0.08	–
Hysterectomy after EA/Mirena	12 months	1 (70)	–	7.39 (0.15 to 372)	0.3	–
		Trials	**Frequency**			
Discontinued Mirena	12 months	3	12/95 (13%)			
	2 years	1	8/30 (27%)			
	3 years	1	9/30 (30%)			
EA after Mirena	12 months	1	4/30 (13%)			
Hysterectomy after Mirena	12 months	1	0/35 (0%)			

	Trials	Frequency (first-generation: max. 95; Mirena: max. 95)	OR (95% CI)[b]	*p*-value	Hetero (*p*)/I^2 (%)
Periprocedure complications					
Uterine perforation (first-generation, Mirena)	5	0; 0	–	–	–
Cervical laceration (first-generation, Mirena)	5	0; 0	–	–	–
Anaesthesia problems (first-generation)	5	0	–	–	–
Excessive bleeding (first-generation)	5	0	–	–	–
Fluid overload (first-generation)	5	0	–	–	–
Visceral damage (first-generation)	5	0	–	–	–
Procedure abandoned (first-generation)	5	0	–	–	–
Converted to hysterectomy (first-generation)	5	0	–	–	–
Failed to insert (Mirena)	5	0	–	–	–
Further complications (< 1 month)					
Urinary tract infection (first-generation)	5	0	–	–	–
Deep-vein thrombosis (first-generation)	5	0	–	–	–
Further bleeding (first-generation)	5	7	–	–	–
Sepsis (first-generation)	5	0	–	–	–
Pyrexia (first-generation)	5	0	–	–	–
Endometriosis (first-generation)	5	2	–	–	–
Haematomata (first-generation)	5	3	–	–	–
Abdominal pain (first-generation)	5	4	–	–	–
Foul discharge (first-generation)	5	0	–	–	–
Visceral damage (first-generation)	5	0	–	–	–
Infection (Mirena)	5	0	–	–	–
Expelled/migrated (Mirena)	5	2	–	–	–
Removed before 3 months (Mirena)	5	4	–	–	–

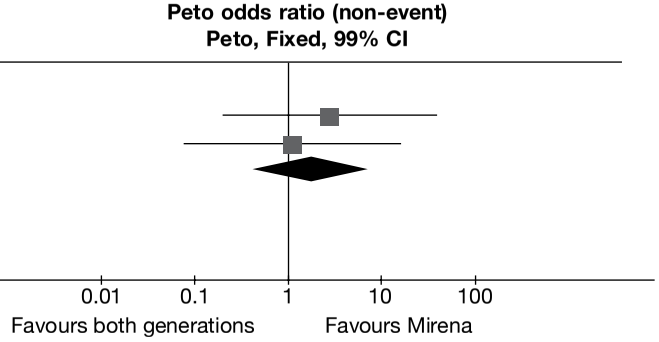
a < 0 favours first-generation EA, > 0 Mirena.
b < 1 favours first-generation EA, > 1 Mirena.

DOI: 10.3310/hta15190

Appendix 7.2 First- and second-generation endometrial ablation versus Mirena

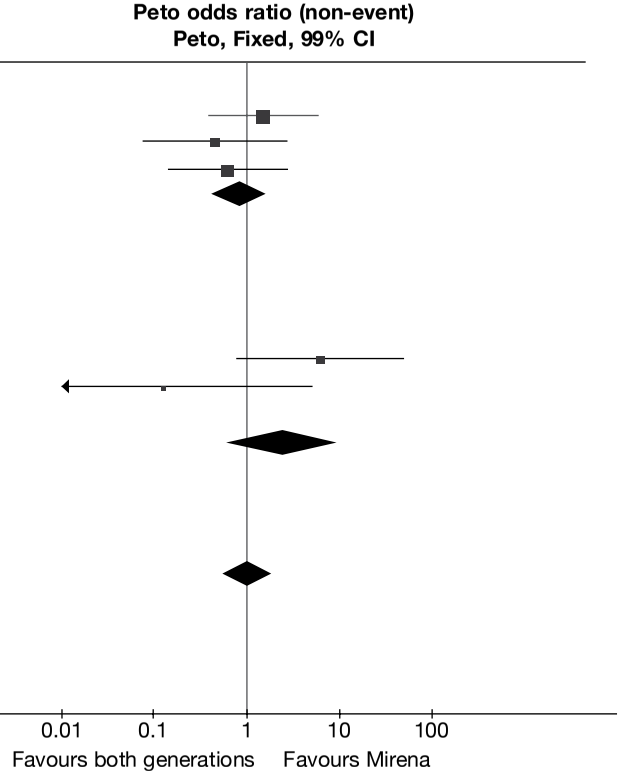
Proportion with amenorrhoea < 12 months

	Both generations		Mirena			
Study or subgroup	**Events**	**Total**	**Events**	**Total**	**Weight**	**Peto odds ratio (non-event) Peto, Fixed, 99% CI**
1.1.2 Second generation						
f. Busfield 2006[107]	1	31	3	32	50.6%	2.77 (0.20 to 38.81)
g. Barrington 2003[81]	2	23	2	21	49.4%	1.10 (0.08 to 15.98)
Subtotal (95% CI)		**54**		**53**	**100.0%**	**1.76 (0.42 to 7.34)**
Total events	3		5			
Heterogeneity: $\chi^2 = 0.40$, df = 1 ($p = 0.53$); $I^2 = 0\%$						
Test for overall effect: $z = 0.77$ ($p = 0.44$)						



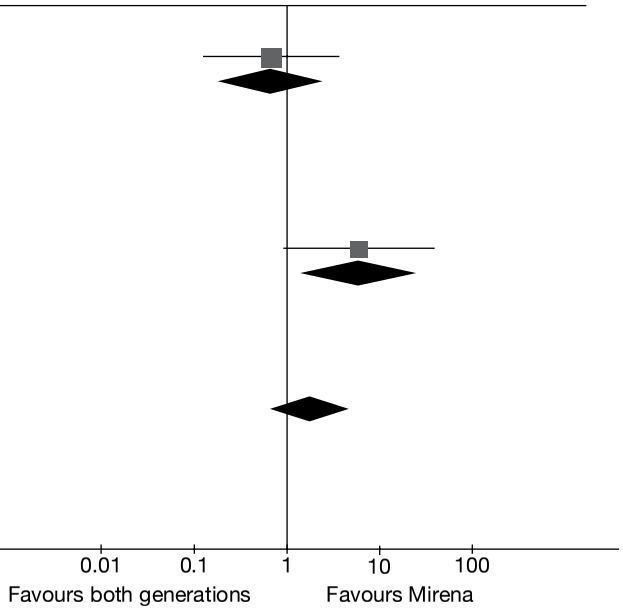
Proportion with amenorrhoea – 12 months

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio (non-event) Peto, Fixed, 99% CI
1.2.1 First generation						
a. Malak 2006[104]	13	30	14	26	33.5%	1.51 (0.38 to 5.94)
b. Kittelsen 1998[105]	7	28	3	24	19.4%	0.45 (0.08 to 2.75)
c. Crosignani 1997[45]	9	35	6	34	28.2%	0.63 (0.14 to 2.79)
Subtotal (95% CI)		**93**		**84**	**81.0%**	**0.84 (0.43 to 1.63)**
Total events	29		23			
Heterogeneity: $\chi^2 = 2.25$, df = 2 ($p = 0.32$); $I^2 = 11\%$						
Test for overall effect: $z = 0.53$ ($p = 0.60$)						
1.2.3 Second generation						
d. Shaw 2007[108]	1	27	6	23	14.4%	6.18 (0.77 to 49.76)
e. Tam 2006[109]	2	15	0	15	4.6%	0.13 (0.00 to 5.13)
f. Busfield 2006[107]	0	25	0	22		Not estimable
Subtotal (95% CI)		**67**		**60**	**19.0%**	**2.43 (0.61 to 9.67)**
Total events	3		6			
Heterogeneity: $\chi^2 = 5.56$, df = 1 ($p = 0.02$); $I^2 = 82\%$						
Test for overall effect: $z = 1.26$ ($p = 0.21$)						
Total (95% CI)		**160**		**144**	**100.0%**	**1.02 (0.56 to 1.87)**
Total events	32		29			
Heterogeneity: $\chi^2 = 9.66$, df = 4 ($p = 0.05$); $I^2 = 59\%$						
Test for overall effect: $z = 0.07$ ($p = 0.94$)						
Test for subgroup differences: $\chi^2 = 1.85$, df = 1 ($p = 0.17$), $I^2 = 45.9\%$						



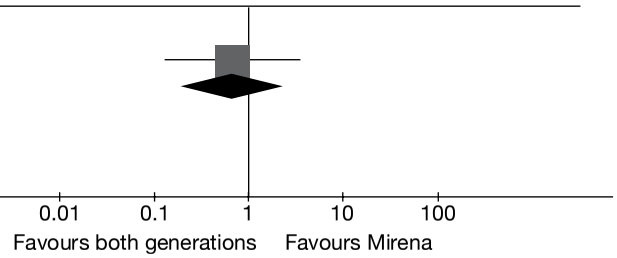
Proportion with amenorrhoea – 2 years

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio (non-event) Peto, Fixed, 99% CI
1.3.3 First generation						
b. Kittelsen 1998[105]	8	24	5	20	55.3%	0.68 (0.12 to 3.66)
Subtotal (95% CI)		**24**		**20**	**55.3%**	**0.68 (0.19 to 2.45)**
Total events	8		5			
Heterogeneity: not applicable						
Test for overall effect: $z = 0.60$ ($p = 0.55$)						
1.3.4 Second generation						
f. Busfield 2006[107]	1	17	9	22	44.7%	5.99 (0.91 to 39.27)
Subtotal (95% CI)		**17**		**22**	**44.7%**	**5.99 (1.43 to 25.05)**
Total events	1		9			
Heterogeneity: not applicable						
Test for overall effect: $z = 2.45$ ($p = 0.01$)						
Total (95% CI)		**41**		**42**	**100.0%**	**1.79 (0.69 to 4.66)**
Total events	9		14			
Heterogeneity: $\chi^2 = 4.94$, df = 1 ($p = 0.03$); $I^2 = 80\%$						
Test for overall effect: $z = 1.20$ ($p = 0.23$)						
Test for subgroup differences: $\chi^2 = 4.94$, df = 1 ($p = 0.03$), $I^2 = 79.8\%$						



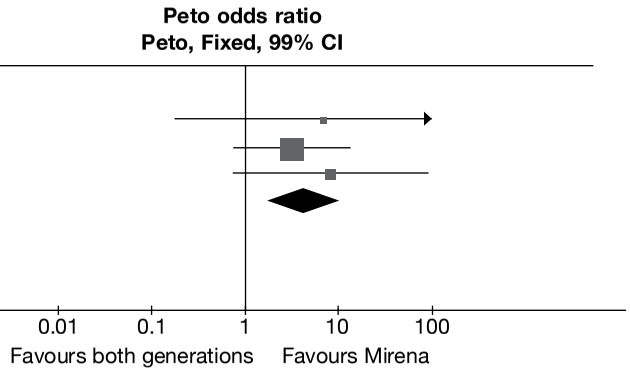
Proportion with amenorrhoea – 3 years

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio (non-event) Peto, Fixed, 99% CI
1.4.4 First generation						
b. Kittelsen 1998[105]	9	22	6	19	100.0%	0.68 (0.13 to 3.53)
Subtotal (95% CI)		**22**		**19**	**100.0%**	**0.68 (0.19 to 2.38)**
Total events	9		6			
Heterogeneity: not applicable						
Test for overall effect: $z = 0.61$ ($p = 0.54$)						



Proportion with heavy bleeding – < 12 months

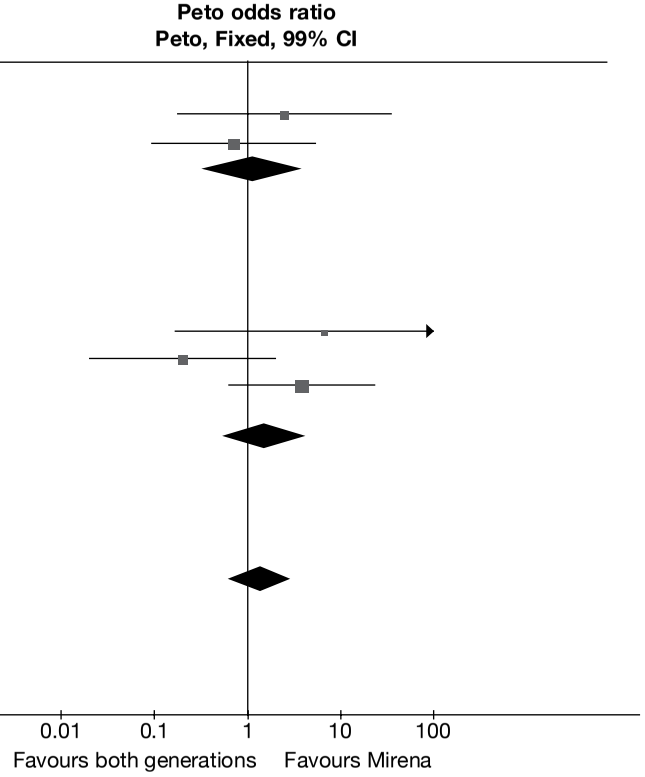
Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.5.2 Second generation						
d. Shaw 2007[108]	2	29	0	26	10.3%	6.91 (0.17 to 274.46)
f. Busfield 2006[107]	12	31	5	32	66.0%	3.17 (0.74 to 13.52)
g. Barrington 2003[81]	5	23	0	21	23.7%	8.24 (0.73 to 92.84)
Subtotal (95% CI)		**83**		**79**	**100.0%**	**4.30 (1.76 to 10.56)**
Total events	19		5			
Heterogeneity: $\chi^2 = 0.88$, df = 2 ($p = 0.64$); $I^2 = 0\%$						
Test for overall effect: $z = 3.19$ ($p = 0.001$)						



DOI: 10.3310/hta15190

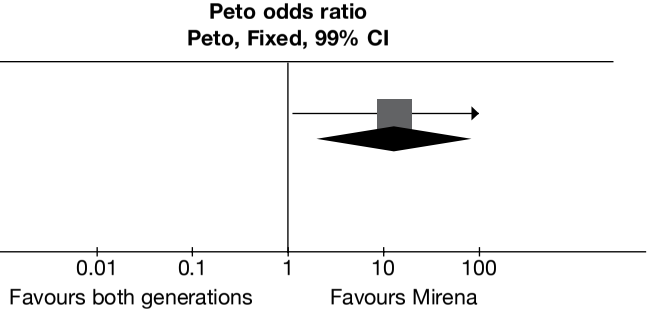
Proportion with heavy bleeding – 12 months

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.6.1 First generation						
a. Malak 2006[104]	3	30	1	26	15.0%	2.49 (0.17 to 35.41)
c. Crosignani 1997[45]	3	35	4	34	25.5%	0.71 (0.09 to 5.44)
Subtotal (95% CI)		**65**		**60**	**40.5%**	**1.13 (0.33 to 3.86)**
Total events	6		5			
Heterogeneity: $\chi^2 = 0.93$, df = 1 ($p = 0.33$); $I^2 = 0\%$						
Test for overall effect: $z = 0.19$ ($p = 0.85$)						
1.6.4 Second generation						
d. Shaw 2007[108]	2	27	0	23	7.8%	6.62 (0.16 to 265.81)
e. Tam 2006[109]	1	15	5	15	19.8%	0.20 (0.02 to 2.02)
f. Busfield 2006[107]	8	25	2	22	32.0%	3.81 (0.62 to 23.54)
h. Soysal 2001[96]	0	35	0	32		Not estimable
Subtotal (95% CI)		**102**		**92**	**59.5%**	**1.54 (0.56 to 4.24)**
Total events	11		7			
Heterogeneity: $\chi^2 = 7.86$, df = 2 ($p = 0.02$); $I^2 = 75\%$						
Test for overall effect: $z = 0.83$ ($p = 0.41$)						
Total (95% CI)		**167**		**152**	**100.0%**	**1.36 (0.62 to 2.97)**
Total events	17		12			
Heterogeneity: $\chi^2 = 8.94$, df = 4 ($p = 0.06$); $I^2 = 55\%$						
Test for overall effect: $z = 0.76$ ($p = 0.45$)						
Test for subgroup differences: $\chi^2 = 0.14$, df = 1 ($p = 0.70$), $I^2 = 0\%$						



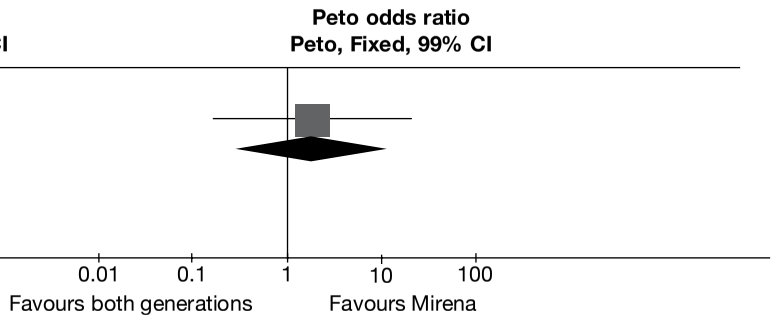
Proportion with heavy bleeding – 2 years

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.7.5 Second generation						
f. Busfield 2006[107]	5	17	0	22	100.0%	12.99 (1.11 to 151.40)
Subtotal (95% CI)		**17**		**22**	**100.0%**	**12.99 (2.00 to 84.16)**
Total events	5		0			
Heterogeneity: not applicable						
Test for overall effect: $z = 2.69$ ($p = 0.007$)						



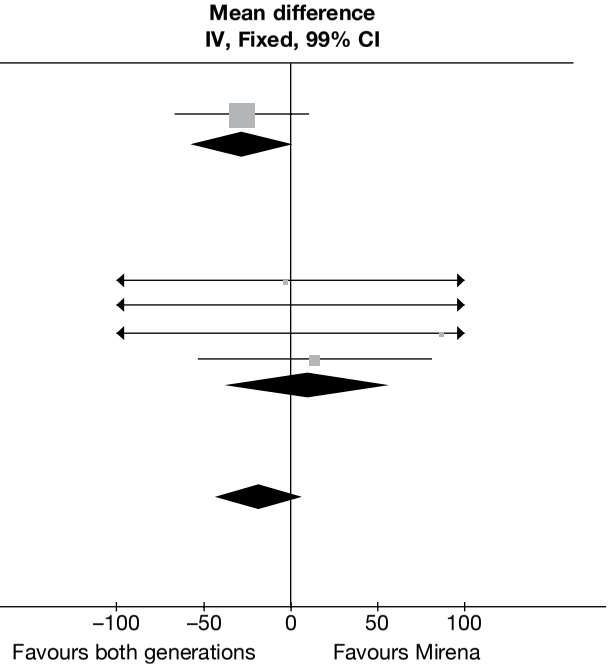
Proportion with heavy bleeding – 3 years

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.8.2 First generation						
b. Kittelsen 1998[105]	3	19	2	22	100.0%	1.84 (0.16 to 21.02)
Subtotal (95% CI)		**19**		**22**	**100.0%**	**1.84 (0.29 to 11.74)**
Total events	3		2			
Heterogeneity: not applicable						
Test for overall effect: $z = 0.65$ ($p = 0.52$)						



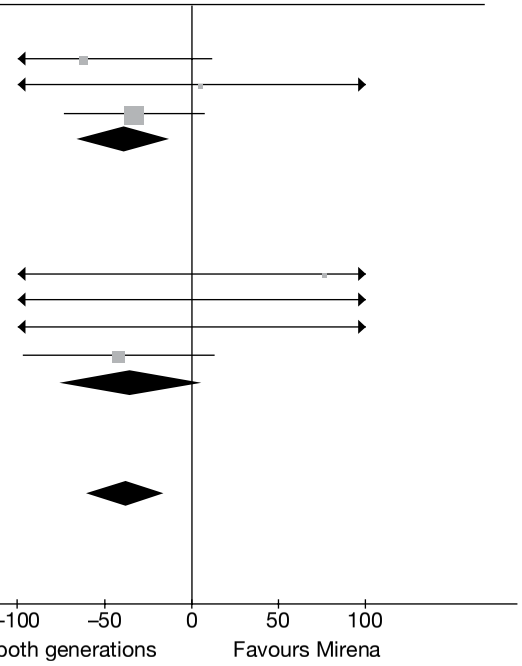
Bleeding/pictorial blood loss assessment chart scores (change from baseline) – < 12 months

	Both generations			Mirena				
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**Mean difference IV, Fixed, 99% CI**
1.17.1 First generation								
c. Crosignani 1997[45]	–173	68	34	–145	55	34	72.1%	–28.00 (–66.63 to 10.63)
Subtotal (95% CI)			**34**			**34**	**72.1%**	**–28.00 (–57.40 to 1.40)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.87$ ($p = 0.06$)								
1.17.2 Second generation								
d. Shaw 2007[108]	–329	376	27	–326	230	29	2.3%	–3.00 (–219.44 to 213.44)
e. Tam 2006[109]	–431	168	3	–92	265	3	0.5%	–339.00 (–805.62 to 127.62)
f. Busfield 2006[107]	–364	460	31	–451	410	32	1.3%	87.00 (–196.09 to 370.09)
g. Barrington 2003[81]	–62	47	23	–76	111	21	23.8%	14.00 (–53.31 to 81.31)
Subtotal (95% CI)			**84**			**85**	**27.9%**	**9.86 (–37.41 to 57.12)**
Heterogeneity: $\chi^2 = 4.25$, df = 3 ($p = 0.24$); $I^2 = 29\%$								
Test for overall effect: $z = 0.41$ ($p = 0.68$)								
Total (95% CI)			**118**			**119**	**100.0%**	**–17.44 (–42.40 to 7.52)**
Heterogeneity: $\chi^2 = 6.03$, df = 4 ($p = 0.20$); $I^2 = 34\%$								
Test for overall effect: $z = 1.37$ ($p = 0.17$)								
Test for subgroup differences: $\chi^2 = 1.78$, df = 1 ($p = 0.18$), $I^2 = 43.7\%$								



Bleeding/pictorial blood loss assessment chart scores (change from baseline) – 12 months

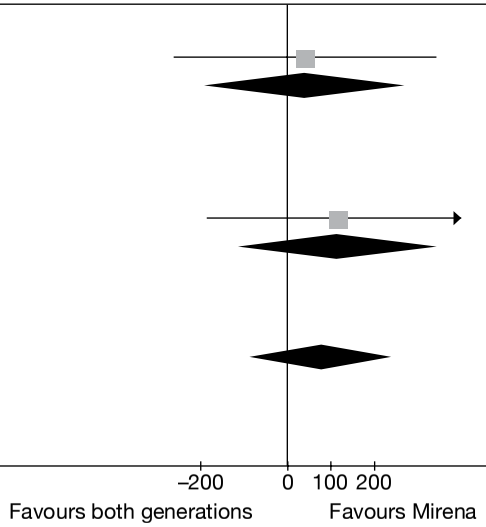
	Both generations			Mirena					
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**Mean difference IV, Fixed, 99% CI**	**Mean difference IV, Fixed, 99% CI**
1.18.2 First generation									
a. Malak 2006[104]	–340	111	30	–278	103	26	15.8%	–62.00 (–135.70 to 11.70)	
b. Kittelsen 1998[105]	–371	456	28	–376	311	24	1.1%	5.00 (–270.70 to 280.70)	
c. Crosignani 1997[45]	–179	67	30	–146	54	30	52.4%	–33.00 (–73.47 to 7.47)	
Subtotal (95% CI)			**88**			**80**	**69.4%**	**–38.99 (–65.76 to –12.22)**	
Heterogeneity: $\chi^2 = 0.96$, df = 2 ($p = 0.62$); $I^2 = 0\%$									
Test for overall effect: $z = 2.85$ ($p = 0.004$)									
1.18.4 Second generation									
d. Shaw 2007[108]	–348	402	20	–424	256	21	1.2%	76.00 (–196.61 to 348.61)	
e. Tam 2006[109]	–504	572	9	–216	269	5	0.3%	–288.00 (–868.71 to 292.71)	
f. Busfield 2006[107]	–360	433	25	–474	375	22	0.9%	114.00 (–189.59 to 417.59)	
h. Soysal 2001[96]	–395	75	35	–353	96	31	28.3%	–42.00 (–97.13 to 13.13)	
Subtotal (95% CI)			**89**			**79**	**30.6%**	**–34.84 (–75.15 to 5.47)**	
Heterogeneity: $\chi^2 = 4.06$, df = 3 ($p = 0.25$); $I^2 = 26\%$									
Test for overall effect: $z = 1.69$ ($p = 0.09$)									
Total (95% CI)			**177**			**159**	**100.0%**	**–37.72 (–60.02 to –15.42)**	
Heterogeneity: $\chi^2 = 5.05$, df = 6 ($p = 0.54$); $I^2 = 0\%$									
Test for overall effect: $z = 3.32$ ($p = 0.0009$)									
Test for subgroup differences: $\chi^2 = 0.03$, df = 1 ($p = 0.87$), $I^2 = 0\%$									



DOI: 10.3310/hta15190

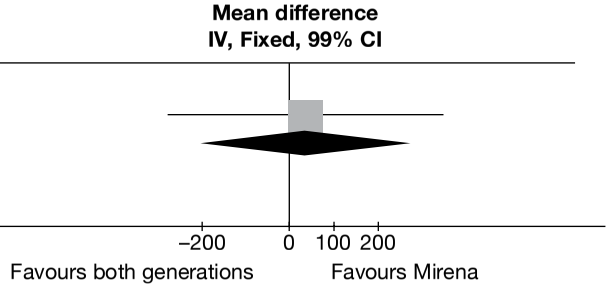
Bleeding/pictorial blood loss assessment chart scores (change from baseline) – 2 years

	Both generations			Mirena				Mean difference	Mean difference
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 99% CI**	**IV, Fixed, 99% CI**
1.19.3 First generation									
b. Kittelsen 1998[105]	–368	443	24	–409	334	20	49.9%	41.00 (–261.10 to 343.10)	
Subtotal (95% CI)			**24**			**20**	**49.9%**	**41.00 (–188.87 to 270.87)**	
Heterogeneity: not applicable									
Test for overall effect: $z = 0.35$ ($p = 0.73$)									
1.19.5 Second generation									
f. Busfield 2006[107]	–354	283	17	–471	445	22	50.1%	117.00 (–184.63 to 418.63)	
Subtotal (95% CI)			**17**			**22**	**50.1%**	**117.00 (–112.51 to 346.51)**	
Heterogeneity: not applicable									
Test for overall effect: $z = 1.00$ ($p = 0.32$)									
Total (95% CI)			**41**			**42**	**100.0%**	**79.06 (–83.36 to 241.47)**	
Heterogeneity: $\chi^2 = 0.21$, df = 1 ($p = 0.65$); $I^2 = 0\%$									
Test for overall effect: $z = 0.95$ ($p = 0.34$)									
Test for subgroup differences: $\chi^2 = 0.21$, df = 1 ($p = 0.65$), $I^2 = 0\%$									



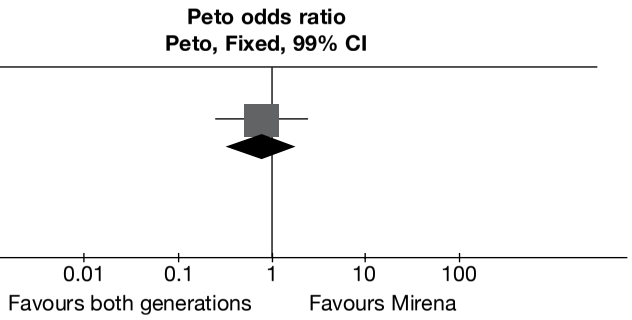
Bleeding/pictorial blood loss assessment chart scores (change from baseline) – 3 years

Study or subgroup	Both generations Mean	SD	Total	Mirena Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.20.4 First generation								
b. Kittelsen 1998[105]	–374	442	22	–411	337	19	100.0%	37.00 (–276.97 to 350.97)
Subtotal (95% CI)			**22**			**19**	**100.0%**	**37.00 (–201.90 to 275.90)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.30$ ($p = 0.76$)								



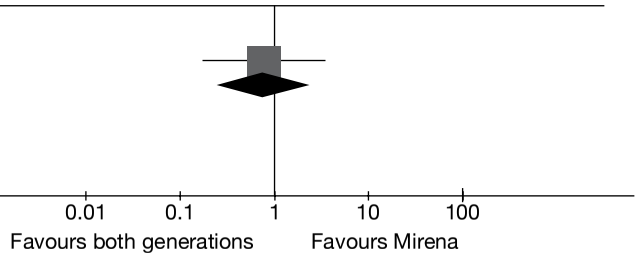
Proportion with dysmenorrhoea – <12 months

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.21.1 Second generation						
f. Busfield 2006[107]	16	41	19	42	100.0%	0.78 (0.25 to 2.43)
Subtotal (95% CI)		**41**		**42**	**100.0%**	**0.78 (0.33 to 1.85)**
Total events	16		19			
Heterogeneity: not applicable						
Test for overall effect: $z = 0.57$ ($p = 0.57$)						



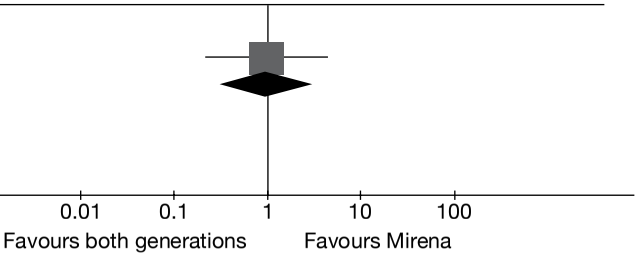
Proportion with dysmenorrhoea – 12 months

	Both generations		Mirena			Peto odds ratio	Peto odds ratio
Study or subgroup	**Events**	**Total**	**Events**	**Total**	**Weight**	**Peto, Fixed, 99% CI**	**Peto, Fixed, 99% CI**
1.22.3 Second generation							
f. Busfield 2006[107]	8	21	12	27	100.0%	0.77 (0.17 to 3.49)	
Subtotal (95% CI)		**21**		**27**	**100.0%**	**0.77 (0.25 to 2.43)**	
Total events	8		12				
Heterogeneity: not applicable							
Test for overall effect: $z = 0.44$ ($p = 0.66$)							



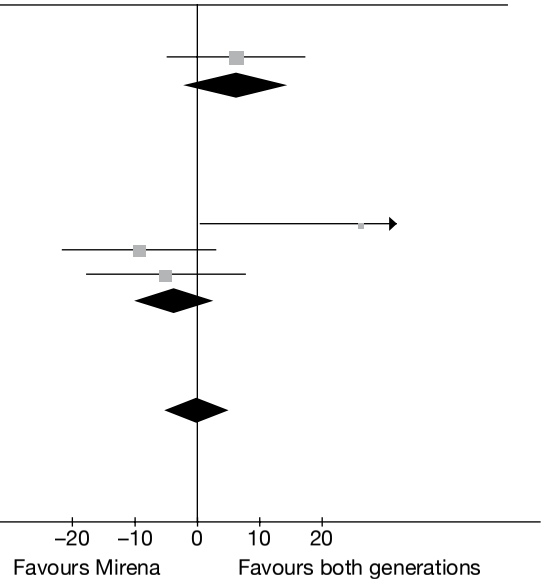
Proportion with dysmenorrhoea – 2 years

	Both generations		Mirena			Peto odds ratio	Peto odds ratio
Study or subgroup	**Events**	**Total**	**Events**	**Total**	**Weight**	**Peto, Fixed, 99% CI**	**Peto, Fixed, 99% CI**
1.23.4 Second generation							
f. Busfield 2006[107]	7	42	7	41	100.0%	0.97 (0.22 to 4.36)	
Subtotal (95% CI)		**42**		**41**	**100.0%**	**0.97 (0.31 to 3.05)**	
Total events	7		7				
Heterogeneity: not applicable							
Test for overall effect: $z = 0.05$ ($p = 0.96$)							



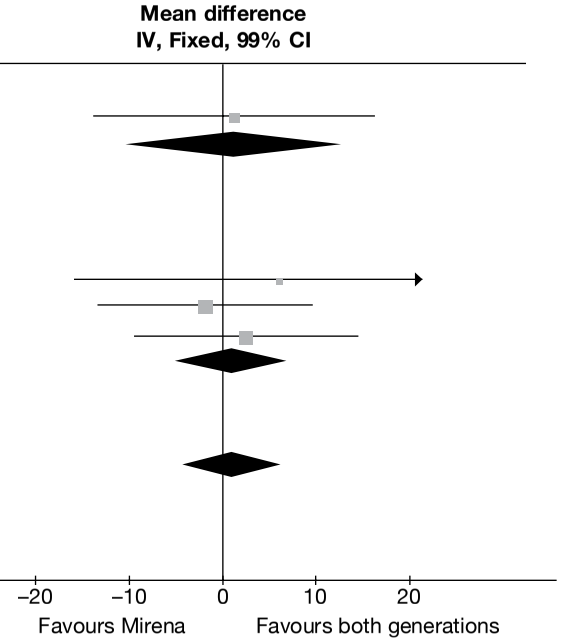
SF-36 general health (absolute values) – 12 months

Study or subgroup	Both generations Mean	SD	Total	Mirena Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.24.1 First generation								
c. Crosignani 1997[45]	70.3	15.1	31	64.1	18.6	31	36.5%	6.20 (–4.88 to 17.28)
Subtotal (95% CI)			**31**			**31**	**36.5%**	**6.20 (–2.23 to 14.63)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.44$ ($p = 0.15$)								
1.24.2 Second generation								
e. Tam 2006[109]	60.3	26.2	9	34.1	15.5	10	6.7%	26.20 (0.40 to 52.00)
f. Busfield 2006[107]	73.3	22.3	32	82.6	15.2	31	29.4%	–9.30 (–21.65 to 3.05)
h. Soysal 2002[106]	47	20	33	52	20	32	27.4%	–5.00 (–17.78 to 7.78)
Subtotal (95% CI)			**74**			**73**	**63.5%**	**–3.68 (–10.07 to 2.71)**
Heterogeneity: $\chi^2 = 10.35$, df = 2 ($p = 0.006$); $I^2 = 81\%$								
Test for overall effect: $z = 1.13$ ($p = 0.26$)								
Total (95% CI)			**105**			**104**	**100.0%**	**–0.08 (–5.17 to 5.02)**
Heterogeneity: $\chi^2 = 13.70$, df = 3 ($p = 0.003$); $I^2 = 78\%$								
Test for overall effect: $z = 0.03$ ($p = 0.98$)								
Test for subgroup differences: $\chi^2 = 3.35$, df = 1 ($p = 0.07$), $I^2 = 70.2\%$								



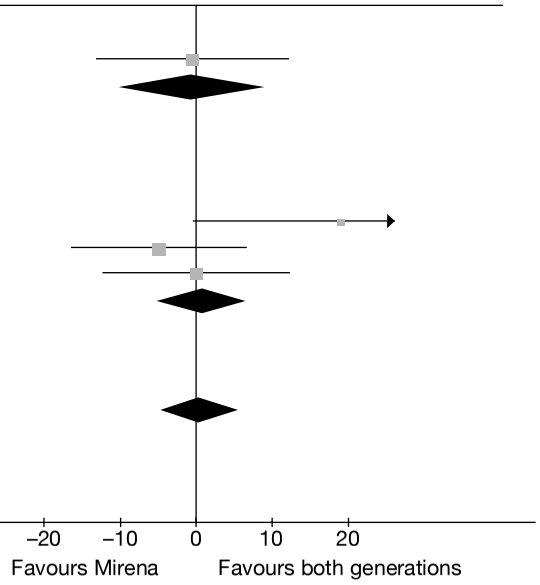
SF-36 physical function (absolute values) – 12 months

Study or subgroup	Both generations Mean	SD	Total	Mirena Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.25.1 First generation								
c. Crosignani 1997[45]	79.2	23.7	31	78	22.4	31	21.0%	1.20 (–13.89 to 16.29)
Subtotal (95% CI)			**31**			**31**	**21.0%**	**1.20 (–10.28 to 12.68)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.20$ ($p = 0.84$)								
1.25.2 Second generation								
e. Tam 2006[109]	76.1	18.8	9	70	17.5	9	9.8%	6.10 (–15.95 to 28.15)
f. Busfield 2006[107]	88.1	20	32	90	15.3	31	36.0%	–1.90 (–13.43 to 9.63)
j. Soysal 2002[106]	75	18.8	33	72.5	18.8	32	33.2%	2.50 (–9.51 to 14.51)
Subtotal (95% CI)			**74**			**72**	**79.0%**	**0.94 (–4.98 to 6.87)**
Heterogeneity: $\chi^2 = 0.88$, df = 2 ($p = 0.64$); $I^2 = 0\%$								
Test for overall effect: $z = 0.31$ ($p = 0.75$)								
Total (95% CI)			**105**			**103**	**100.0%**	**1.00 (–4.27 to 6.26)**
Heterogeneity: $\chi^2 = 0.88$, df = 3 ($p = 0.83$); $I^2 = 0\%$								
Test for overall effect: $z = 0.37$ ($p = 0.71$)								
Test for subgroup differences: $\chi^2 = 0.00$, df = 1 ($p = 0.97$), $I^2 = 0\%$								



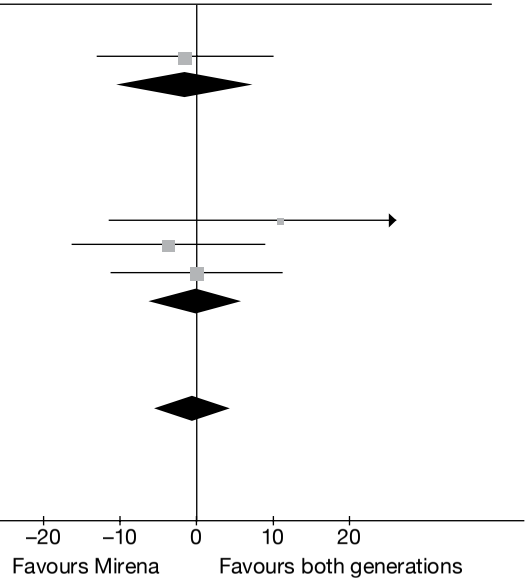
SF-36 mental health (absolute values) – 12 months

	Both generations			Mirena				Mean difference	Mean difference
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 99% CI**	**IV, Fixed, 99% CI**
1.26.1 First generation									
c. Crosignani 1997[45]	59.6	20.5	31	60.1	18.2	31	27.2%	–0.50 (–13.18 to 12.18)	
Subtotal (95% CI)			**31**			**31**	**27.2%**	**–0.50 (–10.15 to 9.15)**	
Heterogeneity: not applicable									
Test for overall effect: $z = 0.10$ ($p = 0.92$)									
1.26.3 Second generation									
e. Tam 2006[109]	65.5	13.3	8	46.4	18.8	10	11.5%	19.10 (–0.42 to 38.62)	
f. Busfield 2006[107]	75	19.5	32	79.9	16	31	32.7%	–4.90 (–16.46 to 6.66)	
j. Soysal 2002[106]	52	19.3	33	52	19.3	32	28.7%	0.00 (–12.33 to 12.33)	
Subtotal (95% CI)			**73**			**73**	**72.8%**	**0.81 (–5.09 to 6.70)**	
Heterogeneity: $\chi^2 = 7.47$, df = 2 ($p = 0.02$); $I^2 = 73\%$									
Test for overall effect: $z = 0.27$ ($p = 0.79$)									
Total (95% CI)			**104**			**104**	**100.0%**	**0.45 (–4.58 to 5.48)**	
Heterogeneity: $\chi^2 = 7.52$, df = 3 ($p = 0.06$); $I^2 = 60\%$									
Test for overall effect: $z = 0.18$ ($p = 0.86$)									
Test for subgroup differences: $\chi^2 = 0.05$, df = 1 ($p = 0.82$), $I^2 = 0\%$									



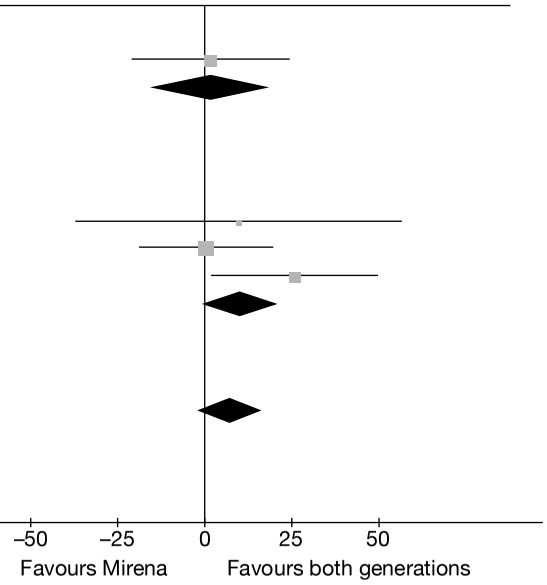
SF-36 vitality (absolute values) – 12 months

	Both generations			Mirena					
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**Mean difference IV, Fixed, 99% CI**	**Mean difference IV, Fixed, 99% CI**
1.27.1 First generation									
c. Crosignani 1997[45]	54.8	20.7	31	56.3	14.1	31	31.6%	–1.50 (–13.09 to 10.09)	
Subtotal (95% CI)			**31**			**31**	**31.6%**	**–1.50 (–10.32 to 7.32)**	
Heterogeneity: not applicable									
Test for overall effect: $z = 0.33$ ($p = 0.74$)									
1.27.4 Second generation									
e. Tam 2006[109]	50	22.2	9	39	14.7	10	8.4%	11.00 (–11.51 to 33.51)	
f. Busfield 2006[107]	63	20.7	32	66.7	18	30	26.4%	–3.70 (–16.37 to 8.97)	
j. Soysal 2002[106]	45	17.6	33	45	17.6	32	33.6%	0.00 (–11.25 to 11.25)	
Subtotal (95% CI)			**74**			**72**	**68.4%**	**–0.08 (–6.08 to 5.91)**	
Heterogeneity: $\chi^2 = 2.15$, df = 2 ($p = 0.34$); $I^2 = 7\%$									
Test for overall effect: $z = 0.03$ ($p = 0.98$)									
Total (95% CI)			**105**			**103**	**100.0%**	**–0.53 (–5.49 to 4.43)**	
Heterogeneity: $\chi^2 = 2.22$, df = 3 ($p = 0.53$); $I^2 = 0\%$									
Test for overall effect: $z = 0.21$ ($p = 0.83$)									
Test for subgroup differences: $\chi^2 = 0.07$, df = 1 ($p = 0.79$), $I^2 = 0\%$									



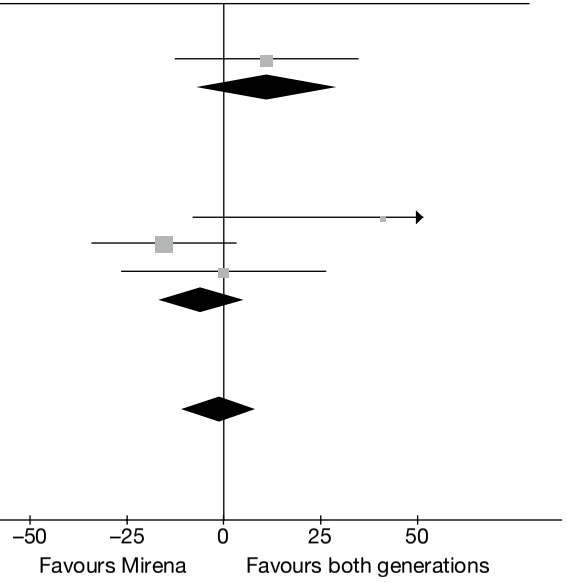
SF-36 physical role limitation (absolute values) – 12 months

Study or subgroup	Both generations Mean	SD	Total	Mirena Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.28.1 First generation								
c. Crosignani 1997[45]	74.2	35.6	31	72.5	33.7	31	28.4%	1.70 (–20.98 to 24.38)
Subtotal (95% CI)			**31**			**31**	**28.4%**	**1.70 (–15.56 to 18.96)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.19$ ($p = 0.85$)								
1.28.5 Second generation								
e. Tam 2006[109]	72.2	36.3	9	62.5	42.9	10	6.7%	9.70 (–37.12 to 56.52)
f. Busfield 2006[107]	87.5	30.5	32	87.9	28.8	31	39.5%	–0.40 (–19.65 to 18.85)
j. Soysal 2002[106]	50	37.5	33	25	37.5	32	25.5%	25.00 (1.04 to 48.96)
Subtotal (95% CI)			**74**			**73**	**71.6%**	**9.57 (–1.30 to 20.45)**
Heterogeneity: $\chi^2 = 4.53$, df = 2 ($p = 0.10$); $I^2 = 56\%$								
Test for overall effect: $z = 1.73$ ($p = 0.08$)								
Total (95% CI)			**105**			**104**	**100.0%**	**7.33 (–1.86 to 16.53)**
Heterogeneity: $\chi^2 = 5.10$, df = 3 ($p = 0.16$); $I^2 = 41\%$								
Test for overall effect: $z = 1.56$ ($p = 0.12$)								
Test for subgroup differences: $\chi^2 = 0.57$, df = 1 ($p = 0.45$), $I^2 = 0\%$								



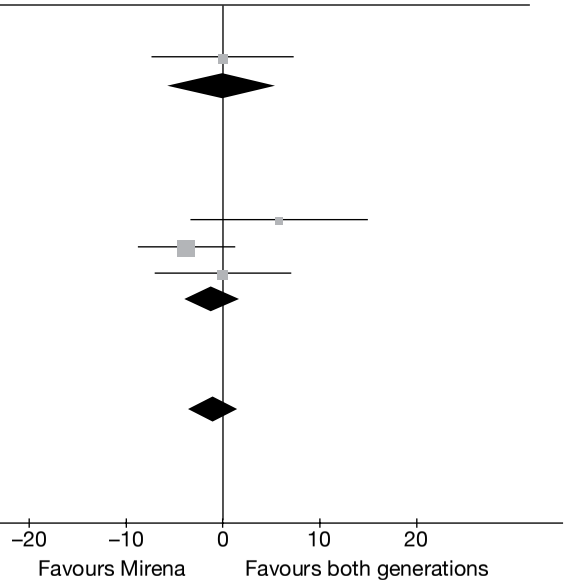
SF-36 emotional role limitation (absolute values) – 12 months

	Both generations			Mirena				Mean difference
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 99% CI**
1.29.1 First generation								
c. Crosignani 1997[45]	72.4	36.8	31	61.3	35.6	31	27.5%	11.10 (–12.59 to 34.79)
Subtotal (95% CI)			**31**			**31**	**27.5%**	**11.10 (–6.92 to 29.12)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.21$ ($p = 0.23$)								
1.29.6 Second generation								
e. Tam 2006[109]	77.8	37.3	9	36.7	45.7	10	6.4%	41.10 (–8.01 to 90.21)
f. Busfield 2006[107]	76	36.2	32	91.4	19.2	31	44.0%	–15.40 (–34.12 to 3.32)
j. Soysal 2002[106]	33.3	41.4	33	33.3	41.4	32	22.1%	0.00 (–26.46 to 26.46)
Subtotal (95% CI)			**74**			**73**	**72.5%**	**–5.72 (–16.83 to 5.38)**
Heterogeneity: $\chi^2 = 8.11$, df = 2 ($p = 0.02$); $I^2 = 75\%$								
Test for overall effect: $z = 1.01$ ($p = 0.31$)								
Total (95% CI)			**105**			**104**	**100.0%**	**–1.10 (–10.55 to 8.36)**
Heterogeneity: $\chi^2 = 10.54$, df = 3 ($p = 0.01$); $I^2 = 72\%$								
Test for overall effect: $z = 0.23$ ($p = 0.82$)								
Test for subgroup differences: $\chi^2 = 2.43$, df = 1 ($p = 0.12$), $I^2 = 58.8\%$								



SF-36 social function (absolute values) – 12 months

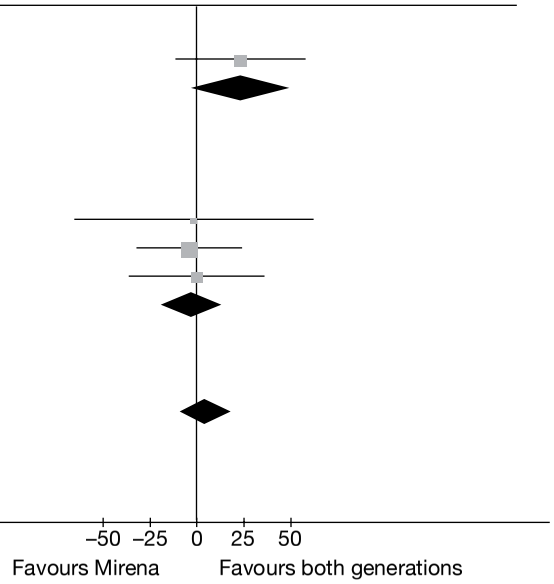
	Both generations			Mirena				
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**Mean difference IV, Fixed, 99% CI**
1.30.1 First generation								
c. Crosignani 1997[45]	69.7	24.1	31	69.8	22.3	31	20.6%	–0.10 (–15.29 to 15.09)
Subtotal (95% CI)			**31**			**31**	**20.6%**	**–0.10 (–11.66 to 11.46)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.02$ ($p = 0.99$)								
1.30.7 Second generation								
e. Tam 2006[109]	75.3	15.5	9	63.3	16.6	10	13.2%	12.00 (–6.97 to 30.97)
f. Busfield 2006[107]	73.6	19.3	32	81.4	12	31	44.0%	–7.80 (–18.19 to 2.59)
j. Soysal 2002[106]	50	22.9	33	50	22.9	32	22.2%	0.00 (–14.63 to 14.63)
Subtotal (95% CI)			**74**			**73**	**79.4%**	**–2.33 (–8.21 to 3.56)**
Heterogeneity: $\chi^2 = 5.79$, df = 2 ($p = 0.06$); $I^2 = 65\%$								
Test for overall effect: $z = 0.77$ ($p = 0.44$)								
Total (95% CI)			**105**			**104**	**100.0%**	**–1.87 (–7.11 to 3.38)**
Heterogeneity: $\chi^2 = 5.90$, df = 3 ($p = 0.12$); $I^2 = 49\%$								
Test for overall effect: $z = 0.70$ ($p = 0.49$)								
Test for subgroup differences: $\chi^2 = 0.11$, df = 1 ($p = 0.74$), $I^2 = 0\%$								



DOI: 10.3310/hta15190

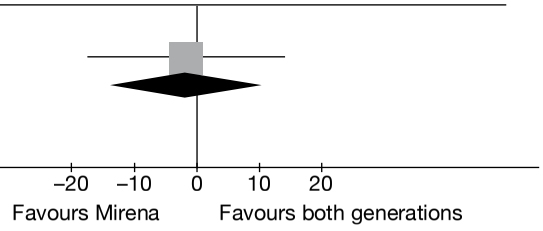
SF-36 bodily pain (absolute values) – 12 months

Study or subgroup	Both generations Mean	SD	Total	Mirena Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.31.1 First generation								
c. Crosignani 1997[45]	70.3	23.3	31	58.9	28	31	26.8%	11.40 (–5.45 to 28.25)
Subtotal (95% CI)			**31**			**31**	**26.8%**	**11.40 (–1.42 to 24.22)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.74$ ($p = 0.08$)								
1.31.8 Second generation								
e. Tam 2006[109]	70.4	24.2	9	71.1	28.3	10	7.9%	–0.70 (–31.73 to 30.33)
f. Busfield 2006[107]	79.5	23.3	32	81.4	18.7	31	40.6%	–1.90 (–15.59 to 11.79)
j. Soysal 2002[106]	51	27.5	33	51	27.5	32	24.7%	0.00 (–17.57 to 17.57)
Subtotal (95% CI)			**74**			**73**	**73.2%**	**–1.13 (–8.89 to 6.63)**
Heterogeneity: $\chi^2 = 0.05$, df = 2 ($p = 0.98$); $I^2 = 0\%$								
Test for overall effect: $z = 0.29$ ($p = 0.78$)								
Total (95% CI)			**105**			**104**	**100.0%**	**2.23 (–4.41 to 8.87)**
Heterogeneity: $\chi^2 = 2.73$, df = 3 ($p = 0.43$); $I^2 = 0\%$								
Test for overall effect: $z = 0.66$ ($p = 0.51$)								
Test for subgroup differences: $\chi^2 = 2.68$, df = 1 ($p = 0.10$), $I^2 = 62.8\%$								



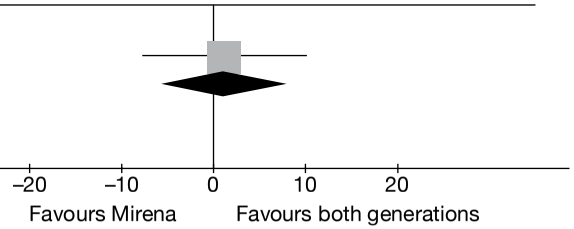
SF-36 general health (absolute values) – 2 years

Study or subgroup	Both generations Mean	SD	Total	Mirena Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.32.2 Second generation								
f. Busfield 2006[107]	77.2	23.2	22	78.9	18.9	27	100.0%	–1.70 (–17.51 to 14.11)
Subtotal (95% CI)			**22**			**27**	**100.0%**	**–1.70 (–13.73 to 10.33)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.28$ ($p = 0.78$)								



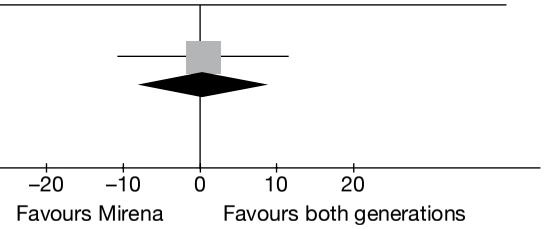
SF-36 physical function (absolute values) – 2 years

Study or subgroup	Both generations Mean	SD	Total	Mirena Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.33.2 Second generation								
f. Busfield 2006[107]	92.3	9	22	91.1	15.3	28	100.0%	1.20 (–7.74 to 10.14)
Subtotal (95% CI)			**22**			**28**	**100.0%**	**1.20 (–5.60 to 8.00)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.35$ ($p = 0.73$)								



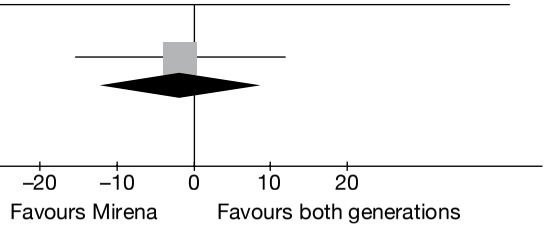
SF-36 mental health (absolute values) – 2 years

Study or subgroup	Both generations Mean	SD	Total	Mirena Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.34.3 Second generation								
f. Busfield 2006[107]	79.4	15.2	21	79	14.8	28	100.0%	0.40 (–10.78 to 11.58)
Subtotal (95% CI)			**21**			**28**	**100.0%**	**0.40 (–8.10 to 8.90)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.09$ ($p = 0.93$)								



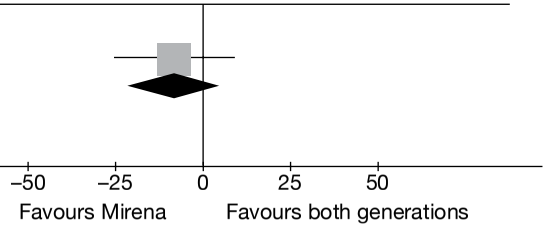
SF-36 vitality (absolute values) – 2 years

Study or subgroup	Both generations Mean	SD	Total	Mirena Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.35.4 Second generation								
f. Busfield 2006[107]	63.6	19.9	21	65.4	16.3	28	100.0%	–1.80 (–15.51 to 11.91)
Subtotal (95% CI)			**21**			**28**	**100.0%**	**–1.80 (–12.24 to 8.64)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.34$ ($p = 0.74$)								



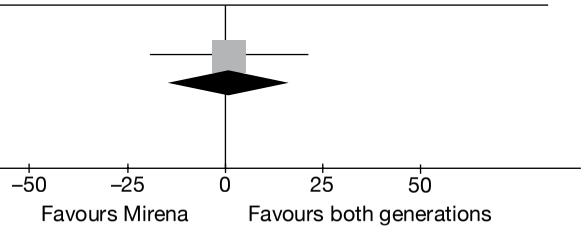
SF-36 physical role limitation (absolute values) – 2 years

Study or subgroup	Both generations Mean	SD	Total	Mirena Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.36.5 Second generation								
f. Busfield 2006[107]	86.4	27.5	22	94.6	17.2	28	100.0%	–8.20 (–25.47 to 9.07)
Subtotal (95% CI)			**22**			**28**	**100.0%**	**–8.20 (–21.34 to 4.94)**
Heterogeneity: not applicable								
Test for overall effect: $z = 1.22$ ($p = 0.22$)								



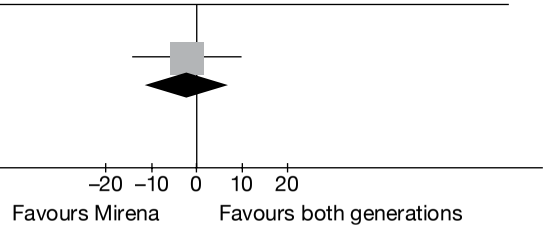
SF-36 emotional role limitation (absolute values) – 2 years

Study or subgroup	Both generations Mean	SD	Total	Mirena Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.37.6 Second generation								
f. Busfield 2006[107]	87.9	26.3	22	86.9	29.2	28	100.0%	1.00 (–19.26 to 21.26)
Subtotal (95% CI)			**22**			**28**	**100.0%**	**1.00 (–14.42 to 16.42)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.13$ ($p = 0.90$)								



SF-36 social function (absolute values) – 2 years

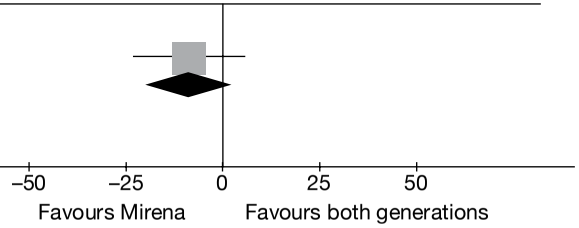
Study or subgroup	Both generations Mean	SD	Total	Mirena Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.38.7 Second generation								
f. Busfield 2006[107]	76.8	18.1	22	79	14	28	100.0%	–2.20 (–14.25 to 9.85)
Subtotal (95% CI)			**22**			**28**	**100.0%**	**–2.20 (–11.37 to 6.97)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.47$ ($p = 0.64$)								



DOI: 10.3310/hta15190

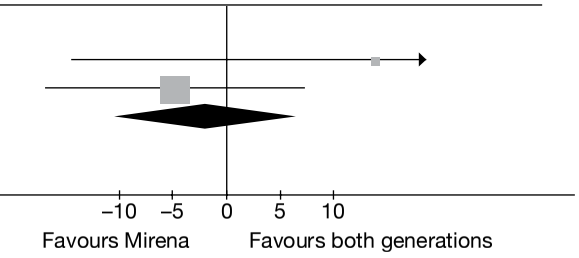
SF-36 bodily pain (absolute values) – 2 years

	Both generations			Mirena				Mean difference	Mean difference IV, Fixed, 99% CI
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 99% CI**	
1.39.8 Second generation									
f. Busfield 2006[107]	74.2	21.5	22	82.9	17.2	28	100.0%	–8.70 (–23.17 to 5.77)	
Subtotal (95% CI)			**22**			**28**	**100.0%**	**–8.70 (–19.71 to 2.31)**	
Heterogeneity: not applicable									
Test for overall effect: $z = 1.55$ ($p = 0.12$)									



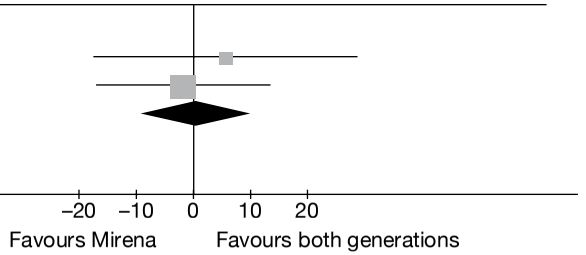
SF-36 general health (change from baseline) – 12 months

	Both generations			Mirena				Mean difference	Mean difference IV, Fixed, 99% CI
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 99% CI**	
1.40.1 Second generation									
e. Tam 2006[109]	15.9	28.3	8	2.1	14.2	10	15.4%	13.80 (–14.45 to 42.05)	
f. Busfield 2006[107]	4	19.3	32	8.8	17.9	31	84.6%	–4.80 (–16.88 to 7.28)	
Subtotal (95% CI)			**40**			**41**	**100.0%**	**–1.93 (–10.38 to 6.52)**	
Heterogeneity: $\chi^2 = 2.43$, df = 1 ($p = 0.12$); $I^2 = 59\%$									
Test for overall effect: $z = 0.45$ ($p = 0.65$)									



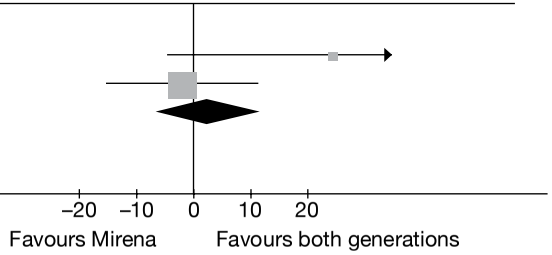
SF-36 physical function (change from baseline) – 12 months

	Both generations			Mirena				
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**Mean difference IV, Fixed, 99% CI**
1.41.2 Second generation								
e. Tam 2006[109]	10	17.9	9	4.4	20.1	9	30.4%	5.60 (–17.51 to 28.71)
f. Busfield 2006[107]	6.9	25.5	32	8.7	21.4	31	69.6%	–1.80 (–17.06 to 13.46)
Subtotal (95% CI)			**41**			**40**	**100.0%**	**0.45 (–9.24 to 10.14)**
Heterogeneity: $\chi^2 = 0.47$, df = 1 ($p = 0.49$); $I^2 = 0\%$								
Test for overall effect: $z = 0.09$ ($p = 0.93$)								



SF-36 mental health (change from baseline) – 12 months

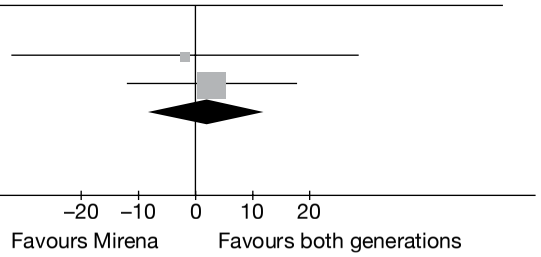
	Both generations			Mirena				
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**Mean difference IV, Fixed, 99% CI**
1.42.3 Second generation								
e. Tam 2006[109]	5	18.5	8	–14.8	20.2	10	17.5%	19.80 (–3.75 to 43.35)
f. Busfield 2006[107]	5.1	16.4	32	6.7	17	31	82.5%	–1.60 (–12.45 to 9.25)
Subtotal (95% CI)			**40**			**41**	**100.0%**	**2.14 (–5.35 to 9.64)**
Heterogeneity: $\chi^2 = 4.52$, df = 1 ($p = 0.03$); $I^2 = 78\%$								
Test for overall effect: $z = 0.56$ ($p = 0.57$)								



DOI: 10.3310/hta15190

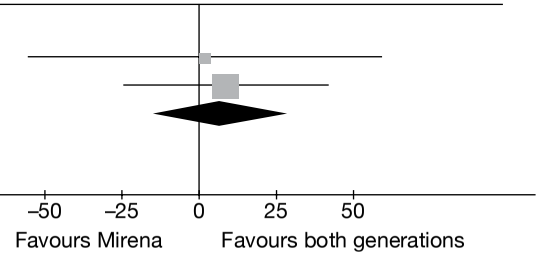
SF-36 vitality (change from baseline) – 12 months

	Both generations			Mirena					
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**Mean difference IV, Fixed, 99% CI**	**Mean difference IV, Fixed, 99% CI**
1.43.4 Second generation									
e. Tam 2006[109]	–1.3	28	8	0.5	20.3	10	19.4%	–1.80 (–32.19 to 28.59)	
f. Busfield 2006[107]	18.4	19.1	32	15.5	25.7	30	80.6%	2.90 (–11.99 to 17.79)	
Subtotal (95% CI)			**40**			**40**	**100.0%**	**1.99 (–8.18 to 12.16)**	
Heterogeneity: $\chi^2 = 0.13$, df = 1 ($p = 0.72$); $I^2 = 0\%$									
Test for overall effect: $z = 0.38$ ($p = 0.70$)									



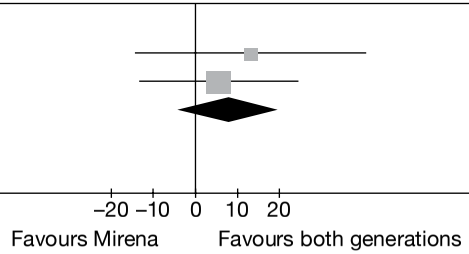
SF-36 physical role limitation (change from baseline) – 12 months

	Both generations			Mirena					
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**Mean difference IV, Fixed, 99% CI**	**Mean difference IV, Fixed, 99% CI**
1.44.5 Second generation									
e. Tam 2006[109]	19.4	49.7	9	17.5	47.2	10	25.2%	1.90 (–55.54 to 59.34)	
f. Busfield 2006[107]	30.5	51.1	32	21.8	51.5	31	74.8%	8.70 (–24.60 to 42.00)	
Subtotal (95% CI)			**41**			**41**	**100.0%**	**6.99 (–14.93 to 28.91)**	
Heterogeneity: $\chi^2 = 0.07$, df = 1 ($p = 0.79$); $I^2 = 0\%$									
Test for overall effect: $z = 0.62$ ($p = 0.53$)									



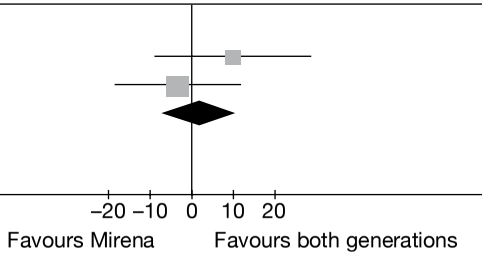
SF-36 bodily pain (change from baseline) – 12 months

	Both generations			Mirena				Mean difference	Mean difference
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 99% CI**	**Mean difference IV, Fixed, 99% CI**
1.45.8 Second generation									
e. Tam 2006[109]	21	26.3	9	7.8	19.6	10	32.1%	13.20 (–14.46 to 40.86)	
f. Busfield 2006[107]	26	30.3	32	20.4	28.3	31	67.9%	5.60 (–13.42 to 24.62)	
Subtotal (95% CI)			**41**			**41**	**100.0%**	**8.04 (–3.88 to 19.97)**	
Heterogeneity: $\chi^2 = 0.34$, df = 1 ($p = 0.56$); $I^2 = 0\%$									
Test for overall effect: $z = 1.32$ ($p = 0.19$)									



SF-36 social function (change from baseline) – 12 months

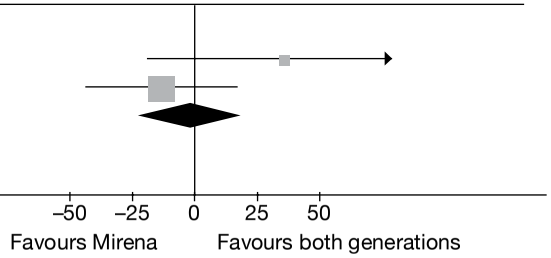
	Both generations			Mirena				Mean difference	Mean difference
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 99% CI**	**Mean difference IV, Fixed, 99% CI**
1.46.7 Second generation									
e. Tam 2006[109]	9.9	16.1	9	0	15.7	10	39.4%	9.90 (–8.93 to 28.73)	
f. Busfield 2006[107]	11.8	25.1	32	15.1	21.6	31	60.6%	–3.30 (–18.48 to 11.88)	
Subtotal (95% CI)			**41**			**41**	**100.0%**	**1.90 (–7.09 to 10.89)**	
Heterogeneity: $\chi^2 = 1.98$, df = 1 ($p = 0.16$); $I^2 = 49\%$									
Test for overall effect: $z = 0.41$ ($p = 0.68$)									



DOI: 10.3310/hta15190

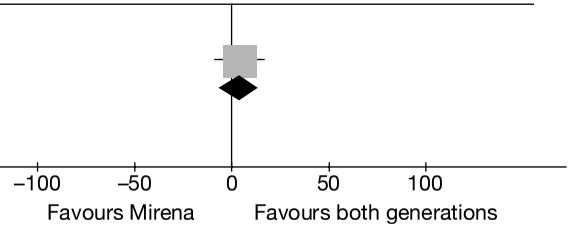
SF-36 emotional role limitation (change from baseline) – 12 months

Study or subgroup	Both generations Mean	SD	Total	Mirena Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.47.6 Second generation								
e. Tam 2006[109]	25.9	54.7	9	–10	35.3	10	23.5%	35.90 (–19.17 to 90.97)
f. Busfield 2006[107]	14.6	51.5	32	28	42.2	31	76.5%	–13.40 (–43.91 to 17.11)
Subtotal (95% CI)			**41**			**41**	**100.0%**	**–1.82 (–22.13 to 18.49)**
Heterogeneity: $\chi^2 = 4.07$, df = 1 ($p = 0.04$); $I^2 = 75\%$								
Test for overall effect: $z = 0.18$ ($p = 0.86$)								



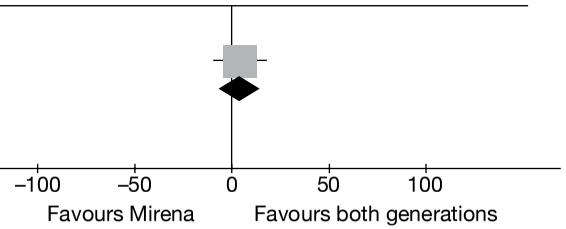
SF-36 general health (change from baseline) – 2 years

Study or subgroup	Both generations Mean	SD	Total	Mirena Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.48.1 Second generation								
f. Busfield 2006[107]	9	15	22	4.9	20.3	27	100.0%	4.10 (–8.90 to 17.10)
Subtotal (95% CI)			**22**			**27**	**100.0%**	**4.10 (–5.80 to 14.00)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.81$ ($p = 0.42$)								



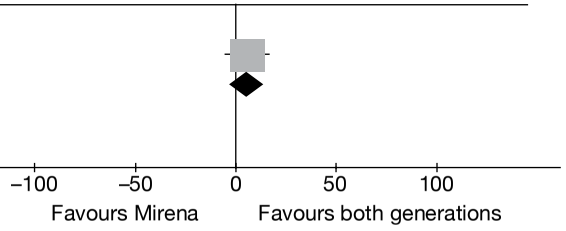
SF-36 physical function (change from baseline) – 2 years

	Both generations			Mirena				Mean difference	Mean difference
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 99% CI**	**Mean difference IV, Fixed, 99% CI**
1.49.2 Second generation									
f. Busfield 2006[107]	13.2	18	22	8.8	19.8	28	100.0%	4.40 (–9.41 to 18.21)	
Subtotal (95% CI)			**22**			**28**	**100.0%**	**4.40 (–6.11 to 14.91)**	
Heterogeneity: not applicable									
Test for overall effect: $z = 0.82$ ($p = 0.41$)									



SF-36 mental health (change from baseline) – 2 years

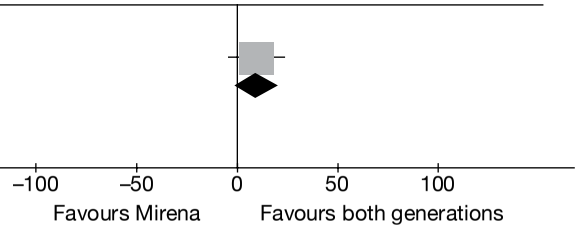
	Both generations			Mirena				Mean difference	Mean difference
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 99% CI**	**Mean difference IV, Fixed, 99% CI**
1.50.3 Second generation									
f. Busfield 2006[107]	10.3	16.2	21	4.6	13.4	28	100.0%	5.70 (–5.50 to 16.90)	
Subtotal (95% CI)			**21**			**28**	**100.0%**	**5.70 (–2.82 to 14.22)**	
Heterogeneity: not applicable									
Test for overall effect: $z = 1.31$ ($p = 0.19$)									



DOI: 10.3310/hta15190

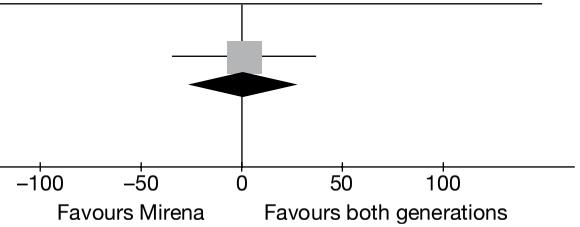
SF-36 vitality (change from baseline) – 2 years

	Both generations			Mirena					
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**Mean difference IV, Fixed, 99% CI**	**Mean difference IV, Fixed, 99% CI**
1.51.4 Second generation									
f. Busfield 2006[107]	21.4	19.3	21	11.8	18.6	28	100.0%	9.60 (–4.53 to 23.73)	
Subtotal (95% CI)			**21**			**28**	**100.0%**	**9.60 (–1.15 to 20.35)**	
Heterogeneity: not applicable									
Test for overall effect: $z = 1.75$ ($p = 0.08$)									



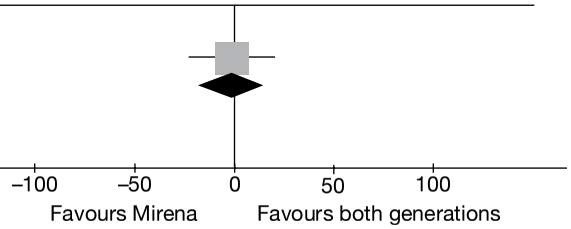
SF-36 physical role limitation (change from baseline) – 2 years

	Both generations			Mirena					
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**Mean difference IV, Fixed, 99% CI**	**Mean difference IV, Fixed, 99% CI**
1.52.5 Second generation									
f. Busfield 2006[107]	26.1	54.3	22	25	40.8	28	100.0%	1.10 (–34.73 to 36.93)	
Subtotal (95% CI)			**22**			**28**	**100.0%**	**1.10 (–26.16 to 28.36)**	
Heterogeneity: not applicable									
Test for overall effect: $z = 0.08$ ($p = 0.94$)									



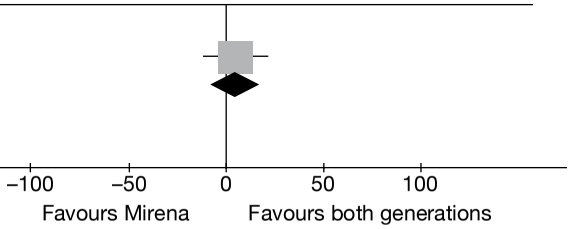
SF-36 bodily pain (change from baseline) – 2 years

	Both generations			Mirena				Mean difference	Mean difference IV, Fixed, 99% CI
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 99% CI**	
1.53.8 Second generation									
f. Busfield 2006[107]	19.7	32.7	22	21	24.8	28	100.0%	–1.30 (–22.94 to 20.34)	
Subtotal (95% CI)			**22**			**28**	**100.0%**	**–1.30 (–17.76 to 15.16)**	
Heterogeneity: not applicable									
Test for overall effect: $z = 0.15$ ($p = 0.88$)									



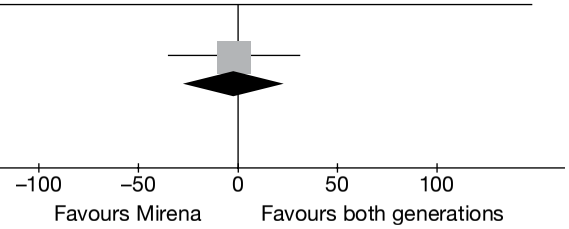
SF-36 social function (change from baseline) – 2 years

	Both generations			Mirena				Mean difference	Mean difference IV, Fixed, 99% CI
Study or subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 99% CI**	
1.54.7 Second generation									
f. Busfield 2006[107]	15.2	26	22	10.3	17.9	28	100.0%	4.90 (–11.83 to 21.63)	
Subtotal (95% CI)			**22**			**28**	**100.0%**	**4.90 (–7.83 to 17.63)**	
Heterogeneity: not applicable									
Test for overall effect: $z = 0.75$ ($p = 0.45$)									



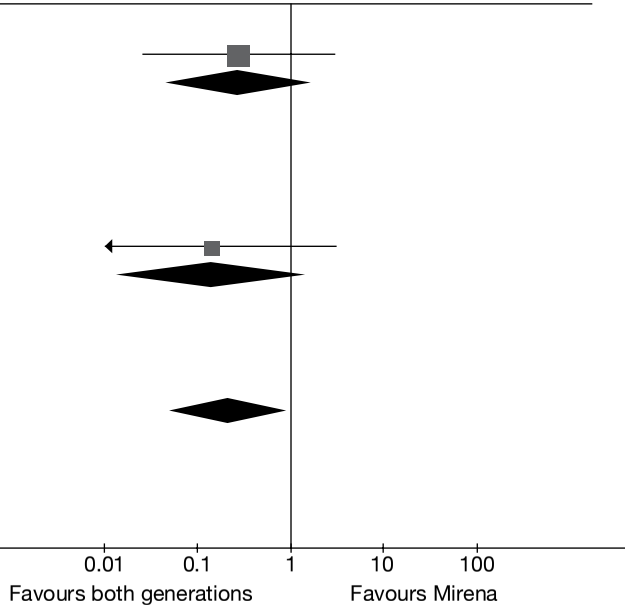
SF-36 emotional role limitation (change from baseline) – 2 years

Study or subgroup	Both generations Mean	SD	Total	Mirena Mean	SD	Total	Weight	Mean difference IV, Fixed, 99% CI
1.55.6 Second generation								
f. Busfield 2006[107]	18.2	42.1	22	20.2	49.1	28	100.0%	–2.00 (–35.25 to 31.25)
Subtotal (95% CI)			**22**			**28**	**100.0%**	**–2.00 (–27.30 to 23.30)**
Heterogeneity: not applicable								
Test for overall effect: $z = 0.15$ ($p = 0.88$)								



Proportion requiring endometrial ablation – 12 months

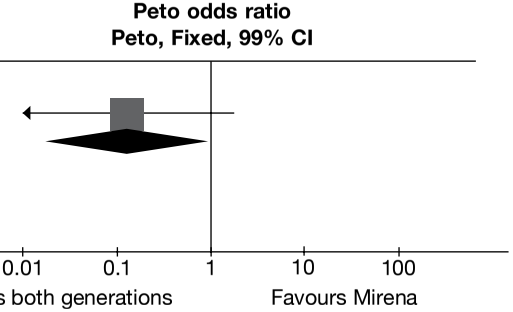
Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.56.1 First generation						
a. Malak 2006[104]	1	30	4	30	62.6%	0.28 (0.03 to 3.00)
Subtotal (95% CI)		**30**		**30**	**62.6%**	**0.28 (0.04 to 1.70)**
Total events	1		4			
Heterogeneity: not applicable						
Test for overall effect: $z = 1.39$ ($p = 0.16$)						
1.56.2 Second generation						
e. Tam 2006[109]	0	15	3	18	37.4%	0.14 (0.01 to 3.09)
Subtotal (95% CI)		**15**		**18**	**37.4%**	**0.14 (0.01 to 1.48)**
Total events	0		3			
Heterogeneity: not applicable						
Test for overall effect: $z = 1.63$ ($p = 0.10$)						
Total (95% CI)		**45**		**48**	**100.0%**	**0.21 (0.05 to 0.90)**
Total events	1		7			
Heterogeneity: $\chi^2 = 0.19$, df = 1 ($p = 0.66$); $I^2 = 0\%$						
Test for overall effect: $z = 2.10$ ($p = 0.04$)						
Test for subgroup differences: $\chi^2 = 0.19$, df = 1 ($p = 0.66$), $I^2 = 0\%$						



DOI: 10.3310/hta15190

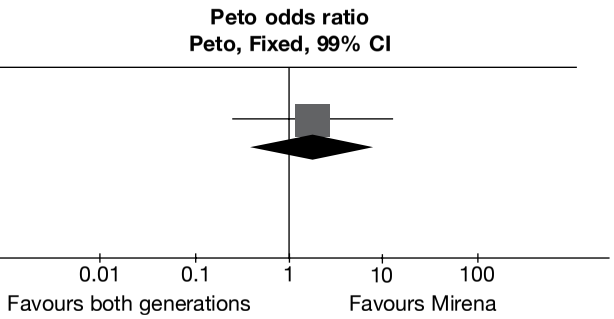
Proportion requiring endometrial ablation – 2 years

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.57.2 Second generation						
f. Busfield 2006[107]	0	41	4	42	100.0%	0.13 (0.01 to 1.77)
Subtotal (95% CI)		**41**		**42**	**100.0%**	**0.13 (0.02 to 0.95)**
Total events	0		4			
Heterogeneity: not applicable						
Test for overall effect: $z = 2.01$ ($p = 0.04$)						



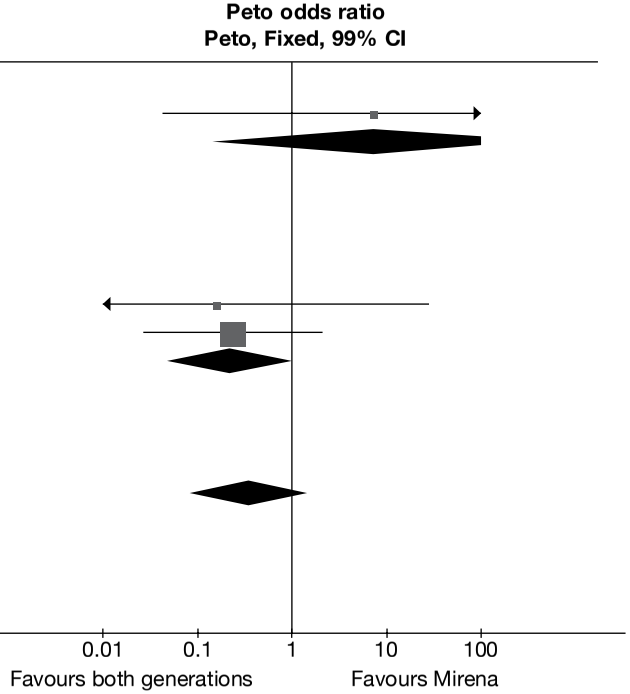
Proportion requiring hysterectomy – <12 months

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.58.1 Second generation						
g. Barrington 2003[81]	5	25	3	25	100.0%	1.79 (0.25 to 12.82)
Subtotal (95% CI)		**25**		**25**	**100.0%**	**1.79 (0.40 to 8.01)**
Total events	5		3			
Heterogeneity: not applicable						
Test for overall effect: $z = 0.76$ ($p = 0.45$)						



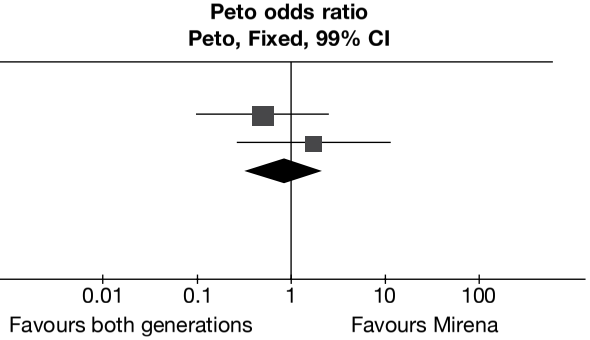
Proportion requiring hysterectomy – 12 months

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.59.1 First generation						
c. Crosignani 1997[45]	1	35	0	35	13.2%	7.39 (0.04 to 1276.22)
Subtotal (95% CI)		**35**		**35**	**13.2%**	**7.39 (0.15 to 372.38)**
Total events	1		0			
Heterogeneity: not applicable						
Test for overall effect: $z = 1.00$ ($p = 0.32$)						
1.59.2 Second generation						
e. Tam 2006[109]	0	15	1	18	13.1%	0.16 (0.00 to 28.21)
h. Soysal 2002[106]	1	36	5	36	73.7%	0.24 (0.03 to 2.11)
Subtotal (95% CI)		**51**		**54**	**86.8%**	**0.22 (0.05 to 1.04)**
Total events	1		6			
Heterogeneity: $\chi^2 = 0.03$, df = 1 ($p = 0.85$); $I^2 = 0\%$						
Test for overall effect: $z = 1.92$ ($p = 0.06$)						
Total (95% CI)		**86**		**89**	**100.0%**	**0.36 (0.09 to 1.48)**
Total events	2		6			
Heterogeneity: $\chi^2 = 2.68$, df = 2 ($p = 0.26$); $I^2 = 25\%$						
Test for overall effect: $z = 1.42$ ($p = 0.16$)						
Test for subgroup differences: $\chi^2 = 2.65$, df = 1 ($p = 0.10$), $I^2 = 62.3\%$						



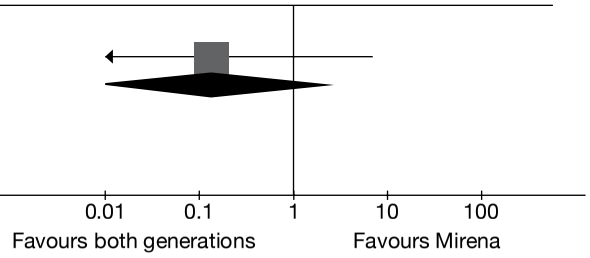
Proportion requiring hysterectomy – 2 years

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.60.3 Second generation						
d. Shaw 2007[108]	4	30	8	33	57.4%	0.50 (0.10 to 2.57)
f. Busfield 2006[107]	5	41	3	42	42.6%	1.77 (0.26 to 11.91)
Subtotal (95% CI)		**71**		**75**	**100.0%**	**0.86 (0.33 to 2.21)**
Total events	9		11			
Heterogeneity: $\chi^2 = 1.69$, df = 1 ($p = 0.19$); $I^2 = 41\%$						
Test for overall effect: $z = 0.32$ ($p = 0.75$)						



Proportion discontinuing Mirena – <12 months

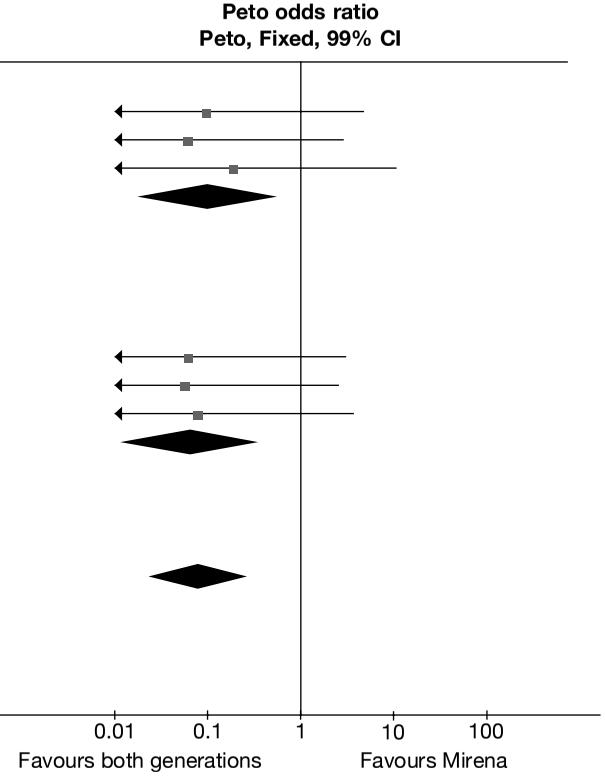
Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Odds ratio Peto, Fixed, 99% CI
1.61.2 Second generation						
f. Busfield 2006[107]	0	41	3	42	100.0%	0.14 (0.00 to 6.96)
Subtotal (95% CI)		**41**		**42**	**100.0%**	**0.14 (0.01 to 2.72)**
Total events	0		3			
Heterogeneity: not applicable						
Test for overall effect: $z = 1.31$ ($p = 0.19$)						



DOI: 10.3310/hta15190

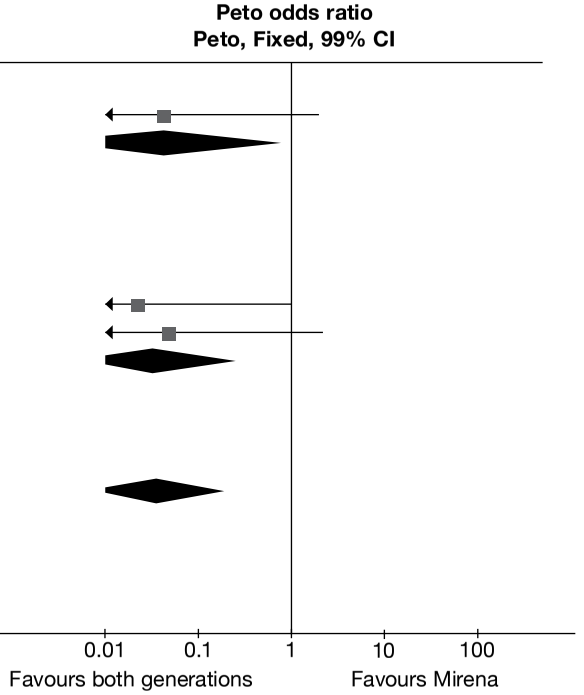
Proportion discontinuing Mirena –12 months

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.62.1 First generation						
a. Malak 2006[104]	0	30	4	30	16.6%	0.10 (0.00 to 4.77)
b. Kittelsen 1998[105]	0	30	6	30	17.1%	0.06 (0.00 to 2.89)
c. Crosignani 1997[45]	0	35	2	35	15.5%	0.19 (0.00 to 10.71)
Subtotal (95% CI)		**95**		**95**	**49.1%**	**0.10 (0.02 to 0.57)**
Total events	0		12			
Heterogeneity: $\chi^2 = 0.27$, df = 2 ($p = 0.87$); $I^2 = 0\%$						
Test for overall effect: $z = 2.59$ ($p = 0.009$)						
1.62.3 Second generation						
e. Tam 2006[109]	0	15	6	18	16.5%	0.06 (0.00 to 3.08)
f. Busfield 2006[107]	0	41	7	42	17.4%	0.06 (0.00 to 2.57)
h. Soysal 2002[106]	0	36	5	36	17.0%	0.08 (0.00 to 3.71)
Subtotal (95% CI)		**92**		**96**	**50.9%**	**0.07 (0.01 to 0.35)**
Total events	0		18			
Heterogeneity: $\chi^2 = 0.02$, df = 2 ($p = 0.99$); $I^2 = 0\%$						
Test for overall effect: $z = 3.16$ ($p = 0.002$)						
Total (95% CI)		**187**		**191**	**100.0%**	**0.08 (0.02 to 0.27)**
Total events	0		30			
Heterogeneity: $\chi^2 = 0.43$, df = 5 ($p = 0.99$); $I^2 = 0\%$						
Test for overall effect: $z = 4.07$ ($p < 0.0001$)						
Test for subgroup differences: $\chi^2 = 0.13$, df = 1 ($p = 0.72$), $I^2 = 0\%$						



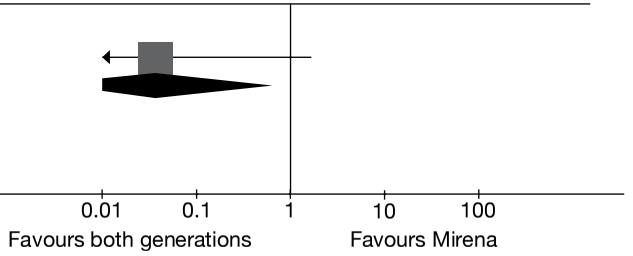
Proportion discontinuing Mirena – 2 years

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.63.2 First generation						
b. Kittelsen 1998[105]	0	30	8	30	33.0%	0.04 (0.00 to 1.97)
Subtotal (95% CI)		**30**		**30**	**33.0%**	**0.04 (0.00 to 0.79)**
Total events	0		8			
Heterogeneity: not applicable						
Test for overall effect: $z = 2.12$ ($p = 0.03$)						
1.63.4 Second generation						
d. Shaw 2007[108]	0	33	13	33	33.6%	0.02 (0.00 to 0.99)
f. Busfield 2006[107]	0	41	8	42	33.4%	0.05 (0.00 to 2.18)
Subtotal (95% CI)		**74**		**75**	**67.0%**	**0.03 (0.00 to 0.26)**
Total events	0		21			
Heterogeneity: $\chi^2 = 0.14$, df = 1 ($p = 0.71$); $I^2 = 0\%$						
Test for overall effect: $z = 3.27$ ($p = 0.001$)						
Total (95% CI)		**104**		**105**	**100.0%**	**0.04 (0.01 to 0.19)**
Total events	0		29			
Heterogeneity: $\chi^2 = 0.16$, df = 2 ($p = 0.92$); $I^2 = 0\%$						
Test for overall effect: $z = 3.90$ ($p < 0.0001$)						
Test for subgroup differences: $\chi^2 = 0.02$, df = 1 ($p = 0.88$), $I^2 = 0\%$						



Proportion discontinuing Mirena – 3 years

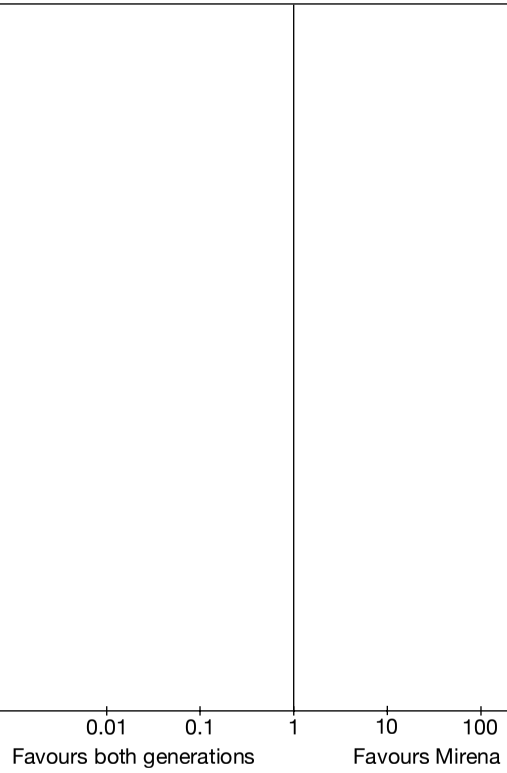
Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Odds ratio Peto, Fixed, 99% CI
1.64.3 First generation						
b. Kittelsen 1998[105]	0	30	9	30	100.0%	0.04 (0.00 to 1.67)
Subtotal (95% CI)		**30**		**30**	**100.0%**	**0.04 (0.00 to 0.67)**
Total events	0		9			
Heterogeneity: not applicable						
Test for overall effect: $z = 2.23$ ($p = 0.03$)						



DOI: 10.3310/hta15190

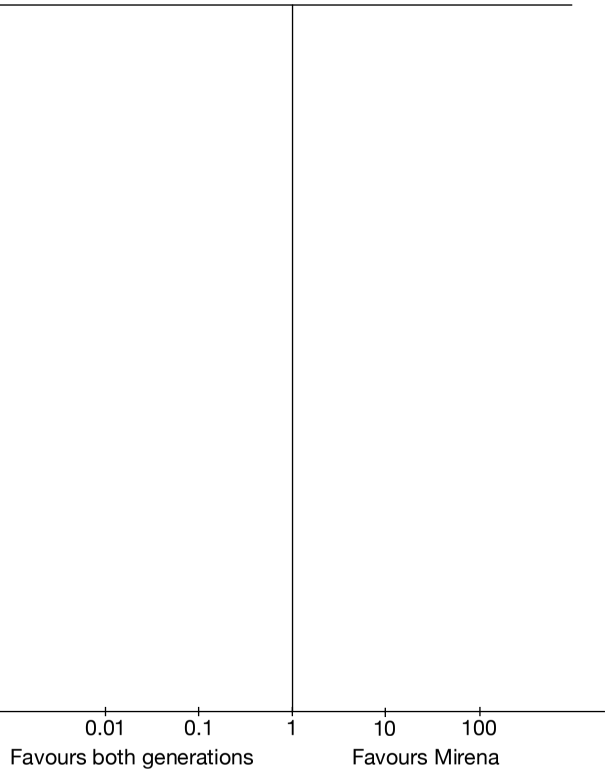
Patients with adverse events – periprocedure (uterine perforation)

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.65.1 First generation						
a. Malak 2006[104]	0	30	0	30		Not estimable
b. Kittelsen 1998[105]	0	30	0	30		Not estimable
c. Crosignani 1997[45]	0	35	0	35		Not estimable
Subtotal (95% CI)		**95**		**95**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
1.65.2 Second generation						
d. Shaw 2007[108]	0	33	0	33		Not estimable
e. Tam 2006[109]	0	0	0	0		Not estimable
f. Busfield 2006[107]	0	0	0	0		Not estimable
g. Barrington 2003[81]	0	0	0	0		Not estimable
h. Soysal 2002[106]	0	0	0	0		Not estimable
Subtotal (95% CI)		**33**		**33**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Total (95% CI)		**128**		**128**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Test for subgroup differences: not applicable						



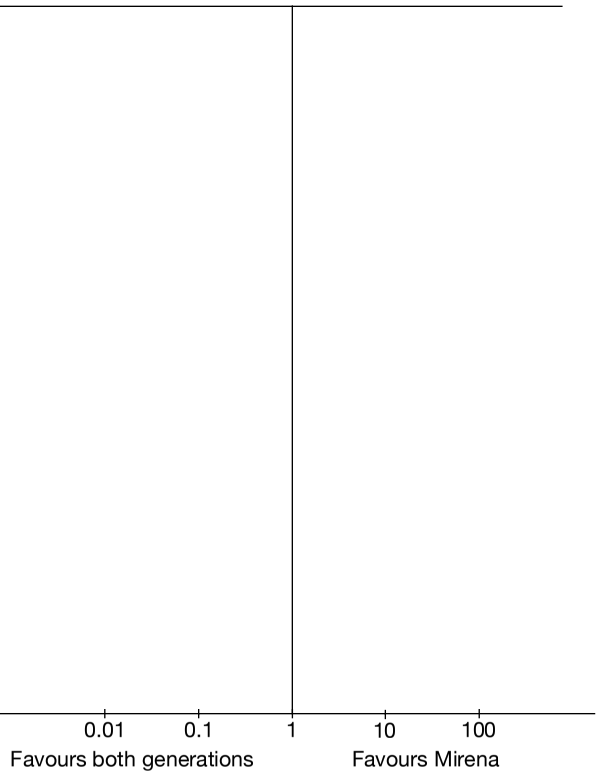
Patients with adverse events – periprocedure (cervical laceration)

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.66.1 First generation						
a. Malak 2006[104]	0	30	0	30		Not estimable
b. Kittelsen 1998[105]	0	30	0	30		Not estimable
c. Crosignani 1997[45]	0	35	0	35		Not estimable
Subtotal (95% CI)		**95**		**95**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
1.66.2 Second generation						
d. Shaw 2007[108]	0	33	0	33		Not estimable
e. Tam 2006[109]	0	22	0	22		Not estimable
f. Busfield 2006[107]	0	41	0	42		Not estimable
g. Barrington 2003[81]	0	25	0	25		Not estimable
h. Soysal 2002[106]	0	36	0	36		Not estimable
Subtotal (95% CI)		**157**		**158**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Total (95% CI)		**252**		**253**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Test for subgroup differences: not applicable						



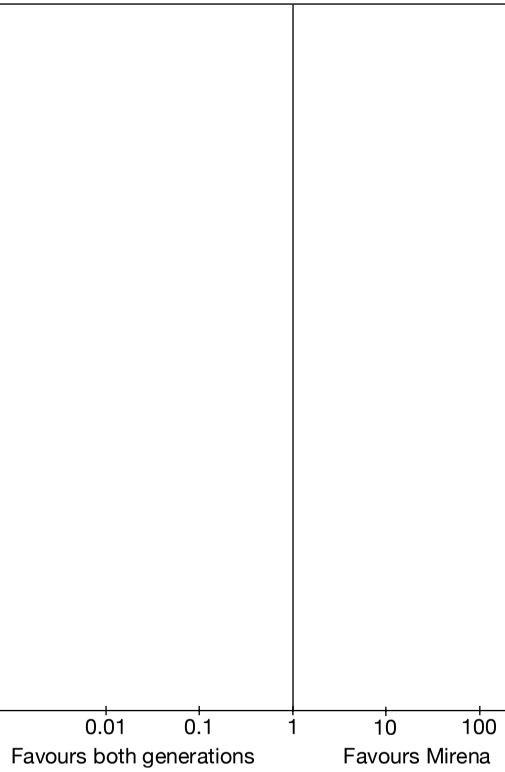
Patients with adverse events – periprocedure (anaesthesia problems)

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.67.1 First generation						
a. Malak 2006[104]	0	30	0	30		Not estimable
b. Kittelsen 1998[105]	0	30	0	30		Not estimable
c. Crosignani 1997[45]	0	35	0	35		Not estimable
Subtotal (95% CI)		**95**		**95**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
1.67.2 Second generation						
d. Shaw 2007[108]	0	33	0	33		Not estimable
e. Tam 2006[109]	0	22	0	22		Not estimable
f. Busfield 2006[107]	0	41	0	42		Not estimable
g. Barrington 2003[81]	0	25	0	25		Not estimable
h. Soysal 2002[106]	0	36	0	36		Not estimable
Subtotal (95% CI)		**157**		**158**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Total (95% CI)		**252**		**253**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Test for subgroup differences: not applicable						



Patients with adverse events – periprocedure (excessive bleeding)

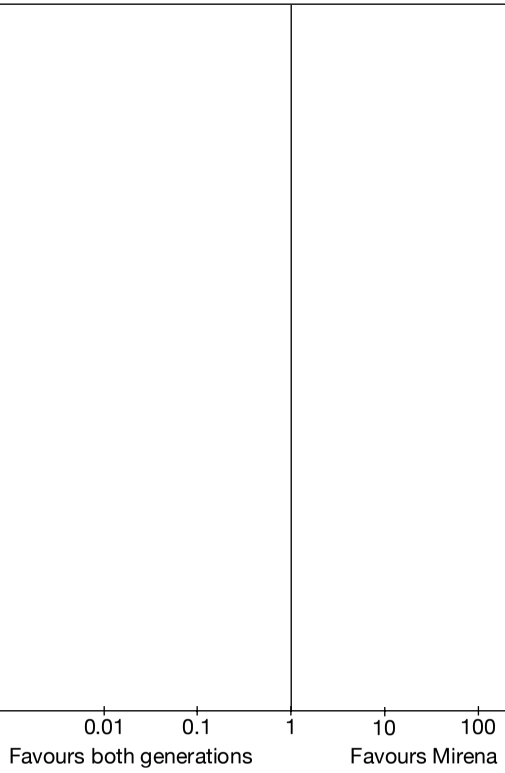
Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.68.1 First generation						
a. Malak 2006[104]	0	30	0	30		Not estimable
b. Kittelsen 1998[105]	0	30	0	30		Not estimable
c. Crosignani 1997[45]	0	35	0	35		Not estimable
Subtotal (95% CI)		**95**		**95**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
1.68.2 Second generation						
d. Shaw 2007[108]	0	33	0	33		Not estimable
e. Tam 2006[109]	0	22	0	22		Not estimable
f. Busfield 2006[107]	0	41	0	42		Not estimable
g. Barrington 2003[81]	0	25	0	25		Not estimable
h. Soysal 2002[106]	0	36	0	36		Not estimable
Subtotal (95% CI)		**157**		**158**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Total (95% CI)		**252**		**253**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Test for subgroup differences: not applicable						



DOI: 10.3310/hta15190

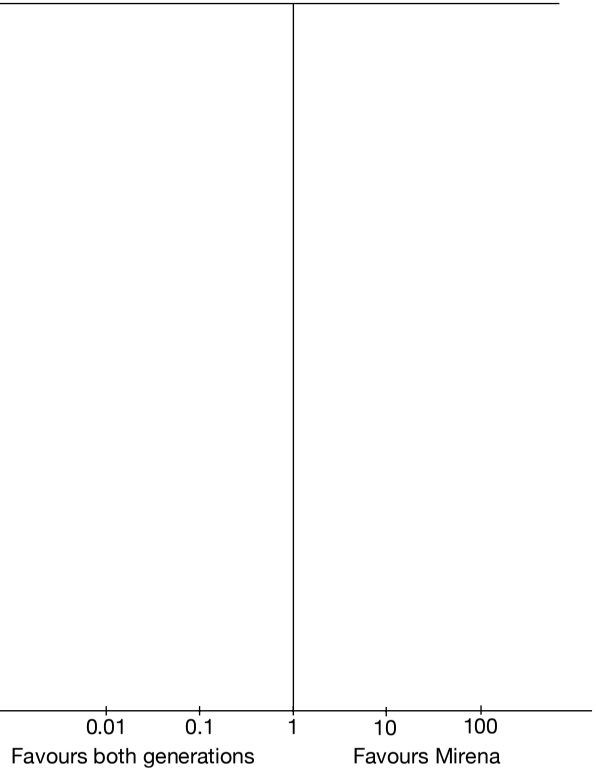
Patients with adverse events – periprocedure (fluid overload)

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.69.1 First generation						
a. Malak 2006[104]	0	30	0	30		Not estimable
b. Kittelsen 1998[105]	0	30	0	30		Not estimable
c. Crosignani 1997[45]	0	35	0	35		Not estimable
Subtotal (95% CI)		**95**		**95**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
1.69.2 Second generation						
d. Shaw 2007[108]	0	33	0	33		Not estimable
e. Tam 2006[109]	0	22	0	22		Not estimable
f. Busfield 2006[107]	0	41	0	42		Not estimable
g. Barrington 2003[81]	0	25	0	25		Not estimable
h. Soysal 2002[106]	0	36	0	36		Not estimable
Subtotal (95% CI)		**157**		**158**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Total (95% CI)		**252**		**253**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Test for subgroup differences: not applicable						



Patients with adverse events – periprocedure (excessive visceral damage)

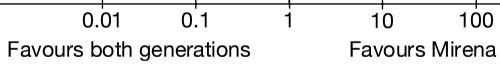
Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.70.1 First generation						
a. Malak 2006[104]	0	30	0	30		Not estimable
b. Kittelsen 1998[105]	0	30	0	30		Not estimable
c. Crosignani 1997[45]	0	35	0	35		Not estimable
Subtotal (95% CI)		**95**		**95**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
1.70.2 Second generation						
d. Shaw 2007[108]	0	33	0	33		Not estimable
e. Tam 2006[109]	0	22	0	22		Not estimable
f. Busfield 2006[107]	0	41	0	42		Not estimable
g. Barrington 2003[81]	0	25	0	25		Not estimable
h. Soysal 2002[106]	0	36	0	36		Not estimable
Subtotal (95% CI)		**157**		**158**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Total (95% CI)		**252**		**253**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Test for subgroup differences: not applicable						



DOI: 10.3310/hta15190

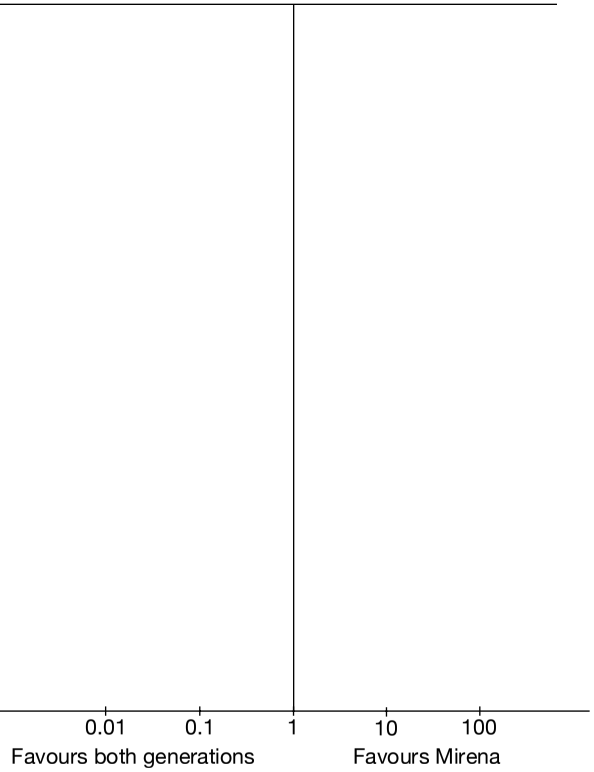
Patients with adverse events – periprocedure (procedure abandoned)

	Both generations		Mirena				
Study or subgroup	**Events**	**Total**	**Events**	**Total**	**Weight**	**Peto odds ratio Peto, Fixed, 99% CI**	**Peto odds ratio Peto, Fixed, 99% CI**
1.71.1 First generation							
a. Malak 2006[104]	0	30	0	30		Not estimable	
b. Kittelsen 1998[105]	0	30	0	30		Not estimable	
c. Crosignani 1997[45]	0	35	0	35		Not estimable	
Subtotal (95% CI)		**95**		**95**		**Not estimable**	
Total events	0		0				
Heterogeneity: not applicable							
Test for overall effect: not applicable							
1.71.2 Second generation							
d. Shaw 2007[108]	0	33	0	33		Not estimable	
e. Tam 2006[109]	0	22	0	22		Not estimable	
f. Busfield 2006[107]	0	41	0	42		Not estimable	
g. Barrington 2003[81]	0	25	0	25		Not estimable	
h. Soysal 2002[106]	0	36	0	36		Not estimable	
Subtotal (95% CI)		**157**		**158**		**Not estimable**	
Total events	0		0				
Heterogeneity: not applicable							
Test for overall effect: not applicable							
Total (95% CI)		**252**		**253**		**Not estimable**	
Total events	0		0				
Heterogeneity: not applicable							
Test for overall effect: not applicable							
Test for subgroup differences: not applicable							



Patients with adverse events – periprocedure (procedure converted to hysterectomy)

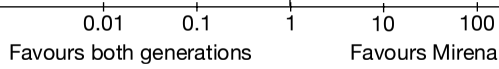
Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.72.1 First generation						
a. Malak 2006[104]	0	30	0	30		Not estimable
b. Kittelsen 1998[105]	0	30	0	30		Not estimable
c. Crosignani 1997[45]	0	35	0	35		Not estimable
Subtotal (95% CI)		**95**		**95**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
1.72.2 Second generation						
d. Shaw 2007[108]	0	33	0	33		Not estimable
e. Tam 2006[109]	0	22	0	22		Not estimable
f. Busfield 2006[107]	0	41	0	42		Not estimable
g. Barrington 2003[81]	0	25	0	25		Not estimable
h. Soysal 2002[106]	0	36	0	36		Not estimable
Subtotal (95% CI)		**157**		**158**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Total (95% CI)		**252**		**253**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Test for subgroup differences: not applicable						



DOI: 10.3310/hta15190

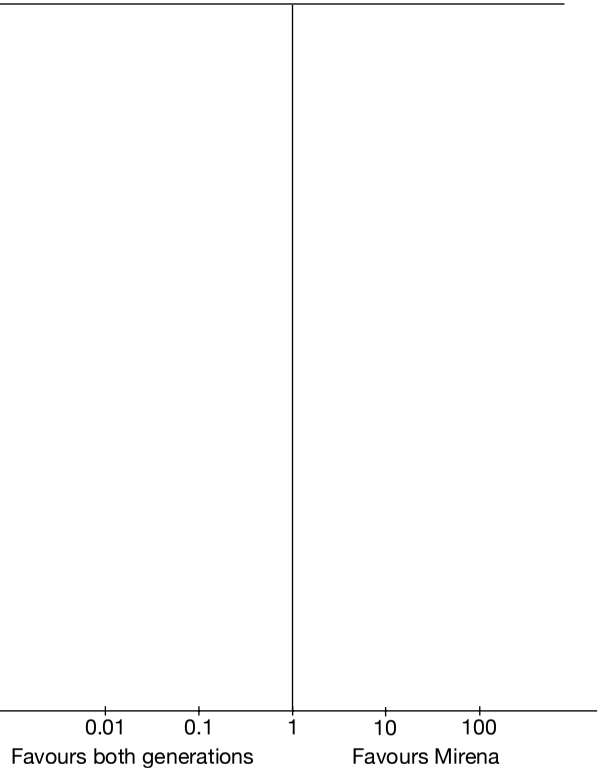
Patients with adverse events – periprocedure (failed insert)

	Both generations		Mirena				
Study or subgroup	**Events**	**Total**	**Events**	**Total**	**Weight**	**Peto odds ratio Peto, Fixed, 99% CI**	**Peto odds ratio Peto, Fixed, 99% CI**
1.73.1 First generation							
a. Malak 2006[104]	0	30	0	30		Not estimable	
b. Kittelsen 1998[105]	0	30	0	30		Not estimable	
c. Crosignani 1997[45]	0	35	0	35		Not estimable	
Subtotal (95% CI)		**95**		**95**		**Not estimable**	
Total events	0		0				
Heterogeneity: not applicable							
Test for overall effect: not applicable							
1.73.2 Second generation							
d. Shaw 2007[108]	0	33	0	33		Not estimable	
e. Tam 2006[109]	0	22	0	22		Not estimable	
f. Busfield 2006[107]	0	41	0	42		Not estimable	
g. Barrington 2003[81]	0	25	0	25		Not estimable	
h. Soysal 2002[106]	0	36	0	36		Not estimable	
Subtotal (95% CI)		**157**		**158**		**Not estimable**	
Total events	0		0				
Heterogeneity: not applicable							
Test for overall effect: not applicable							
Total (95% CI)		**252**		**253**		**Not estimable**	
Total events	0		0				
Heterogeneity: not applicable							
Test for overall effect: not applicable							
Test for subgroup differences: not applicable							



Patients with adverse events – postoperatively (urinary tract infection)

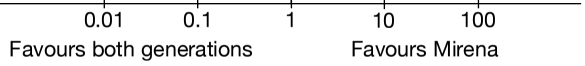
Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.74.1 First generation						
a. Malak 2006[104]	0	30	0	30		Not estimable
b. Kittelsen 1998[105]	0	30	0	30		Not estimable
c. Crosignani 1997[45]	0	35	0	35		Not estimable
Subtotal (95% CI)		**95**		**95**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
1.74.2 Second generation						
d. Shaw 2007[108]	0	33	0	33		Not estimable
e. Tam 2006[109]	0	22	0	22		Not estimable
f. Busfield 2006[107]	0	41	0	42		Not estimable
g. Barrington 2003[81]	0	25	0	25		Not estimable
h. Soysal 2002[106]	0	36	0	36		Not estimable
Subtotal (95% CI)		**157**		**158**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Total (95% CI)		**252**		**253**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Test for subgroup differences: not applicable						



DOI: 10.3310/hta15190

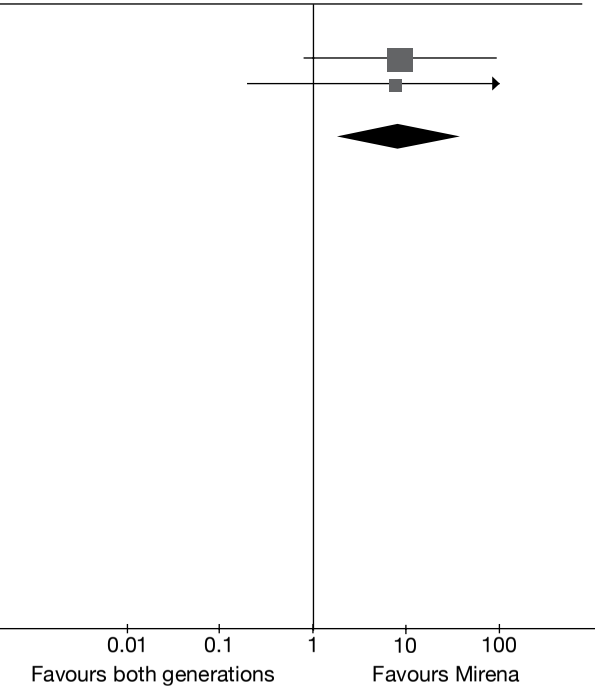
Patients with adverse events – postoperatively (deep-vein thrombosis)

	Both generations		Mirena			Peto odds ratio	Peto odds ratio
Study or subgroup	**Events**	**Total**	**Events**	**Total**	**Weight**	**Peto, Fixed, 99% CI**	**Peto, Fixed, 99% CI**
1.75.1 First generation							
a. Malak 2006[104]	0	30	0	30		Not estimable	
b. Kittelsen 1998[105]	0	30	0	30		Not estimable	
c. Crosignani 1997[45]	0	35	0	35		Not estimable	
Subtotal (95% CI)		**95**		**95**		**Not estimable**	
Total events	0		0				
Heterogeneity: not applicable							
Test for overall effect: not applicable							
1.75.2 Second generation							
d. Shaw 2007[108]	0	33	0	33		Not estimable	
e. Tam 2006[109]	0	22	0	22		Not estimable	
f. Busfield 2006[107]	0	41	0	42		Not estimable	
g. Barrington 2003[81]	0	25	0	25		Not estimable	
h. Soysal 2002[106]	0	36	0	36		Not estimable	
Subtotal (95% CI)		**157**		**158**		**Not estimable**	
Total events	0		0				
Heterogeneity: not applicable							
Test for overall effect: not applicable							
Total (95% CI)		**252**		**253**		**Not estimable**	
Total events	0		0				
Heterogeneity: not applicable							
Test for overall effect: not applicable							
Test for subgroup differences: not applicable							



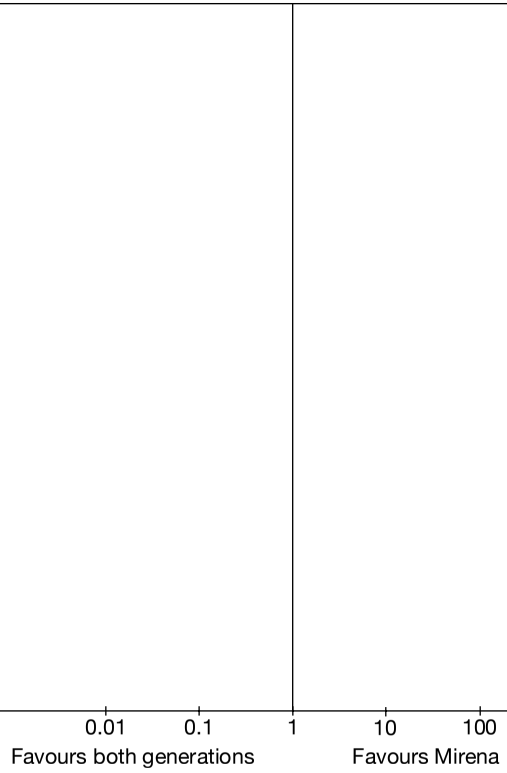
Patients with adverse events – postoperatively (further bleeding)

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.76.1 First generation						
a. Malak 2006[104]	5	30	0	30	70.3%	8.55 (0.79 to 92.91)
b. Kittelsen 1998[105]	2	30	0	30	29.7%	7.65 (0.19 to 301.43)
c. Crosignani 1997[45]	0	35	0	35		Not estimable
Subtotal (95% CI)		**95**		**95**	**100.0%**	**8.27 (1.80 to 37.91)**
Total events	7		0			
Heterogeneity: $\chi^2 = 0.00$, df = 1 ($p = 0.95$); $I^2 = 0\%$						
Test for overall effect: $z = 2.72$ ($p = 0.007$)						
1.76.2 Second generation						
d. Shaw 2007[108]	0	33	0	33		Not estimable
e. Tam 2006[109]	0	22	0	22		Not estimable
f. Busfield 2006[107]	0	41	0	42		Not estimable
g. Barrington 2003[81]	0	25	0	25		Not estimable
h. Soysal 2002[106]	0	36	0	36		Not estimable
Subtotal (95% CI)		**157**		**158**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Test for subgroup differences: not applicable						



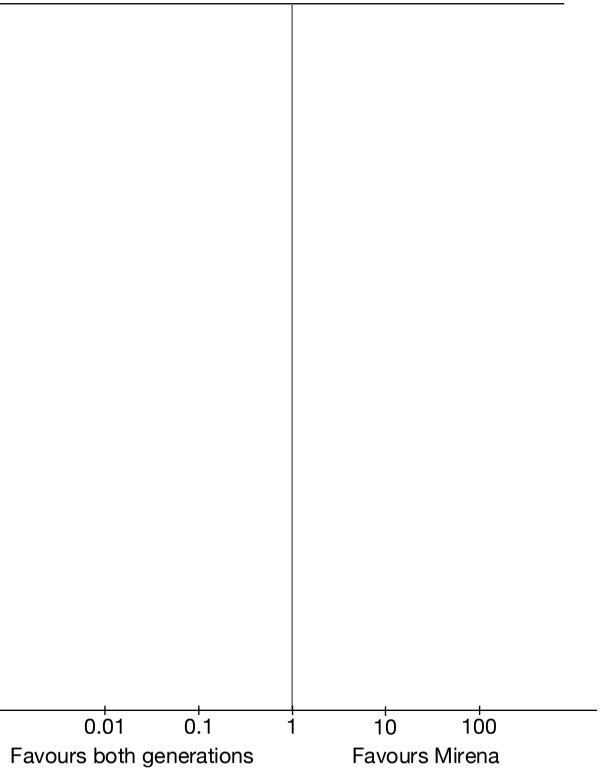
Patients with adverse events – postoperatively (sepsis)

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.77.1 First generation						
a. Malak 2006[104]	0	30	0	30		Not estimable
b. Kittelsen 1998[105]	0	30	0	30		Not estimable
c. Crosignani 1997[45]	0	35	0	35		Not estimable
Subtotal (95% CI)		**95**		**95**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
1.77.2 Second generation						
d. Shaw 2007[108]	0	33	0	33		Not estimable
e. Tam 2006[109]	0	22	0	22		Not estimable
f. Busfield 2006[107]	0	41	0	42		Not estimable
g. Barrington 2003[81]	0	25	0	25		Not estimable
h. Soysal 2002[106]	0	36	0	36		Not estimable
Subtotal (95% CI)		**157**		**158**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Total (95% CI)		**252**		**253**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Test for subgroup differences: not applicable						



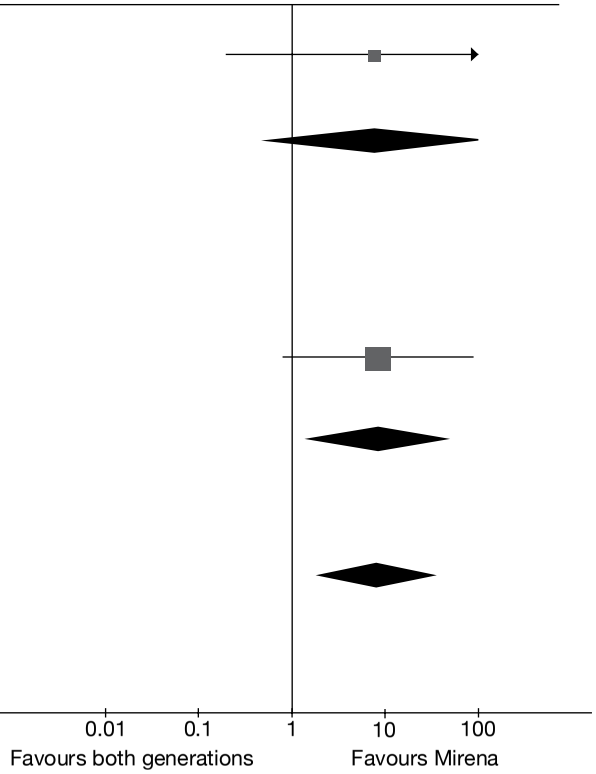
Patients with adverse events – postoperatively (pyrexia)

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.78.1 First generation						
a. Malak 2006[104]	0	30	0	30		Not estimable
b. Kittelsen 1998[105]	0	30	0	30		Not estimable
c. Crosignani 1997[45]	0	35	0	35		Not estimable
Subtotal (95% CI)		**95**		**95**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
1.78.2 Second generation						
d. Shaw 2007[108]	0	33	0	33		Not estimable
e. Tam 2006[109]	0	22	0	22		Not estimable
f. Busfield 2006[107]	0	41	0	42		Not estimable
g. Barrington 2003[81]	0	25	0	25		Not estimable
h. Soysal 2002[106]	0	36	0	36		Not estimable
Subtotal (95% CI)		**157**		**158**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Total (95% CI)		**252**		**253**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Test for subgroup differences: not applicable						



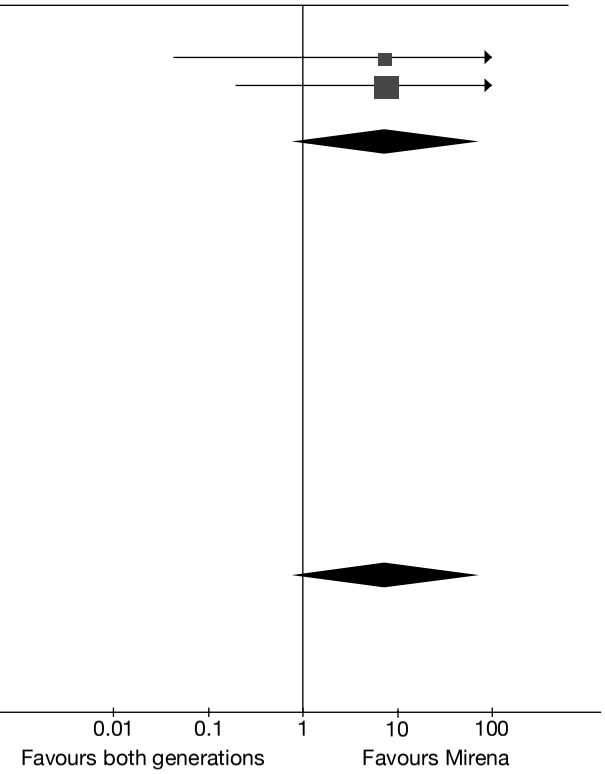
Patients with adverse events – postoperatively (endometriosis)

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.79.1 First generation						
a. Malak 2006[104]	2	30	0	30	29.3%	7.65 (0.19 to 301.43)
b. Kittelsen 1998[105]	0	30	0	30		Not estimable
c. Crosignani 1997[45]	0	35	0	35		Not estimable
Subtotal (95% CI)		**95**		**95**	**29.3%**	**7.65 (0.47 to 125.22)**
Total events	2		0			
Heterogeneity: not applicable						
Test for overall effect: $z = 1.43$ ($p = 0.15$)						
1.79.2 Second generation						
d. Shaw 2007[108]	0	33	0	33		Not estimable
e. Tam 2006[109]	0	22	0	22		Not estimable
f. Busfield 2006[107]	5	41	0	42	70.7%	8.40 (0.79 to 89.18)
g. Barrington 2003[81]	0	25	0	25		Not estimable
h. Soysal 2002[106]	0	36	0	36		Not estimable
Subtotal (95% CI)		**157**		**158**	**70.7%**	**8.40 (1.39 to 50.69)**
Total events	5		0			
Heterogeneity: not applicable						
Test for overall effect: $z = 2.32$ ($p = 0.02$)						
Total (95% CI)		**252**		**253**	**100.0%**	**8.17 (1.80 to 37.07)**
Total events	7		0			
Heterogeneity: $\chi^2 = 0.00$, df = 1 ($p = 0.96$); $I^2 = 0\%$						
Test for overall effect: $z = 2.72$ ($p = 0.006$)						
Test for subgroup differences: $\chi^2 = 0.00$, df = 1 ($p = 0.96$); $I^2 = 0\%$						



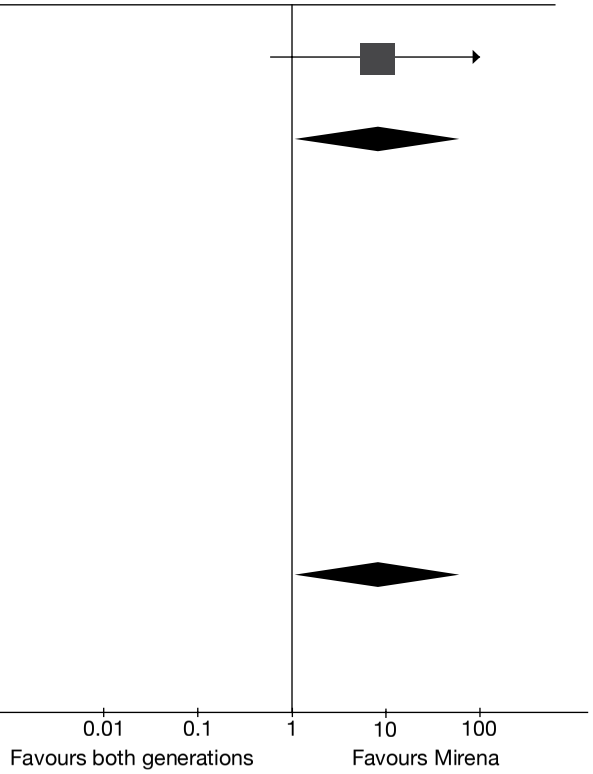
Patients with adverse events – postoperatively (haematometra)

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.80.1 First generation						
a. Malak 2006[104]	1	30	0	30	33.7%	7.39 (0.04 to 1276.22)
b. Kittelsen 1998[105]	2	30	0	30	66.3%	7.65 (0.19 to 301.43)
c. Crosignani 1997[45]	0	35	0	35		Not estimable
Subtotal (95% CI)		**95**		**95**	**100.0%**	**7.56 (0.78 to 73.62)**
Total events	3		0			
Heterogeneity: $\chi^2 = 0.00$, df = 1 ($p = 0.99$); $I^2 = 0\%$						
Test for overall effect: $z = 1.74$ ($p = 0.08$)						
1.80.2 Second generation						
d. Shaw 2007[108]	0	33	0	33		Not estimable
e. Tam 2006[109]	0	22	0	22		Not estimable
f. Busfield 2006[107]	0	41	0	42		Not estimable
g. Barrington 2003[81]	0	25	0	25		Not estimable
h. Soysal 2002[106]	0	36	0	36		Not estimable
Subtotal (95% CI)		**157**		**158**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Total (95% CI)		**252**		**253**	**100.0%**	**7.56 (0.78 to 73.62)**
Total events	3		0			
Heterogeneity: $\chi^2 = 0.00$, df = 1 ($p = 0.99$); $I^2 = 0\%$						
Test for overall effect: $z = 1.74$ ($p = 0.08$)						
Test for subgroup differences: not applicable						



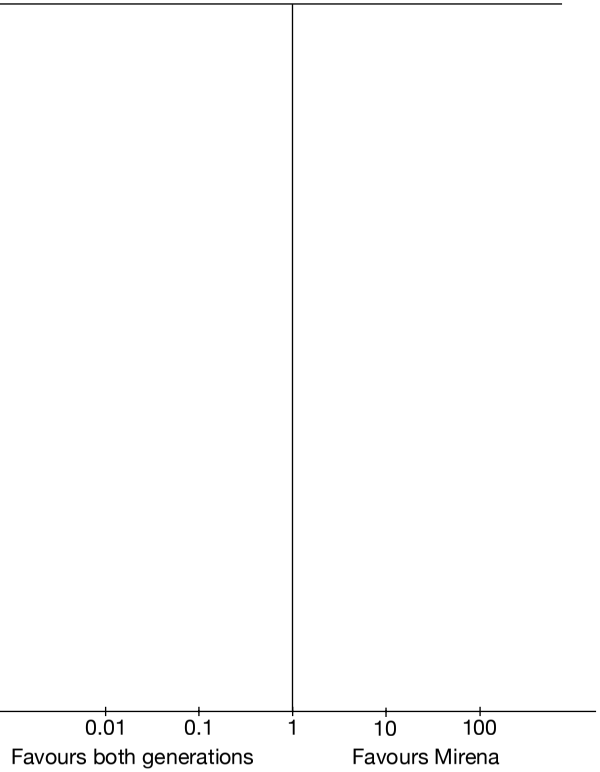
Patients with adverse events – postoperatively (abdominal pain)

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.81.1 First generation						
a. Malak 2006[104]	4	30	0	30	100.0%	8.22 (0.58 to 115.71)
b. Kittelsen 1998[105]	0	30	0	30		Not estimable
c. Crosignani 1997[45]	0	35	0	35		Not estimable
Subtotal (95% CI)		**95**		**95**	**100.0%**	**8.22 (1.10 to 61.49)**
Total events	4		0			
Heterogeneity: not applicable						
Test for overall effect: $z = 2.05$ ($p = 0.04$)						
1.81.2 Second generation						
d. Shaw 2007[108]	0	30	0	33		Not estimable
e. Tam 2006[109]	0	22	0	22		Not estimable
f. Busfield 2006[107]	0	41	0	42		Not estimable
g. Barrington 2003[81]	0	25	0	25		Not estimable
h. Soysal 2002[106]	0	36	0	36		Not estimable
Subtotal (95% CI)		**154**		**158**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Total (95% CI)		**249**		**253**	**100.0%**	**8.22 (1.10 to 61.49)**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: $z = 2.05$ ($p = 0.04$)						
Test for subgroup differences: not applicable						



Patients with adverse events – postoperatively (foul discharge)

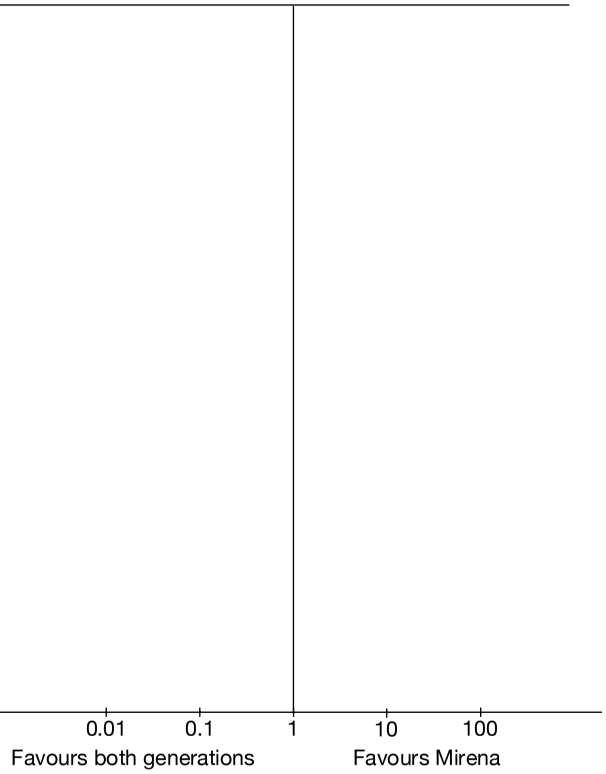
Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.82.1 First generation						
a. Malak 2006[104]	0	30	0	30		Not estimable
b. Kittelsen 1998[105]	0	30	0	30		Not estimable
c. Crosignani 1997[45]	0	35	0	35		Not estimable
Subtotal (95% CI)		**95**		**95**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
1.82.2 Second generation						
d. Shaw 2007[108]	0	33	0	33		Not estimable
e. Tam 2006[109]	0	22	0	22		Not estimable
f. Busfield 2006[107]	0	41	0	42		Not estimable
g. Barrington 2003[81]	0	25	0	25		Not estimable
h. Soysal 2002[106]	0	36	0	36		Not estimable
Subtotal (95% CI)		**157**		**158**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Total (95% CI)		**252**		**253**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Test for subgroup differences: not applicable						



DOI: 10.3310/hta15190

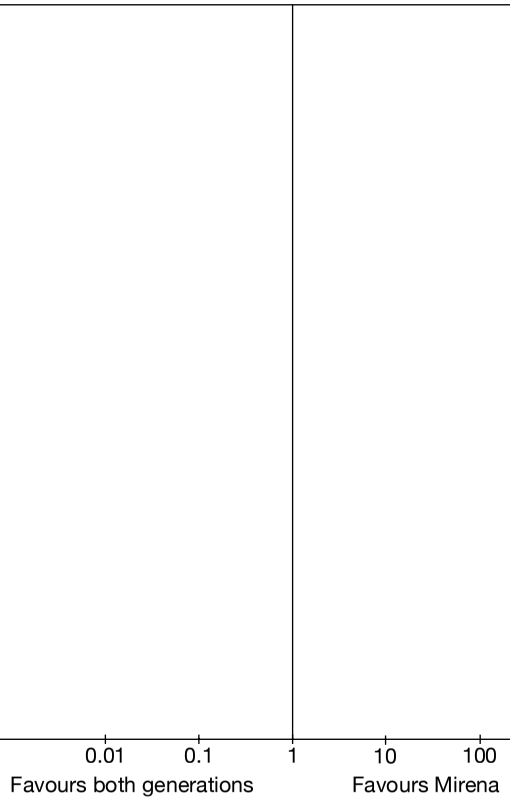
Patients with adverse events – postoperatively (visceral damage)

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.83.1 First generation						
a. Malak 2006[104]	0	30	0	30		Not estimable
b. Kittelsen 1998[105]	0	30	0	30		Not estimable
c. Crosignani 1997[45]	0	35	0	35		Not estimable
Subtotal (95% CI)		**95**		**95**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
1.83.2 Second generation						
d. Shaw 2007[108]	0	33	0	33		Not estimable
e. Tam 2006[109]	0	22	0	22		Not estimable
f. Busfield 2006[107]	0	41	0	42		Not estimable
g. Barrington 2003[81]	0	25	0	25		Not estimable
h. Soysal 2002[106]	0	36	0	36		Not estimable
Subtotal (95% CI)		**157**		**158**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Total (95% CI)		**252**		**253**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Test for subgroup differences: not applicable						



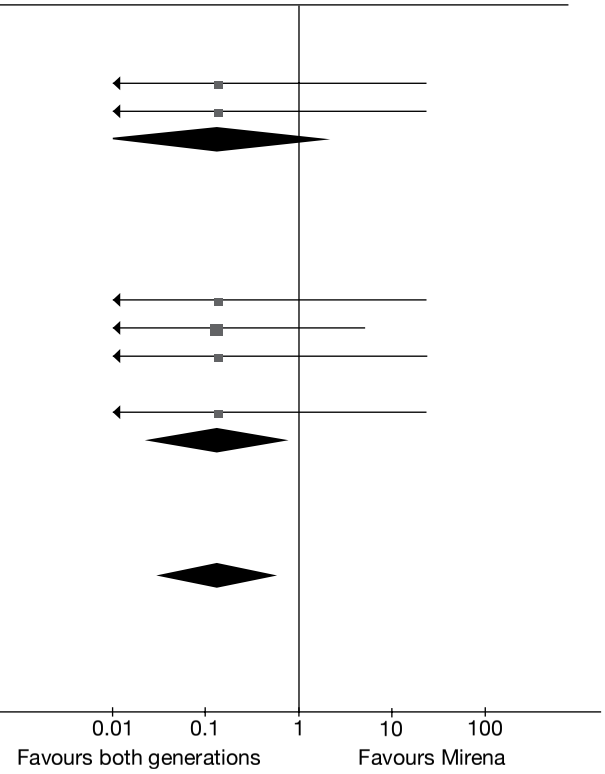
Patients with adverse events – postoperatively (infection)

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.84.1 First generation						
a. Malak 2006[104]	0	30	0	30		Not estimable
b. Kittelsen 1998[105]	0	30	0	30		Not estimable
c. Crosignani 1997[45]	0	35	0	35		Not estimable
Subtotal (95% CI)		**95**		**95**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
1.84.2 Second generation						
d. Shaw 2007[108]	0	33	0	33		Not estimable
e. Tam 2006[109]	0	22	0	22		Not estimable
f. Busfield 2006[107]	0	41	0	42		Not estimable
g. Barrington 2003[81]	0	25	0	25		Not estimable
h. Soysal 2002[106]	0	36	0	36		Not estimable
Subtotal (95% CI)		**157**		**158**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Test for subgroup differences: not applicable						
Total (95% CI)		**252**		**253**		**Not estimable**
Total events	0		0			
Heterogeneity: not applicable						
Test for overall effect: not applicable						
Test for subgroup differences: not applicable						



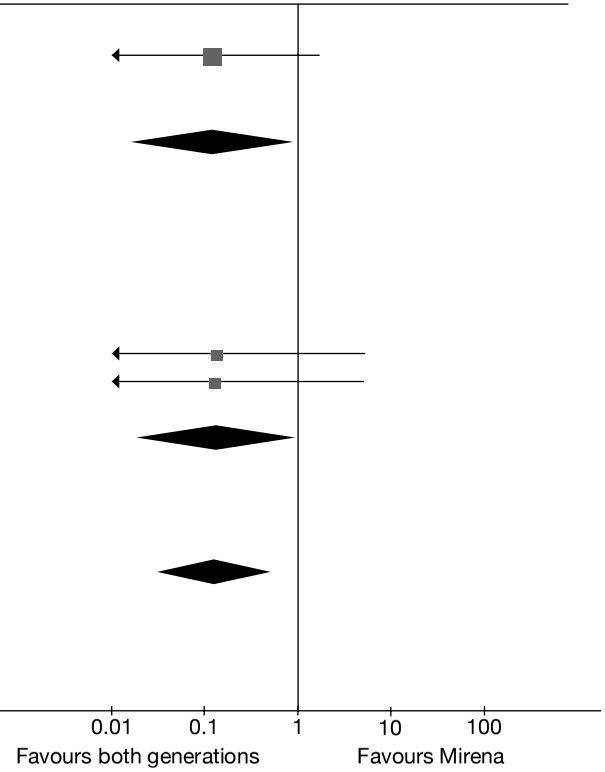
Patients with adverse events – postoperatively (migrated coil)

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.85.1 First generation						
a. Malak 2006[104]	0	30	0	30		Not estimable
b. Kittelsen 1998[105]	0	30	1	30	14.4%	0.14 (0.00 to 23.37)
c. Crosignani 1997[45]	0	35	1	35	14.4%	0.14 (0.00 to 23.37)
Subtotal (95% CI)		**95**		**95**	**28.8%**	**0.14 (0.01 to 2.16)**
Total events	0		2			
Heterogeneity: $\chi^2 = 0.00$, df = 1 ($p = 1.00$); $I^2 = 0\%$						
Test for overall effect: $z = 1.41$ ($p = 0.16$)						
1.85.2 Second generation						
d. Shaw 2007[108]	0	33	1	33	14.4%	0.14 (0.00 to 23.37)
e. Tam 2006[109]	0	22	2	22	28.1%	0.13 (0.00 to 5.15)
f. Busfield 2006[107]	0	41	1	42	14.4%	0.14 (0.00 to 23.95)
g. Barrington 2003[81]	0	25	0	25		Not estimable
h. Soysal 2002[106]	0	36	1	36	14.4%	0.14 (0.00 to 23.37)
Subtotal (95% CI)		**157**		**158**	**71.2%**	**0.13 (0.02 to 0.78)**
Total events	0		5			
Heterogeneity: $\chi^2 = 0.00$, df = 3 ($p = 1.00$); $I^2 = 0\%$						
Test for overall effect: $z = 2.24$ ($p = 0.03$)						
Total (95% CI)		**252**		**253**	**100.0%**	**0.13 (0.03 to 0.59)**
Total events	0		7			
Heterogeneity: $\chi^2 = 0.00$, df = 5 ($p = 1.00$); $I^2 = 0\%$						
Test for overall effect: $z = 2.65$ ($p = 0.008$)						
Test for subgroup differences: $\chi^2 = 0.00$, df = 1 ($p = 0.99$); $I^2 = 0\%$						



Patients with adverse events – postoperatively (removed <3 months)

Study or subgroup	Both generations Events	Total	Mirena Events	Total	Weight	Peto odds ratio Peto, Fixed, 99% CI
1.86.1 First generation						
a. Malak 2006[104]	0	30	4	30	49.1%	0.12 (0.01 to 1.71)
b. Kittelsen 1998[105]	0	30	0	30		Not estimable
c. Crosignani 1997[45]	0	35	0	35		Not estimable
Subtotal (95% CI)		**95**		**95**	**49.1%**	**0.12 (0.02 to 0.91)**
Total events	0		4			
Heterogeneity: not applicable						
Test for overall effect: $z = 2.05$ ($p = 0.04$)						
1.86.2 Second generation						
d. Shaw 2007[108]	0	33	0	33		Not estimable
e. Tam 2006[109]	0	22	0	22		Not estimable
f. Busfield 2006[107]	0	41	2	42	25.6%	0.14 (0.00 to 5.28)
g. Barrington 2003[81]	0	25	2	25	25.3%	0.13 (0.00 to 5.15)
h. Soysal 2002[106]	0	36	0	36		Not estimable
Subtotal (95% CI)		**157**		**158**	**50.9%**	**0.13 (0.02 to 0.96)**
Total events	0		4			
Heterogeneity: $\chi^2 = 0.00$, df = 1 ($p = 0.98$); $I^2 = 0\%$						
Test for overall effect: $z = 2.00$ ($p = 0.05$)						
Total (95% CI)		**252**		**253**	**100.0%**	**0.13 (0.03 to 0.52)**
Total events	0		8			
Heterogeneity: $\chi^2 = 0.00$, df = 2 ($p = 1.00$); $I^2 = 0\%$						
Test for overall effect: $z = 2.87$ ($p = 0.004$)						
Test for subgroup differences: $\chi^2 = 0.00$, df = 1 ($p = 0.95$); $I^2 = 0\%$						



Appendix 8

Pooled results for Mirena versus second-generation endometrial destruction

Appendix 8.1 Quality of life and clinical outcomes

	Time point	Trials (no.)	WMD (95% CI)[a]	OR (95%CI)[b]	*p*-value	Hetero (*p*)/I^2 (%)
Proportion amenorrhoea	6 months	2 (107)	–	1.76 (0.42 to 7.34)	0.4	0.5/0
	12 months	3 (127)	–	2.43 (0.61 to 9.67)	0.2	0.02/82
	2 years	1 (39)	–	5.99 (1.43 to 25.1)	0.01	–
Proportion with heavy bleeding	6 months	3 (162)	–	4.30 (1.76 to 10.6)	0.001	0.6/0
	12 months	4 (200)	–	1.54 (0.56 to 4.24)	0.4	0.02/75
	2 years	1 (39)	–	13.0 (2.00 to 84.2)	0.007	–
Bleeding score (change)	6 months	4 (169)	10 (–37 to 57)	–	0.7	0.2/29
	12 months	4 (168)	–35 (–75 to 5)	–	0.09	0.3/26
	2 years	1 (39)	117 (–113 to 347)	–	0.7	–
Proportion dysmenorrhoea	6 months	1 (83)	–	0.78 (0.33 to 1.85)	0.6	–
	12 months	1 (48)	–	0.77 (0.25 to 2.43)	0.7	–
	2 years	1 (83)	–	0.97 (0.31 to 3.05)	1.0	–
SF-36 general health (absolute)	12 months	3 (147)	3.7 (–2.7 to 10.1)	–	0.3	0.006/81
SF-36 physical function (absolute)		3 (146)	–0.9 (–6.9 to 5.0)	–	0.8	0.6/0
SF-36 role physical (absolute)		3 (147)	–9.6 (–20.5 to 1.3)	–	0.08	0.1/56
SF-36 role emotional (absolute)		3 (147)	5.7 (–5.4 to 16.8)	–	0.3	0.02/75
SF-36 mental health (absolute)		3 (146)	–0.8 (–6.7 to 5.1)	–	0.8	0.02/73
SF-36 social function (absolute)		3 (147)	2.3 (–3.6 to 8.2)	–	0.4	0.06/65
SF-36 vitality (absolute)		3 (146)	0.01 (–5.9 to 6.1)	–	1.0	0.3/7
SF-36 pain (absolute)		3 (147)	1.1 (–6.6 to 8.9)	–	0.8	1.0/0
SF-36 general health (absolute)	2 years	1 (49)	1.7 (–10.3 to 13.7)	–	0.8	–
SF-36 physical function (absolute)		1 (50)	–1.2 (–8.0 to 5.6)	–	0.7	–
SF-36 role physical (absolute)		1 (50)	8.2 (–4.9 to 21.3)	–	0.2	–
SF-36 role emotional (absolute)		1 (50)	–1.0 (–16.4 to 14.4)	–	0.9	–
SF-36 mental health (absolute)		1 (49)	–0.4 (–8.9 to 8.1)	–	0.9	–
SF-36 social function (absolute)		1 (50)	2.2 (–7.0 to 11.4)	–	0.6	–
SF-36 vitality (absolute)		1 (49)	1.8 (–8.6 to 12.2)	–	0.7	–
SF-36 pain (absolute)		1 (50)	8.7 (–2.3 to 19.7)	–	0.1	–

continued

	Time point	Trials (no.)	WMD (95% CI)[a]	OR (95%CI)[b]	*p*-value	Hetero (*p*)/I^2 (%)
SF-36 general health (change)	12 months	2 (81)	1.9 (–6.5 to 10.4)	–	0.7	0.1/59
SF-36 physical function (change)		2 (81)	–0.5 (–10.1 to 9.2)	–	0.9	0.5/0
SF-36 role physical (change)		2 (82)	–7.0 (–28.9 to 14.9)	–	0.5	0.8/0
SF-36 role emotional (change)		2 (82)	1.8 (–18.5 to 22.1)	–	0.9	0.04/75
SF-36 mental health (change)		2 (81)	–2.1 (–9.6 to 5.4)	–	0.6	0.03/78
SF-36 social function (change)		2 (82)	–1.9 (–10.9 to 7.1)	–	0.7	0.2/49
SF-36 vitality (change)		2 (80)	–2.0 (–12.2 to 8.2)	–	0.7	0.7/0
SF-36 pain (change)		2 (82)	–8.0 (–20.0 to 3.9)	–	0.2	0.6/0
SF-36 general health (change)	2 years	1 (49)	–4.1 (–14.0 to 5.8)	–	0.4	–
SF-36 physical function (change)		1 (50)	–4.4 (–14.9 to 6.1)	–	0.4	–
SF-36 role physical (change)		1 (50)	–1.1 (–28.4 to 26.2)	–	0.9	–
SF-36 role emotional (change)		1 (50)	2.0 (–23.3 to 27.3)	–	0.9	–
SF-36 mental health (change)		1 (49)	–5.7 (–14.2 to 2.8)	–	0.2	–
SF-36 social function (change)		1 (50)	–4.9 (–17.6 to 7.8)	–	0.5	–
SF-36 vitality (change)		1 (49)	–9.6 (–20.4 to 1.2)	–	0.08	–
SF-36 pain (change)		1 (50)	1.3 (–17.8 to 15.2)	–	0.9	–
Hysterectomy after EA/Mirena	6 months	1 (50)	–	1.79 (0.40 to 8.01)	0.5	–
	12 months	2 (105)	–	0.22 (0.05 to 1.04)	0.06	0.9/0
	2 years	2 (146)	–	0.86 (0.33 to 0.21)	0.8	0.2/41
		Trials	**Frequency**			
Discontinued Mirena	6 months	1	3/42 (7%)			
	12 months	3	18/96 (19%)			
	2 years	2	21/75 (28%)			
EA after Mirena	12 months	1	3/18 (17%)			
	2 years	1	4/42 (10%)			
	6 months	1	3/25 (12%)			
Hysterectomy after Mirena	12 months	2	6/54 (11%)			
	2 years	2	11/75 (15%)			

	Trials	Frequency (second-generation: max. 157; Mirena: max. 158)	OR (95% CI)[b]	*p*-value	Hetero (*p*)/I^2 (%)
Periprocedure complications					
Uterine perforation (second-generation, Mirena)	5	0; 0	–	–	–
Cervical laceration (second-generation, Mirena)	5	0; 0	–	–	–
Anaesthesia problems (second-generation)	5	0	–	–	–
Excessive bleeding (second-generation)	5	0	–	–	–
Fluid overload (second-generation)	5	0	–	–	–
Visceral damage (second-generation)	5	0	–	–	–
Procedure abandoned (second-generation)	5	0	–	–	–
Converted to hysterectomy (second-generation)	5	0	–	–	–
Failed to insert (Mirena)	5	0	–	–	–

	Trials	Frequency (second-generation: max. 157; Mirena: max. 158)	OR (95% CI)[b]	*p*-value	Hetero (*p*)/I^2 (%)
Further complications (< 1 month)					
Urinary tract infection (second-generation)	5	0	–	–	–
Deep-vein thrombosis (second-generation)	5	0	–	–	–
Further bleeding (second-generation)	5	0	–	–	–
Sepsis (second-generation)	5	0	–	–	–
Pyrexia (second-generation)	5	0	–	–	–
Endometriosis (second-generation)	5	0	–	–	–
Haematomata (second-generation)	5	0	–	–	–
Abdominal pain (second-generation)	5	0	–	–	–
Foul discharge (second-generation)	5	0	–	–	–
Visceral damage (second-generation)	5	0	–	–	–
Infection (Mirena)	5	0	–	–	–
Expelled/migrated (Mirena)	5	5	–	–	–
Removed before 3 months (MIrena)	5	4	–	–	–

a < 0 favours second-generation EA, > 0 favours Mirena.
b < 1 favours second-generation EA, > 1 favours Mirena.

Appendix 9 (for Chapter 5)

Survey of gynaecologists with expertise in minimal access surgery

Dear Dr _____

We would value your opinion as an expert in gynaecological surgery on the outcome of a recent Department of Health (Health Technology Assessment Panel)-funded systematic review of the evidence for the treatment of heavy menstrual bleeding (HMB). Our aim was to evaluate the comparative effectiveness and cost-effectiveness of hysterectomy, Mirena® and second-generation endometrial ablation (microwave, balloon and NovaSure). We aggregated and analysed results of trials comparing endometrial ablations with hysterectomy and each other and also trials comparing Mirena® with hysterectomy and with ablation. There were very few trials in the last category.

We would be grateful if you could read the summary of our key findings on the comparative clinical effectiveness of the alternative treatments for HMB and answer the questions below.

Your answers will provide much needed guidance to us in interpreting the results of our review and will inform the recommendation in our final report to the HTA.

Thank you for taking the time to read this letter

On behalf of the HMB IPD Collaborative Group

Kevin Cooper, Patrick Chien, Peter O'Donovan, Khalid Khan, Siladitya Bhattacharya

Our findings (based on individual patient data and aggregated data meta-analysis of randomised trials) suggest:

* At 12 months after treatment, more women (21/382 or 12.6% vs 57/454 or 5.3%) were dissatisfied with first-generation hysteroscopic techniques than hysterectomy (OR 2.46; 95% CI 1.54 to 3.93; $p = 0.0002$), but hospital stay (WMD 3.0 days; 95% CI 2.9 to 3.1 days; $p < 0.00001$) and time to resumption of normal activities (WMD 5.2 days; 95% CI 4.7 to 5.7 days; $p < 0.00001$) were longer for hysterectomy.

* Indirect estimates (*Figure 6*) suggest hysterectomy is also preferable to second-generation ED (OR 2.32; 95% CI 1.27 to 4.24; $p = 0.006$) in terms of patient dissatisfaction.

* Hysterectomy is cheaper and more effective than either first- or second-generation endometrial ablation but carries a higher risk of complications.

* Satisfaction rates were comparable between first- and second-generation techniques (OR 1.20; 95% CI 0.88 to 1.62; $p = 0.2$), although second-generation techniques were quicker (WMD 14.5 minutes; 95% CI 13.7 to 15.3; $p < 0.00001$) and women recovered sooner (WMD 0.48 days; 95% CI 0.20 to 0.75; $p = 0.0008$) with fewer procedural complications.

* Second-generation techniques are cheaper and more effective than first-generation techniques.

* There are few comparisons of Mirena® versus more invasive procedures. The few data available suggest that Mirena® is potentially cheaper and more effective than first-generation ablation techniques with rates of satisfaction which are similar to second-generation ED (18.1% vs 22.5%; OR 0.76; 95% CI 0.38–1.53; $p = 0.4$).

* Owing to a paucity of trials, the evidence to suggest hysterectomy is preferable to Mirena® is weak (OR 2.22; 95% CI 0.94–5.29; $p = 0.07$). In a single study comparing hysterectomy with Mirena®, QoL was similar in both groups although residual pelvic pain was less common after hysterectomy.

* Hysterectomy is more expensive than Mirena® (ICER = 1600).

Based on these data, could you please answer the questions below by deleting all responses other than the most appropriate one.

Please send in your answers by replying to this email:

1: What would you consider to be first-line treatment in women with HMB and failed oral medical treatment associated with no obvious clinical abnormalities.

N.B. Any uterine fibroids present are < 3 cm and do not impinge on the endometrial cavity.

i. * Mirena®
ii. * First-generation endometrial ablation
iii. * Second-generation endometrial ablation
iv. * Hysterectomy

2: If the first treatment fails, what in your view should be the next treatment:

i. * Mirena®
ii. * Second-generation ablation
iii. * First-generation ablation (e.g. rollerball)
iv. * Hysterectomy

3: If the second treatment fails what, in your view, should be the next line treatment

i. * Mirena®
ii. * Repeat second-generation ablation
iii. * First-generation ablation (e.g. rollerball)
iv. * Hysterectomy

Appendix 10

PRISMA checklist

Section/topic	#	Checklist item	Reported on page #
Title			
Title	1	Identify the report as a systematic review, meta-analysis or both	11
Abstract			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number	N/A
Introduction			
Rationale	3	Describe the rationale for the review in the context of what is already known	11
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS)	11
Methods			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g. URL), and, if available, provide registration information including registration number	11
Eligibility criteria	6	Specify study characteristics (e.g. PICOS, length of follow-up) and report characteristics (e.g. years considered, language, publication status) used as criteria for eligibility, giving rationale	11
Information sources	7	Describe all information sources (e.g. databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched	11
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated	*Appendix 1*
Study selection	9	State the process for selecting studies (i.e. screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis)	12
Data collection process	10	Describe method of data extraction from reports (e.g. piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators	12
Data items	11	List and define all variables for which data were sought (e.g. PICOS, funding sources) and any assumptions and simplifications made	12, 13
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis	13
Summary measures	13	State the principal summary measures (e.g. risk ratio, difference in means)	13
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g. I^2) for each meta-analysis	13
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g. publication bias, selective reporting within studies)	13
Additional analyses	16	Describe methods of additional analyses (e.g. sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified	13, 14

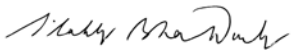
continued

Section/topic	#	Checklist item	Reported on page #
Results			
Study selection	17	Give numbers of studies screened, assessed for eligibility and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram	13, 14, *Figure 1*
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations	*Appendix 2*
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see *Item 12*)	15,16, *Figure 2*, *Appendix 3*
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group; (b) effect estimates and confidence intervals, ideally with a forest plot	17–22, *Figures 3–6*, *Appendices 4–8*
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency	17–22, *Figures 3–6*, *Appendices 4–8*
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see *Item 15*)	17–22
Additional analysis	23	Give results of additional analyses, if done (e.g. sensitivity or subgroup analyses, meta-regression) (see *Item 16*)	17–22
Discussion			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g. health-care providers, users and policy makers)	22, 23
Limitations	25	Discuss limitations at study and outcome level (e.g. risk of bias) and at review level (e.g. incomplete retrieval of identified research, reporting bias)	23
Conclusions	26	Provide a general interpretation of the results in the context of other evidence and implications for future research	24
Funding			
Funding	27	Describe sources of funding for the systematic review and other support (e.g. supply of data); role of funders for the systematic review	N/A

For more information, visit: www.prisma-statement.org.

Appendix 11

Protocol



Signed: Siladitya Bhattacharya

Dated 5 February 2010

The effectiveness of hysterectomy, ablation and levonorgestrel-releasing intra-uterine device in the management of heavy menstrual bleeding

Background

Heavy menstrual bleeding (menorrhagia) is a common problem. It affects nearly one-third of women (Corrado, 1990; Rees, 1991) and prompts 5% of all women of reproductive age to consult their general practitioners with menstrual problems. Menstrual disorders account for 20% of gynaecology outpatient referrals and are responsible for over 23,000 hysterectomies each year in England. One in five women in the United Kingdom is likely to have had a hysterectomy by the age of 55 years (Vessey *et al.*, 1992). HMB affects many aspects of everyday life – including work as well as social activities – and leads to a measurable reduction in QoL.

A literature search was undertaken using the Cochrane Library, MEDLINE (1966–2006), EMBASE (1980 to July 2006) and CINAHL (1982 to July 2006) using the following terms: menorrhagia, hypermenorrhea, (excessive) menstrual blood loss, heavy menstrual bleeding, dysfunctional uterine bleeding, hysterectomy, vaginal hysterectomy, total abdominal hysterectomy, subtotal abdominal hysterectomy, laparoscopic hysterectomy, transcervical resection of the endometrium, transcervical resection of the endometrium, endometrial ablation, laser ablation, hysteroscopy, electrosurgery, rollerball, (thermal) balloon, hypertherm(ia), thermotherapy, photodynamic therapy, phototherapy, cryoablation, microwave endometrial ablation, radiofrequency, saline irrigation, laser interstitial, ThermaChoice®, Cavaterm™, ELITT, Vesta, NovaSure, Microsulis, Cryogen. The metaregister of controlled trials and the ISRCTN register were searched for any trials with menorrhagia and endometrial ablation as keywords.

Current recommendations in the UK promote medical methods for the initial management of HMB. Mefenamic acid, tranexamic acid and the combined oral pill are considered to be suitable first-line drugs [Royal College of Obstetricians and Gynaecologists (RCOG) guideline, 1998]. The LNG-releasing IUS (Mirena®) is an effective non-surgical treatment which is reversible and fertility-sparing. It reduces estimated menstrual blood loss by up to 96% by 12 months, with up to 44% of users reporting amenorrhoea (Milsom *et al.*, 1991; Lahteenmaki *et al.*, 1998), at a cost which is one-third that for hysterectomy (Hurskainen *et al.*, 2001). Despite the availability of these options, long-term medical treatment is unsuccessful or unacceptable in many and surgery is required (Cooper *et al.*, 2001).

Hysterectomy offers a definitive treatment for menorrhagia and guarantees amenorrhoea, but it is particularly invasive and carries significant morbidity (Lethaby *et al.*, 1999). Overall 1 in 30 women suffers a major adverse event, and the mortality rate is 0.4–1.1 per 1000 operations. The need for GA, prolonged hospital stay and delayed recovery also makes hysterectomy an expensive treatment (Cameron *et al.*, 1996).

Endometrial ablative techniques aimed at destruction of the functionally active endometrium along with some of the underlying myometrium (Duffy *et al.*, 1991; Duffy *et al.*, 1992) offer a conservative surgical alternative to hysterectomy. The first-generation ablative techniques including endometrial laser ablation (ELA) (Goldrath and Fuller, 1981; Davis, 1989), TCRE (Magos *et al.*, 1989) and REA were all endoscopic procedures. Although they do not guarantee amenorrhoea, their effectiveness (in comparison with hysterectomy – the existing gold standard) has been demonstrated in a number of RCTs (Gannon *et al.*, 1991; Dwyer *et al.*, 1993; Pinion *et al.*, 1994, O'Connor *et al.*, 1997; Crosignani *et al.*, 1997; Aberdeen Endometrial Ablation Trials Group, 1999).

National audits (Overton *et al.*, 1997; Scottish Hysteroscopy Audit Group, 1995) revealed that although first-generation ablative techniques were less morbid than hysterectomy they were associated with a number of complications including uterine perforation, cervical laceration, false passage creation, haemorrhage, sepsis and bowel injury. In addition they were also related to fluid overload associated with the use of 1.5% urological glycine (non-ionic) irrigation fluid in TCRE and RBA, resulting in serious and occasionally fatal consequences due to hyponatraemia (Arrief and Ayus, 1993; Rosenberg, 1995). Mortality from these techniques has been estimated at 0.26 per 1000 (Overton *et al.*, 1997; Scottish Hysteroscopy Audit Group, 1995).

Second-generation ablative techniques represent simpler, quicker and potentially more efficient means of treating menorrhagia, which require less skill on the part of the operator. Examples of second-generation ablative techniques are fluid-filled TBEA, radiofrequency (thermoregulated) balloon EA, hydrothermal EA, three-dimensional bipolar radiofrequency EA, MEA, diode laser hyperthermy, cryoablation and photodynamic therapy. The most common techniques in the UK are TBEA (ThermaChoice and Cavaterm) (Loffer, 2001; Loffer and Grainger, 2002; Meyer *et al.*, 1998) and MEA (Cooper *et al.*, 1999; Bain *et al.*, 2002), while the NovaSure device (Cooper *et al.*, 2002) is gaining in popularity. TBEA destroys the endometrium by means of heated liquid within a balloon inserted into the uterine cavity. It cannot be used in women with large or irregular uterine cavities. MEA uses microwave energy (at a frequency of 9.2 GHz) to destroy the endometrium. Complications associated with second-generation techniques include equipment failure, uterine infection, perforation, visceral burn, bleeding and cyclical pain. A limited number of randomised trials indicate that these procedures appear to be as effective as first-generation ablative techniques (Lethaby *et al.*, 2005). In addition, some have the added benefit of being performed under LA.

The introduction of new EA techniques over the last two decades has been accompanied by a series of RCTs aimed at evaluating their clinical effectiveness and cost-effectiveness. Initially, first-generation EA techniques such as TCRE and laser ablation were compared with hysterectomy (Lethaby *et al.*, 1999). Subsequent trials, which compared alternative first-generation techniques such as TCRE, laser EA and REA, established TCRE as the gold standard for this group of treatments. As less invasive and more user-friendly second-generation techniques such as MEA became available, these were compared with earlier methods of ablation like TCRE and REA. Although not all techniques have been subjected to head-to-head comparisons in the context of randomised trials, an overview of the literature demonstrates that MEA (second generation) has

been shown to be comparable with TCRE (first generation) – which, in turn, has been shown to be an effective alternative to hysterectomy (gold standard). However, questions about long-term clinical effectiveness and cost implications of alternative forms of surgical treatment remain unanswered. Published data report no more than 5 years of follow-up (Aberdeen Endometrial Ablation Trials Group, 1999; Cooper *et al.*, 2005). Inevitably, some women treated by EA will eventually require repeat ablation or hysterectomy. Following hysterectomy, a proportion of women will also develop further complications such as postsurgical adhesions and pelvic floor dysfunction, which may lead to further surgery. The necessity for a head-to-head comparison between the two most common second-generation methods – MEA and TBEA – has been identified (NICE, 2004). Our group has recently completed recruitment to such a trial involving over 200 women funded by the Chief Scientist Office Scotland (CZH/4/117) (Sambrook, unpublished). Given the widespread use of ablative techniques as first-line surgical treatment for menorrhagia at the present time, it is uncertain whether it is either necessary or feasible to compare second-generation techniques directly with hysterectomy in a new randomised trial, which is unlikely to produce any meaningful results for another 4–5 years. At the same time, the need to obtain comparative information on long-term outcomes is clearly accepted, as is the need to identify the best technique for individual women.

From a clinical perspective, we believe that the most relevant research questions at the present time are:

i. How do the currently used ablative techniques and the Mirena IUS system compare with hysterectomy in the medium to long term?
ii. Which among the commonly used second-generation ablation techniques is the most effective and cost-effective?
iii. Are there subgroups of women who are most likely to benefit from hysterectomy, Mirena or specific types of ablation?

We propose to address these questions by analysis of data from national data sets and randomised trials. We plan to assess long-term outcomes by means of record linkage and follow-up of randomised cohorts, and perform IPD meta-analysis of existing trial data. The output will be used to create a model for the utilisation and costs of the different treatments, which can inform an algorithm for clinical decision making.

Overall aims of the project:

1. To determine, using data from record linkage and follow-up of randomised and non-randomised cohorts of British women, long-term effects of various second-generation ablative techniques and hysterectomy in terms of failure rates, complications and further surgery.
2. To determine, using IPD meta-analysis of existing RCTs, short- to medium-term effects of various second-generation ablative techniques, Mirena IUS and hysterectomy, including exploration of outcomes in clinical subgroups.
3. To undertake a model-based clinical effectiveness and cost-effectiveness analysis comparing Mirena IUS and various second-generation ablative techniques with hysterectomy using output from the above analyses and to conduct extensive sensitivity analyses to explore robustness of the results to the assumptions made.
4. To devise a parsimonious algorithm for clinical decision making regarding the choice of surgery for women with HMB with failed medical treatment.

Record linkage study protocol

Research Group (Aberdeen), Professor S Bhattacharya,[1] Dr K Cooper,[2] Dr P Chien,[3] Professor A Lee[1] and Dr V Timmuraju[1]

[1]University of Aberdeen, Aberdeen Maternity Hospital, Foresterhill, Aberdeen, UK
[2]University of Aberdeen, Department of Obstetrics & Gynaecology, Grampian Hospitals, Aberdeen, UK
[3]Ninewells Hospital, Dundee, UK

Aim

To determine, using data from record linkage and follow-up of randomised and non-randomised cohorts of British women, long-term effects of various second-generation ablative techniques and hysterectomy in terms of failure rates, complications, QoL and sexual function.

This will be addressed by means of:

Analysis of a large population-based anonymised, observational data set generated by the ISD Scotland, in order to identify medium- and long-term effects of hysterectomy and second-generation EA techniques. This will overcome some of the potential limitations of data from trials which are based on relatively small numbers of women. This is thus an area where observational data will be invaluable in assessing outcomes in all categories of women rather than the highly selected group who have been recruited to trials.

This aim has had to be modified as long-term data on QoL and sexual function as well as variables listed in the previous analysis plan (uterine size, presence of fibroids, coexisting gynaecological pathology) are not available in the ISD data set.

Predictor variables which are available in the ISD data set include age, type of procedure, CARSTAIR quintile for social deprivation, year of operation and cancer.

Analytical approach

Data sets

Population-based routinely collected data will be used in the analysis to meet this objective. We have confirmation of availability of access to population-based data in Scotland. An initial search within the ISD data set has identified over 40,000 hysterectomies (1985–2005) and 14,000 ablative techniques (1989–2005) performed in women with DUB. This includes a subset of women randomised to alternative treatments for menorrhagia. The custodians of the ISD registry have given their approval to proceed along these lines and have agreed to generate an anonymised data set for analysis.

Analysis

Descriptive statistics will be used to summarise each of the outcomes and potential predictor variables (age, type of procedure, uterine size, presence of fibroids, coexisting gynaecological pathology). Appropriate univariate analyses (two sample *t*-test, chi-squared test and non-parametric tests) will be used initially to examine the association between these potential predictors and the outcomes of interest (repeat surgery, hysterectomy, other pelvic surgery).

Multiple logistic regression techniques will be used to examine the mutually adjusted effects of potential predictors identified in the univariate analysis. The predictive ability of the models will not be assessed by determination of the area under the ROC curve owing to the unavailability of the predictor variables (uterine size, presence of fibroids, coexisting gynaecological pathology). Comparison of the predictive ability of models incorporating only two variables using area under the ROC curve was therefore deemed inappropriate. The analysis will generally be carried out stratified by the women's age group.

Appropriate univariate analyses (chi-squared test; *t*-test) will examine the association between the ISD-linked Scottish randomised trial women and future retreatment. The women will be analysed by appropriate subgroups. Multiple logistic regression will be used to quantify the risk of treatment failure among subgroups of women after adjustment for confounders such as age, CARSTAIR quintiles, year of operation and cancer.

Sample size

From the ISD data set, we envisage assembling a cohort of at least 13,000 women post ablation and 40,000 post hysterectomy. With a data set of 13,000 ablations, the two-sided 95% CI around an estimated prevalence of retreatment of 25% would be 24.3%–25.7%.

The effectiveness of hysterectomy, ablation and levonorgestrel-releasing intra-uterine device: individual patient data meta-analysis

The International HMB (Heavy Menstrual Bleeding) IPD Meta-analysis Collaborative Group, Management Group Aberdeen, UK, S Bhattacharya,[1] K Cooper,[2] KS Khan,[3] J Daniels,[3] L Middleton,[4] R Champaneria[3] and R Gray[4]

[1]University of Aberdeen, Aberdeen Maternity Hospital, Foresterhill, Aberdeen, UK
[2]University of Aberdeen, Department of Obstetrics & Gynaecology, Grampian Hospitals NHS Trust, Foresterhill, Aberdeen, UK
[3]Birmingham Women's Hospital, Edgbaston, Birmingham, UK
[4]Birmingham Clinical Trials Unit, Robert Aitken Institute, University of Birmingham, Birmingham, UK

Aim

To determine, using IPD meta-analysis of existing RCTs, short- to medium-term effects of various second-generation ablative techniques, Mirena IUS and hysterectomy, including exploration of outcomes in clinical subgroups.

Objectives

To assess the comparative effectiveness of hysterectomy, ablative techniques and LNG IUS for the treatment of menorrhagia using the following comparisons:

1. hysterectomy versus ablation
2. ablation versus ablation (comparison of different techniques)
3. ablation versus LNG IUS
4. hysterectomy versus LNG IUS.

Eligibility

Types of studies

Studies will only be included if they are RCTs with adequate randomisation concealment, excluding quasi-randomisation and non-randomisation.

Types of participants

Inclusion criteria

Participants in the trials will be included in IPD meta-analysis if women have menorrhagia or abnormal/excessive/prolonged uterine bleeding that is unresponsive to medical treatment without obvious clinically detectable underlying pathology.

As many of the trials have been pragmatic, prior hysteroscopy will not have been performed. Thus, they will include women with small fibroids.

Exclusion criteria

Participants in the trial who have uterine bleeding caused by polyps and other uterine pathologies will not be included in the main IPD meta-analysis or, if considered necessary, will be analysed as a subgroup

Types of intervention

Randomised controlled trials comparing hysterectomy, endometrial resection or ablation, and LNG IUS in any of the combinations laid out in the *Objectives* section. *Table 1* shows the range of interventions that will be included.

Types of outcome measures

Primary outcomes

The primary outcome of interest is subjective reduction in MBL. Any studies that do not include a measurement of MBL will be excluded. MBL can be assessed in a number of ways including VAS or PBAC.

Secondary outcomes

Other outcomes will be collected for meta-analysis to investigate the effect of the interventions on other aspects of HMB, adverse effects and resource implications. These will include:

- patient satisfaction
- safety of procedure (morbidity, adverse effects, operative complications)
- length of operating time
- length of hospital stay
- fluid deficit
- pain
- anxiety, depression, sexual functioning
- long-term complications
- QoL
- health-related QoL
- pre-menstrual symptoms
- repeated surgery for HMB.

TABLE 1 Intervention groups and surgical techniques

Intervention	Type	Trade name
Hysterectomy	Total (both the body of uterus and cervix removed)	
	Subtotal (the body of the uterus is removed, leaving the cervix in place)	
	± Salpingo-oophorectomy	
	± Bilateral salpingo-oophorectomies	
	Wertheim (will be excluded) (body of uterus and cervix, part of the vagina, fallopian tubes, usually the ovaries, parametrium – the broad ligament below the fallopian tubes – and lymph glands and fatty tissue in the pelvis removed. This type of hysterectomy is also called a radical hysterectomy)	
Ablation – endometrial	First generation	
	TCRE	
	RBI	
	Laser (Nd:YAG)	
	Second generation	
	Thermal balloon	ThermaChoice, Cavaterm
	Hydrothermal	
	3D bipolar radiofrequency	
	Microwave	NovaSure
	Diode laser hyperthermy	
	Cryoablation	
	Photodynamic therapy	
LNG IUS	LNG IUS	Mirena coil

Methods

An overview of the process of collecting and synthesising data is shown in *Figure 1*.

Literature searching

An original literature search was undertaken using the Cochrane Library, MEDLINE (1966–2007), EMBASE (1980 to July 2007) and CINAHL (1982 to July 2007).

To select studies of surgical interventions for menorrhagia the following search terms were used: menorrhagia, hypermenorrhea, (excessive) menstrual blood loss, heavy menstrual bleeding, dysfunctional uterine bleeding, hysterectomy, vaginal hysterectomy, total abdominal hysterectomy, subtotal abdominal hysterectomy, laparoscopic hysterectomy, transcervical resection of the endometrium, TCRE, endometrial ablation, laser ablation, hysteroscopy, electrosurgery, rollerball, (thermal) balloon, hypertherm(ia), thermotherapy, photodynamic therapy, phototherapy, cryoablation, microwave endometrial ablation, radiofrequency, saline irrigation, laser interstitial, ThermaChoice, Cavaterm, ELITT, Vesta, NovaSure, Microsulis, Cryogen.

To identify any ongoing RCTs the following were searched: the Meta-Register of Controlled Trials and the ISRCTN register with 'menorrhagia' and 'endometrial ablation' as keywords.

All identified trials are shown in *Appendix A*.

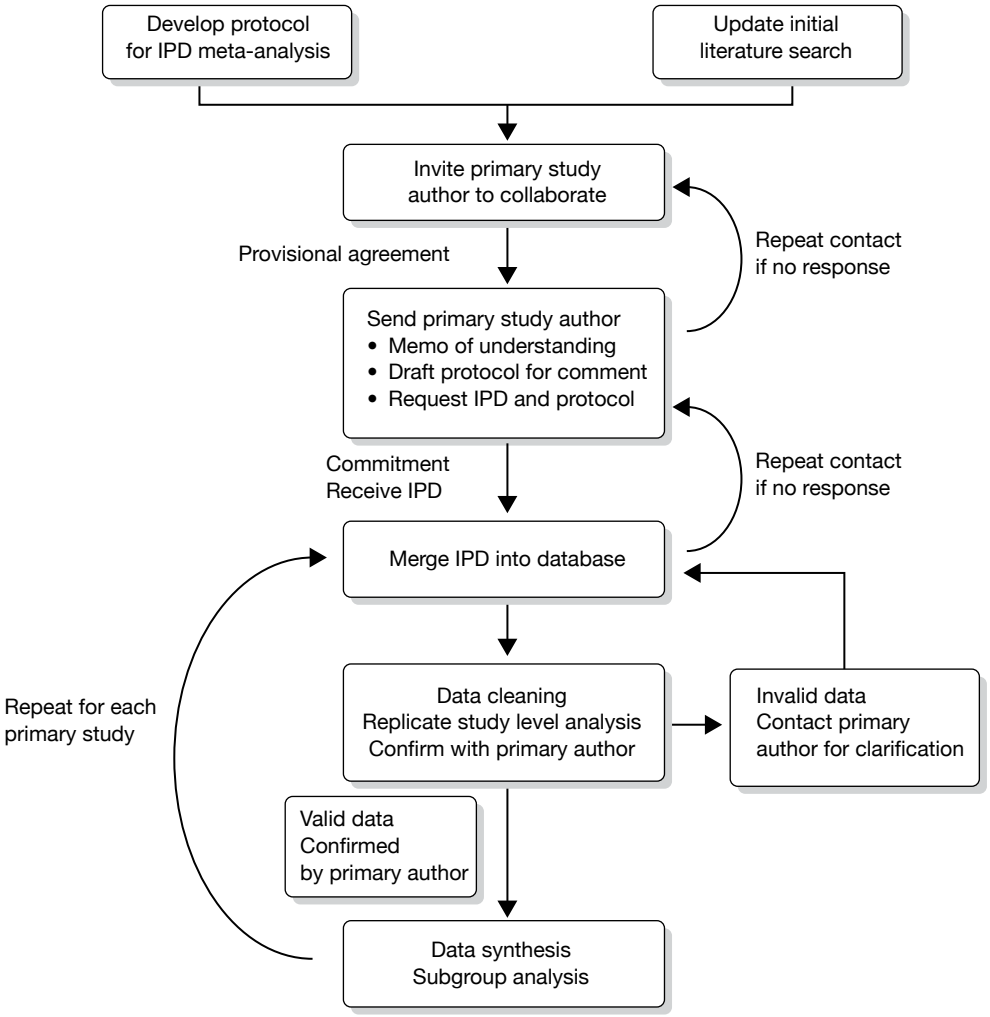


FIGURE 1 Summary of steps in undertaking the HMB IPD meta-analysis.

The search will be repeated every three months throughout the project to ensure any newly published studies are identified. *Appendix B* gives the full search strategy.

Once the collaborative group has been established, investigators from the identified studies will be asked to review the included study list to identify any studies that might have been missed.

Collection of IPD from authors of primary RCTs

Initial contact has already been made with the first named author of the included primary studies. Authors that have not as yet responded to the initial invitation will be sent another letter. If attempts from investigators within the collaboration fail, they may be contacted via the British or International Society for Gynaecological Endoscopy. Confirmation of commitment to the collaboration and ability to supply IPD will then be sought. The responding authors will be sent the overview protocol and a request to send the trial data set, original study protocol and data collection forms. The data can be supplied in either a Microsoft ACCESS database (preferred choice) or a Microsoft EXCEL spreadsheet.

Inclusion in the collaborative group and provision of data will be covered by a Memorandum of Understanding.

Data requested will include primary and secondary outcomes. In addition, the baseline demographic and clinical details listed below will need to be collected:

- age at randomisation
- parity
- uterine cavity length
- presence of fibroids and/or polyps
- number of previous caesarean sections.

All data received will be incorporated into an overview database, taking care to preserve any referential integrity within relational databases. All the data supplied will be subjected to range and consistency checks. Any missing data, obvious errors, inconsistencies between variables or outlying values will be queried and rectified as necessary by correspondence with the investigators. Study level analysis will be repeated to verified published results.

Once the data have been checked and validated, the original authors will be contacted to confirm their acceptance of individual study results before proceeding to the meta-analysis. If the integrity of the data/study is questionable they may be excluded from the analysis.

Data synthesis

Statistical analysis will be carried out on all the patients ever randomised, and will be based on the ITT principle. Results from separate trials will be combined and analysed using suitable methods, including Mantel–Haenszel for dichotomous outcomes at pre-specified time points and multilevel modelling techniques for continuous repeated measurements. The latter method maximises power and allows us to estimate overall treatment effects over time. Trial of origin will be included as a fixed or random effect as deemed appropriate.

Owing to different scales of measurement in individual studies, it is anticipated that the standardised mean difference (SMD) will be used for continuous data. It may also be necessary to convert data on different scales using an appropriate transformation, for example the standard correction factor of $\Pi/3$ to convert from SMD to log odds ratio.

Initially, analyses will be performed using the direct comparisons only (hysterectomy versus ablation, ablation versus ablation and LNG IUS versus ablation). However, it is anticipated that there may be a limited number of direct comparisons available. In this case, a method of adjusted indirect comparison will be used to estimate comparative efficacy. In simple terms, this approach enables a comparison of interventions A and B if both have been compared with C. This will allow us to explore the ranking of treatment effectiveness.

Subgroup analysis

Subgroup analyses, if not carefully planned, can lead to misleading results, for example owing to the play of chance with multiple testing. Extreme caution will be used in interpretation of subgroup results. Any subgroup analysis will be limited to the following parameters:

1. intervention
2. ± pathology
3. age < 35, 35–45 and > 45 years
4. uterine cavity length < 8 cm, 8–10 cm and > 10 cm
5. presence or absence of submucous fibroids > 2 cm
6. previous ablation/treatment
7. nulliparous
8. mode of delivery (i.e. caesarean section).

HMB IPD meta-analysis Collaborative Group organisation

Management of the Collaborative Group

The Birmingham Clinical Trials Unit will act as the group secretariat for the IPD meta-analysis and will hold the main database. All data will be held securely and treated with the strictest of confidence. The overview will be managed by a small group including grant holders and research staff employed on the project grant listed below:

Siladitya Bhattacharya	Lead investigator, overall responsibility for Overview Group
Kevin Cooper	Clinical lead, BSGE (British Society for Gynaecological Endoscopy) representative, contact with authors
Khalid S Khan	Clinical lead, methodology
Richard Gray	Methodology and analysis
Jane Daniels	Project management
Lee Middleton	Overview statistician
Rita Champaneria	Overview systematic reviewer

Memorandum of Understanding for the collaborative group

The activities of the IPD meta-analysis will be governed by an initial Memorandum of Understanding, to be agreed by all collaborators within this group including primary triallists and secondary researchers, at the start of the project. The Memorandum of Understanding will set out the aims, scope, responsibilities and tasks required of all investigators.

Relationships with the other components of the guidelines development group

The IPD meta-analysis is a component of a larger project aiming to generate evidence-based, cost-effective clinical guidelines. The results of the IPD meta-analysis will be incorporated into a decision analytic model, which will then inform the development of guidelines. The International HMB IPD Meta-analysis Collaborative Group will not be directly involved in these processes, other than lead investigators from the Management Group.

Outputs

Outputs from this project will be:

- IPD meta-analysis of direct comparisons of interventions
- indirect comparison of rankings of different types of ablations
- input for the health economics model
- development of methodological methods for IPD meta-analyses
- identification of the need for more primary research (in areas where clinical uncertainties remain).

Publication policy

The results from the IPD meta-analysis will be presented at a collaborators meeting. Any subsequent articles on the results of the meta-analysis will be published under the name of the collaborative group – the International HMB IPD Meta-analysis Collaborative Group. It will also be circulated to the collaborators for comment, amendments and approval before finally being submitted. In the case of any disagreement, the following fundamental principle will be applied: that the report should provide the meta-analysis results, presenting all of the available evidence, but will not include any interpretations of the data, except those that are unanimously decided upon by all collaborators. Any collaborating group is free to withdraw its data at any stage.

Cost-effectiveness analysis

Aim

To undertake a cost-effectiveness analysis of hysterectomy versus second-generation ablative techniques and alternative forms of second-generation ablation using information generated from the above analyses.

This project will involve the development of a decision analytic simulation model as a framework for conducting cost-effectiveness and cost–utility analyses and associated value of information analyses (Felli and Hazen, 1998; Claxton *et al.*, 2001). The economic evaluations will inform current treatment policy in this clinical area, while the value of the information component will serve to highlight future research needs and agendas, and inform possible future research funding decisions. A modelling framework is ideally suited to demonstrate and explore the importance of the inherent uncertainty.

The model development process will use, as a starting point, the recently published menorrhagia clinical pathway Markov model (Garside *et al.*, 2004). This model, generated by researchers at the University of Exeter, formed the basis of the national coverage decision by NICE on microwave and thermal balloon EA for menorrhagia. Any requirements for structural model adjustments will be determined through:

- consideration of other recent HMB models (such as the model developed as part of the NICE HMB guideline prepared by the National Collaborating Centre for Women's and Children's Health – draft out for consultation currently)
- consultation within the research team, drawing on the requisite clinical and modelling expertise; and with appropriate external advisers (such as those involved in the modelling work reported in Garside *et al.*, 2004).

The principal clinical data to be used in populating the model will be drawn from other aspects of our research work, namely the individual patient meta-analyses and data from both national registers and follow-up of existing RCTs (as detailed earlier in this proposal). Assuming that a Markov model is found to be appropriate, it will be constructed using TREEAGE PRO (TreeAge Software Inc., Williamstown, MA, USA) software. This is a widely used and highly user-friendly software package ideally suited to the construction and analysis of decision tree and Markov models.

The economic evaluation will adopt a broad perspective and seek to include consideration of costs incurred by the health sector, by patients and by the economy more broadly in terms of productivity issues. An incremental approach will be adopted with a focus on additional costs and gain in benefits associated with a move away from current practice to alternative treatment strategies. The cost-effectiveness component of the work will report results in terms of an ICER of cost per woman successfully treated and cost per hysterectomy avoided. However, QoL data suitable for use in a cost–utility framework are available from published sources (for example, Sculpher, 1998) and so the economic evaluation will additionally present results in terms of incremental cost per QALY gained. Resource use will be estimated from the existing published evidence and additional cost data will be sought from other sources such as the annual review of unit health and social care costs (by the University of Kent) and national schedule for reference costs.

The results of the cost–utility analysis (CUA) will be presented using cost-effectiveness acceptability curves to reflect sampling variation and uncertainties in the appropriate threshold cost-effectiveness value. We shall also include a value of information analysis to quantify the total uncertainty in terms of the value of removing that uncertainty. As appropriate, we shall include partial value of information analysis calculations. In addition to this probabilistic sensitivity analysis on our base-case model, we shall include a range of alternative analyses to explore the robustness of these results to plausible variations in key assumptions and variations in the analytical methods used, and to consider the broader issue of the generalisability of the results.

To develop an algorithm for clinical decision making in women with heavy menstrual bleeding

Aim

To devise a parsimonious algorithm for clinical decision making, regarding the choice of surgery for women with HMB with failed medical treatment.

The call for proposals asks for patient perspectives to be taken into account. For many patients, the choice is likely to be straightforward if there is absolute certainty about comparative outcomes. Where such certainty is lacking, the ultimate decision may be influenced by personal preference. In this proposal we have planned to produce clinical algorithms which will guide practice, without overriding a clear preference a particular patient may have. We accept that, for an algorithm to be useful in a pragmatic context, it should be flexible enough to accommodate consumer preference. We therefore plan to develop algorithms for a typical (default) situation in a way which is highly sensitive to the needs and preferences of individual patients.

We will use consensus development processes to produce an interim or indicative algorithm. A hybrid method (modified Delphi technique) incorporating a postal questionnaire for the first round of ratings followed by a meeting where the second round of ratings occurs is the preferred technique for this project.

Delphi participants will include a panel (of about 15–20 respondents) selected from the following groups of stakeholders: general practitioners, general gynaecologists, gynaecologists with a special interest in minimal-access surgery (members of the British Gynaecological Endoscopy Society) and representatives from the Royal College of Obstetrics and Gynaecology. A questionnaire will be developed for the consensus process, based on the results from clinical effectiveness and cost-effectiveness data. Participants will initially complete the questionnaire by post/email. Potential loss to follow-up will be minimised by postal/email and telephone repeat reminders.

A subset of individuals will subsequently attend a facilitated face-to-face meeting (unless a consensus emerges from responses to postal questionnaires). At this meeting, each participant will receive a new copy of the questionnaire with a reminder of their own initial ratings and the distribution of ratings for the group as a whole. Each item will be discussed in turn and reasons for any differences explored, after which participants will privately re-rate the questions. Participants at the face-to-face group meeting will also include two patient representatives.

Amendment

Following consultation with HTA this part of the protocol (to develop an algorithm for clinical decision making in women with HMB) was amended. The amendment was approved by the NETSCC consultant advisor on 5 February 2010

The last research question, i.e. development of a clinical algorithm, was felt to be important at the time of submission of the application in 2006. Its current relevance is questionable as NICE has already issued a guideline on HMB which incorporates an algorithm for investigation and treatment. This guideline is due to be revised soon and is expected to take into account the results from this HTA project. In view of the presence of the existing NICE guideline as well as the imminent deadline for submission of the final report (14 February), we felt that a modification of the original protocol was appropriate.

A questionnaire survey of 18 stakeholders (15 clinical experts) was undertaken in January 2010. Responses from 15 experts have shown remarkable consensus in terms of decision making in HMB of unknown origin. Nine out of 10 responders indicated that, on the basis of the effectiveness and cost data generated by this project, they would favour Mirena, followed by second-generation ablation techniques, followed by hysterectomy as first-, second- and third-line approaches to HMB. Under these circumstances the value of a formal consensus process, involving a face-to-face meeting of experts who seem to be in general agreement, seems limited. Instead, based on the responses received, and input from a panel of consumer representatives, we intend to provide a simple clinical algorithm.

References

Abbott J, Hawe J, Hunter D, Garry R. A double-blind randomized trial comparing the Cavaterm and the NovaSure endometrial ablation systems for the treatment of dysfunctional uterine bleeding. *Fertil Steril* 2003;**80**:203–8.

Aberdeen Endometrial Ablation Trials Group. A randomised trial of endometrial ablation versus hysterectomy for the treatment of dysfunctional uterine bleeding: outcome at four years. *Br J Obstet Gynaecol* 1999;**106**:360–6.

Arieff AI, Ayus JC. Endometrial ablation complicated by fatal hyponatremic encephalopathy. *JAMA* 1993;**270**:1230–2.

Bain C, Cooper KG, Parkin DE. Microwave endometrial ablation versus endometrial resection: a randomized controlled trial. *Obstet Gynecol* 2002;**99**:983–7.

Barrington JW, Arunkalaivanan AS, Abdel-Fattah M. Comparison between the levonorgestrel intrauterine system (LNG-IUS) and thermal balloon ablation in the treatment of menorrhagia. *Eur J Obstet Gynecol Reprod Biol* 2003;**108**:72–4.

Bhattacharya S, Cameron IM, Parkin DE, Abramovich DR, Mollison J, Pinion SB, *et al.* A pragmatic randomised comparison of transcervical resection of the endometrium with endometrial laser ablation for the treatment of menorrhagia. *Br J Obstet Gynaecol* 1997;**104**:601–7.

Bongers MY, Bourdrez P, Mol BWJ, Heintz APM, Brolmann HAM. Randomised controlled trial of bipolar radio-frequency endometrial ablation and balloon endometrial ablation. *BJOG* 2004;**111**:1095–102.

Boujida VH, Philipsen T, Pelle J, Joergensen JC. Five-year follow-up of endometrial ablation: endometrial coagulation versus endometrial resection. *Obstet Gynecol* 2002;**99**:988–92.

Brown PM, Farquhar CM, Lethaby A, Sadler LC, Johnson NP. Cost-effectiveness analysis of levonorgestrel intrauterine system and thermal balloon ablation for heavy menstrual bleeding. *BJOG* 2006;**113**:797–803.

Brun JL, Burlet G, Galand B, Quereux C, Bernard P. Cavaterm thermal balloon endometrial ablation versus hysteroscopic endometrial resection to treat menorrhagia: the French, multicenter, randomized study. *J Minim Invasive Gynecol* 2006;**13**:424–30.

Busfield RA, Farquhar CM, Sowter MC, Lethaby A, Sprecher M, Yu Y, *et al.* A randomised trial comparing the levonorgestrel intrauterine system and thermal balloon ablation for heavy menstrual bleeding. *BJOG* 2006;**113**:257–63.

Cameron IM, Mollison J, Pinion SB, Atherton-Naji A, Buckingham K, Torgerson D. A cost comparison of hysterectomy and hysteroscopic surgery for the treatment of menorrhagia. *Eur J Obstet Gynecol Reprod Biol* 1996;**70**:87–92.

Claxton K, Neumann PJ, Araki S, Weinstein MC. Bayesian value-of-information analysis. *Int J Technol Assess Health Care* 2001;**17**:38–55.

Cooper J, Gimpelson R, Laberge P, Galen D, Garza-Leal JG, Scott J, *et al.* A randomized, multicenter trial of safety and efficacy of the NovaSure system in the treatment of menorrhagia. *J Am Assoc Gynecol Laparosc* 2002;**9**:418–28.

Cooper JM, Anderson TL, Fortin CA, Jack SA, Plentl MB. Microwave endometrial ablation vs rollerball electroablation for menorrhagia: a multicenter randomized trial. *J Am Assoc Gynecol Laparosc* 2004;**11**:394–403.

Cooper KG, Bain C, Parkin DE. Comparison of microwave endometrial ablation and transcervical resection of the endometrium for treatment of heavy menstrual loss: a randomised trial. *Lancet* 1999;**354**:1859–63.

Cooper KG, Jack SA, Parkin DE, Grant AM. Five-year follow up of women randomised to medical management or transcervical resection of the endometrium as treatment for heavy menses. *BJOG* 2001;**108**:1222–8.

Cooper KG, Bain C, Lawrie L, Parkin DE. A randomised comparison of microwave endometrial ablation with transcervical resection of the endometrium; follow up at a minimum of five years. *BJOG* 2005;**112**:470–5.

Corson SL. A multicenter evaluation of endometrial ablation by hydrothermablator and rollerball for treatment of menorrhagia. *J Am Assoc Gynecol Laparosc* 2001;**8**:359–67.

Corson SL, Brill AI, Brooks PG, Cooper JM, Indman PD, Liu JH, *et al.* One-year results of the Vesta system for endometrial ablation. *J Am Assoc Gynecol Laparosc* 2000;**7**:489–97

Crosignani PG, Vercellini P, Apolone G, De Giorgi O, Cortesi I, Meschia M. Endometrial resection versus vaginal hysterectomy for menorrhagia: long-term clinical and quality-of-life outcomes. *Am J Obstet Gynecol* 1997;**177**:95–101.

Davis JA. Hysteroscopic endometrial ablation with the neodymium-YAG laser. *Br J Obstet Gynaecol* 1989;**96**:928–32.

Dickersin K, Munro MG, Clark M, Langenberg P, Scherer R, Frick K, *et al.* Hysterectomy compared with endometrial ablation for dysfunctional uterine bleeding: a randomized controlled trial. *Obstet Gynecol* 2007;**110**:1279–89.

Dickersin K, Munro M, Langenberg P, Scherer R, Frick KD, Weber AM, *et al.* Surgical Treatments Outcomes Project for Dysfunctional Uterine Bleeding (STOP-DUB): design and methods. *Control Clin Trials* 2003;**24**:591–609.

Duffy S, Reid PC, Smith JH, Sharp F. In vitro studies of uterine electrosurgery. *Obstet Gynecol* 1991;**78**:213–20.

Duffy S, Reid PC, Sharp F. In-vivo studies of uterine electrosurgery. *Br J Obstet Gynaecol* 1992;**99**:579–82.

Duleba AJ, Heppard MC, Soderstrom RM, Townsend DE. A randomized study comparing endometrial cryoablation and rollerball electroablation for treatment of dysfunctional uterine bleeding. *J Am Assoc Gynecol Laparosc* 2003;**10**:17–26.

Dwyer N, Hutton J, Stirrat GM.. Randomised controlled trial comparing endometrial resection with abdominal hysterectomy for the surgical treatment of menorrhagia. *Br J Obstet Gynaecol* 1993;**100**:237–43.

Felli JC, Hazen GB. Sensitivity analysis and the expected value of perfect information. *Med Decis Making* 1998;**18**:95–109. *Also correction Med Decis Making* 2003;**23**:97.

Gannon M, Holt EM, Fairbank J. A randomised controlled trial comparing endometrial resection and abdominal hysterectomy for the treatment of menorrhagia. *BMJ* 1991;**303**:1362–4.

Garside R, Stein K, Wyatt K, Round A, Pitt M. A cost–utility analysis of microwave and thermal balloon endometrial ablation techniques for the treatment of heavy menstrual bleeding. *BJOG* 2004;**111**:1103–14. Review.

Goldrath MH, Fuller TASS. Laser photovaporisation of endometrium for the treatment of menorrhagia. *Am J Obstet Gynecol* 1981;**140**:14–19.

Halmesmaki K, Hurskainen R, Teperi J, Grenman S, Kivela A, Kujansuu E, *et al.* The effect of hysterectomy or levonorgestrel-releasing intrauterine system on sexual functioning among women with menorrhagia: a 5-year randomised controlled trial. *BJOG* 2007;**114**:563–8.

Hawe J, Abbott J, Hunter D, Phillips G, Garry R. A randomised controlled trial comparing the Cavaterm endometrial ablation system with the Nd:YAG laser for the treatment of dysfunctional uterine bleeding. *BJOG* 2003;**110**:350–7.

Hurskainen R, Teperi J, Rissanen P, Aalto AM, Grenman S, Kivela A, *et al.* Quality of life and cost-effectiveness of levonorgestrel-releasing intrauterine system versus hysterectomy for treatment of menorrhagia: a randomised trial. *Lancet* 2001;**357**:273–7.

Kittelsen N, Istre O. A randomized study comparing levonorgestrel intrauterine system (LNG IUS) and transcervical resection of the endometrium (TCRE) in the treatment of menorrhagia: preliminary results. *Gynaecol Endosc* 1998;**7**:61–5.

Kleijn JH, Engels R, Bourdrez P, Mol BW, Bongers MY. Five-year follow up of a randomised controlled trial comparing NovaSure and ThermaChoice endometrial ablation. *BJOG* 2008;**115**:193–8.

Lahteenmaki P, Haukkamaa M, Puolakka J, Riikonen U, Sainio S, Suvisaari J, *et al.* Open randomised study of use of levonorgestrel releasing intrauterine system as alternative to hysterectomy. *BMJ* 1998;**316**:1122–6.

Lethaby A, Hickey M. Endometrial destruction techniques for heavy menstrual bleeding. *Cochrane Database Syst Rev* 2005;**4**:CD001501.

Lethaby A, Shepperd S, Cooke I, Farquhar C. Endometrial resection and ablation versus hysterectomy for heavy menstrual bleeding. *Cochrane Database Syst Rev* 1999;**2**:CD000329.

Lin H. Comparison between microwave endometrial ablation and total hysterectomy. *Chin Med J* 2006;**119**:1195–7.

Loffer FD. Three-year comparison of thermal balloon and rollerball ablation in treatment of menorrhagia. *J Am Assoc Gynecol Laparosc* 2001;**8**:48–54.

Loffer FD, Grainger D. Five-year follow-up of patients participating in a randomized trial of uterine balloon therapy versus rollerball ablation for treatment of menorrhagia. *J Am Assoc Gynecol Laparosc* 2002;**9**:429–35.

Magos A, Baumann R, Turnbull AC. Transcervical resection of the endometrium in women with menorrhagia. *BMJ* 1989;**298**:1209–12.

Malak KA. Management of menorrhagia with the levonorgestrel intrauterine system versus endometrial resection. *Gynaecol Surg* 2006;**3**:275–80.

McClure N, Marners M, Healy DL, Hill DJ, Lawrence AS, Wingfield M, *et al.* A quantitative assessment of endometrial electrocautery in the management of menorrhagia and a comparative report of argon laser endometrial ablation. *Gynaecol Endoscopy* 1992;**1**:199–202.

Meyer WR, Walsh BW, Grainger DA, Peacock LM, Loffer FD, Steege JF. Thermal balloon and rollerball ablation to treat menorrhagia: a multicenter comparison. *Obstet Gynecol* 1998;**92**:98–103.

Milsom I, Andersson K, Andersch B, Rybo G. A comparison of flurbiprofen, tranexamic acid and a levonorgestrel releasing intrauterine contraceptive device in the treatment of idiopathic menorrhagia. *Am J Obstet Gynecol* 1991;**164**:879–83.

National Institute for Clinical Excellence. *Fluid-filled thermal balloon and microwave endometrial ablation for heavy menstrual bleeding.* Technology Appraisal Guidance No. 78; 2004.

O'Connor H, Broadbent Magos AL, McPherson K. Medical Research Council randomised trial of endometrial resection versus hysterectomy in the management of menorrhagia. *Lancet* 1997;**349**:891–901.

Overton C, Hargreaves J, Maresh M. A national survey of the complications of endometrial destruction for menstrual disorders: the MISTLETOE study. Minimally Invasive Surgical Techniques – Laser, EndoThermal or Endorescetion. *Br J Obstet Gynaecol* 1997;**104**:1351–9.

Pellicano M, Guida M, Acunzo G, Cirillo D, Bifulco G, Nappi C. Hysteroscopic transcervical endometrial resection versus thermal destruction for menorrhagia: a prospective randomized trial on satisfaction rate. *Am J Obstet Gynecol* 2002;**187**:545–50.

Perino A, Castelli A, Cucinella G, Biondo A, Pane A, Venezia R. A randomized comparison of endometrial laser intrauterine thermotherapy and hysteroscopic endometrial resection. *Fertil Steril* 2004;**82**:731–4.

Pinion SB, Parkin DE, Abramovich DR, Naji A, Alexander DA, Russell IT, *et al.* Randomised trial of hysterectomy, endometrial laser ablation, and transcervical endometrial resection for dysfunctional uterine bleeding. *BMJ* 1994;**309**:979–83.

Rees MCP. The role of menstrual blood loss measurements in management of complaints of excessive menstrual bleeding. *BMJ* 1991;**98**:327–8.

Romer T. [The treatment of recurrent menorrhagias – Cavaterm-balloon-coagulation versus Rollerball-endometrial ablation – a prospective randomized comparative study]. *Zentralblatt fur Gynakologie* 1998;**120**:511–14.

Rosenberg MK. Hyponatremic encephalopathy after rollerball endometrial ablation. *Anesth Analg* 1995;**80**:1046–8.

Royal College of Obstetricians and Gynaecologists. *The initial management of menorrhagia (Evidence-based Guidelines No. 1)*. London; 1998.

Sambrook A. A randomised trial of microwave endometrial ablation versus thermal balloon ablation. (Unpublished.)

Sambrook AM, Bain C, Parkin DE, Cooper KG. A randomised comparison of microwave endometrial ablation with transcervical resection of the endometrium: follow up at a minimum of 10 years. *BJOG* 2009;**116**:1033–7.

Scottish Hysteroscopy Audit Group. A Scottish audit of hysteroscopic surgery for menorrhagia: complications and follow up. *Br J Obstet Gynaecol* 1995;**102**:249–54.

Sculpher M. A cost-utility analysis of abdominal hysterectomy versus transcervical endometrial resection for the surgical treatment of menorrhagia. *Int J Technol Assess Health Care* 1998;**14**:302–19.

Shaw RW, Symonds IM, Tamizian O, Chaplain J, Mukhopadhyay S. Randomised comparative trial of thermal balloon ablation and levonorgestrel intrauterine system in patients with idiopathic menorrhagia. *Aust N Z J Obstet Gynaecol* 2007;**47**:335–40.

Soysal M, Soysal S, Ozer S. A randomized controlled trial of levonorgestrel releasing IUD and thermal balloon ablation in the treatment of menorrhagia. *Zentralbl Gynakol* 2002;**124**:213–19.

Soysal ME, Soysal SK, Vicdan K. Thermal balloon ablation in myoma-induced menorrhagia under local anesthesia. *Gynecol Obstet Invest* 2001;**51**:128–33.

Tam WH, Yuen PM, Shan Ng DP, Leung PL, Lok IH, Rogers MS. Health status function after treatment with thermal balloon endometrial ablation and levonorgestrel intrauterine system for idiopathic menorrhagia: a randomized study. *Gynecol Obstet Invest* 2006;**62**:84–8.

Vessey MP, Villard-Mackintosh L, McPherson K, Coulter A, Yeates D. The epidemiology of hysterectomy: findings in a large cohort study. *Br J Obstet Gynaecol* 1992;**99**:402–7.

Vercellini P, Oldani S, Yaylayan L, Zaina B, De Giorgi O, Crosignani PG. Randomized comparison of vaporizing electrode and cutting loop for endometrial ablation. *Obstet Gynecol* 1999;**94**:521-7.

Zupi E, Zullo F, Marconi D, Sbracia M, Pellicano M, Solima E, *et al.* Hysteroscopic endometrial resection versus laparoscopic supracervical hysterectomy for menorrhagia: a prospective randomized trial. *Obstet Gynecol* 2003;**188**:7–12.

Appendix A

TABLE 1a Characteristics of available trials[a] (hysterectomy vs ablation)

Study reference, number randomised	Country	Eligibility criteria	Randomised comparison	Outcome measures	Measure of outcome	Response
Crosignanani, 1997 *n* = 92	Italy	Women < 50 years Failed medical treatment Uterine size < 12 weeks Submucous fibroid < 3 cm	Vaginal hysterectomy vs TCRE	Satisfaction MBL QoL Duration of surgery Hospital stay Return to work Retreatment (further surgery)	 Minutes Days Weeks 	Not as yet, but trying to contact via Vercellini group
Dickersin, 2007 *n* = 242	USA	Premenopausal women with DUB aged ≥ 18 years	Hysterectomy vs ablation	Menstrual status QoL	 EuroQoL (EQ-5D)	Yes, willing to collaborate

continued

TABLE 1a Characteristics of available trials[a] (hysterectomy vs ablation) *(continued)*

Study reference, number randomised	Country	Eligibility criteria	Randomised comparison	Outcome measures	Measure of outcome	Response
Dwyer, 1993 *n* = 200	Weston-Super-Mare, UK	Age < 52 years Failed medical treatment Uterus < 12 weeks	Abdominal hysterectomy vs TCRE	Patient satisfaction (4 months and 2.8 years) MBL (subjective) QoL at 2.8 years Hospital stay Return to work Retreatment (further surgery) Total resource use at 2.8 years	 Days Weeks £ 	Not as yet
Gannon, 1991 *n* = 54	Ireland, UK	Women median age 40 years Failed medical treatment Uterine size < 12 weeks Submucous fibroid < 3 cm Endometrial prep	Abdominal hysterectomy vs TCRE	MBL Duration of surgery Hospital stay Return to work Retreatment (further surgery) Resource use for surgery	 Minutes Days Days £	Yes, willing to collaborate
O'Connor, 1997 *n* = 202	London, UK	Women age 30–50 years Failed medical treatment Uterine size < 12 weeks Submucous fibroid < 5 cm	Abdominal hysterectomy (28) + vaginal hysterectomy (28) vs TCRE	Patient satisfaction (2 years) MBL QoL at 2 years Hospital stay Retreatment (further surgery)	 Days 	Yes, NOT willing to collaborate
Pinion, 1994 *n* = 204	Dundee, UK	Women age < 50 years Failed medical treatment Uterine size < 10 weeks	Abdominal hysterectomy vs TCRE + ELA	Patient satisfaction (1 and 4 years) MBL QoL Hospital stay Return to work Retreatment (further surgery) Health service and patient costs	 VAS Number of nights in hospital Weeks/ months Weeks/ months £	Yes, willing to collaborate
Zupi, 2003 *n* = 181	Italy	Women age < 50 years Failed medical treatment Weight < 100 kg	TCRE vs hysterectomy	Patient satisfaction MBL	 	

a In addition to the above trials we have identified a further abstract of a study published in the *Chinese Medical Journal* (Lin, 2006). We have requested the full paper and need to verify whether this was a randomised trial and therefore suitable for inclusion.

TABLE 1b Characteristics of available trials (ablation versus ablation)

Study reference, number randomised	Country	Eligibility criteria	Randomised comparison	Outcome measures	Measure of outcome	Response
Trials comparing first-generation ablative techniques						
Bhattacharya, 1997 *n* = 372	Aberdeen, UK	Age < 50 years Mean age 41 years Uterine size < 10 weeks Clinical diagnosis of DUB Normal histology	TCRE + rollerball vs laser	Satisfaction at 1 year Amenorrhoea Duration of surgery Complications Retreatment	 Minutes 	Yes, willing to collaborate
Boujida, 2002 *n* = 120	Denmark	Age > 35 years	TCRE vs rollerball endometrial coagulation	Hysterectomy rate 5 years later Days with bleeding Recommend treatment	 Days 	Not as yet, but still trying to make contact
McClure, 1992 *n* = 38	Ireland	Mean age 42 years Menorrhagia unresponsive to medical treatment MBL > 70 ml	TCRE + rollerball vs laser (argon)	MBL reduction Amenorrhoea Duration of surgery Complications	MBL (> 70 ml) Minutes 	Yes, willing to collaborate
Trials comparing first- with second-generation ablative techniques						
Brun, 2006 *n* = 51	France	Higham blood loss score > 100	TCRE vs Cavaterm TBEA	Amenorrhoea Higham bleeding score	 Higham bleeding score	Yes, willing to collaborate
Cooper, 1999 *n* = 263	Aberdeen, UK	Mean age 41 years Uterine size < 10 weeks Clinical diagnosis of DUB Normal histology	TCRE + rollerball vs MEA	PBAC Satisfaction at 1 year QoL (SF-36) Amenorrhoea Duration of surgery Postoperative stay Return to work Complications Retreatment	PBAC SF-36 Minutes Hours Days 	Yes, willing to collaborate
Cooper, 2002 *n* = 265	USA	Age 25–50 years Menorrhagia (PBAC > 150) Failed medical treatment	NovaSure vs wire loop resection + rollerball	PBAC Duration of surgery Sedation Complications	PBAC Minutes 	Deceased, but industry willing to collaborate
Cooper, 2004 *n* = 322	USA	Mean age 41 years Age > 30 years Failed/refused medical treatment PBAC > 185 Uterine cavity 6–14 cm	Microwave vs rollerball	PBAC > 75 Satisfaction QoL (SF-36) Amenorrhoea Duration of surgery Sedation Complications	PBAC SF-36 Minutes 	Deceased, but industry willing to collaborate
Corson, 2000 *n* = 276	USA	PBAC > 150 Distorted uterine cavity Cavity length > 9.75 cm	Vesta balloon vs TCRE + rollerball	PBAC: proportion > 76 Amenorrhoea Adverse events	PBAC 	Not as yet

continued

TABLE 1b Characteristics of available trials (ablation versus ablation) *(continued)*

Study reference, number randomised	Country	Eligibility criteria	Randomised comparison	Outcome measures	Measure of outcome	Response
Corson, 2001 *n* = 276	USA	Age 30–50 years Myomas < 4 cm	Rollerball vs HTA (hydrotherm ablator)	PBAC Menstrual diary Amenorrhoea Proportion with PBAC < 75 QoL Retreatment	PBAC PBAC PBAC SF-36 	Not as yet
Duleba, 2003 *n* = 279	USA	Age 30–50 years PBAC > 150 Uterine cavity > 10 cm Intramural myomas < 2 cm	Rollerball vs endometrial cryoablation	PBAC Menstrual diary Bleeding and pain Satisfaction	PBAC PBAC PBAC 	Not as yet
Hawe, 2003 *n* = 72	UK	Age 29–51 years Uterine length < 12 cm	Cavaterm TBEA vs Nd:Yag laser	Amenorrhoea QoL (SF-12) Satisfaction VAS pain Operative details + complications	 SF-12 VAS 	Yes, willing to collaborate
Meyer, 1998 *n* = 272	USA	Age 29–50 years PBAC score > 150 Ineffective medical therapy Uterine cavity size 4–10 cm	Rollerball vs TBEA (ThermaChoice)	Satisfaction PBAC Complications Duration of surgery Retreatment rate	 PBAC Minutes 	Yes, willing to collaborate
Pellicano, 2002 *n* = 82		Mean age 43 years Age < 50 years Weight < 100 kg Uterine size < 12 weeks	TCRE vs Cavaterm TBEA	Satisfaction Complications Duration of surgery Retreatment rate	 Minutes 	Not as yet
Perino, 2004 *n* = 116	Italy	Age 36–48 years DUB	TCRE vs ELITT (endometrial laser intrauterine thermal therapy)	Amenorrhoea Complications Duration of surgery Retreatment rate	VAS Minutes 	Yes, willing to collaborate
Romer, 1998 *n* = 20	Germany	Age 35–52 years	Rollerball vs Cavaterm TBEA	Satisfaction Amenorrhoea	 VAS	Not as yet
Soysal, 2001 *n* = 96	Turkey	Age 40–49 years	Rollerball vs TBEA	Satisfaction Amenorrhoea Complications Duration of surgery	 PBAC 	Not as yet
van Zon-Rabelonk, 2003 *n* = 139	Netherlands	Age unreported	Rollerball vs TBEA	Technical safety Reduction in menstrual bleeding	 	Yes, willing to collaborate
Vercellini, 1999 *n* = 46	Italy	Age > 35 years Unterine size < 12 weeks Normal cavity	TCRE vs vaporising electrode	Satisfaction Amenorrhoea Complications Duration of surgery PBAC	 PBAC Minutes PBAC	Not as yet

TABLE 1b Characteristics of available trials (ablation versus ablation) *(continued)*

Study reference, number randomised	Country	Eligibility criteria	Randomised comparison	Outcome measures	Measure of outcome	Response
Trials comparing second-generation ablative techniques						
Abbott, 2003 *n* = 57	Australia	Mean ages + 40.5 years (Novasure) and 40.5 years (Cavaterm) DUB Uterine length < 12 cm	Novasure vs Cavaterm TBEA	Amenorrhoea QoL Satisfaction acceptability	EuroQoL (EQ-5D)	Yes, willing to collaborateVAS
Bongers, 2004 *n* = 126 5-year report published, Kleijn *et al.*, 2008	Netherlands	Mean age 43 years PBAC > 150 Uterine length 6–12 cm	Novasure vs ThermaChoice TBEA	Amenorrhoea Satisfaction Duration of surgery Retreatment	PBAC Minutes	Yes, willing to collaborate
Clark, 2007	Birmingham, UK	Unpublished	NovaSure vs ThermaChoice			Yes, willing to collaborate
Sambrook, 2009 *n* = 240	Aberdeen, UK		ThermaChoice TBEA vs MEA	QoL Satisfaction PBAC	PBAC	Yes, willing to collaborate

TABLE 1c Characteristics of available trials (Mirena versus ablation)

Study reference, number randomised	Country	Eligibility criteria	Randomised comparison	Outcome measures	Measure of outcome	Response
Barrington, 2003 *n* = 44	Devon, UK	Menorrhagia refractory to medical treatment Uterine length < 12 cm	LNG IUS Mirena vs TBEA	PBAC score Improvement in bleeding Need for further treatment	PBAC	Yes, NOT willing to collaborate
Busfield, 2006 *n* = 79 Cost-effectiveness study carried out by Brown *et al.*, 2006	New Zealand	Heavy menstrual bleeding Age 25–50 years Regular cycle	LNG-IUS vs TBEA	Menstrual blood loss Patient satisfaction QoL Menstrual symptoms Treatment side effects	PBAC SF-36	Yes, willing to collaborate
Crosignani, 1997 *n* = 70	Italy	Age 38–53 years MBL > 80 ml/cycle Uterine size < 8 weeks	TCRE	PBAC Patient satisfaction SF-36 Amenorrhoea at 12 months	SF-36	Contact again via Vercellini group
Kittelsen, 1998 *n* = 53	Norway	Age 30–49 years PBAC > 100 Regular uterine cavity	LNG IUS Mirena vs TCRE	PBAC	PBAC	Not as yet
Malak, 2006 *n* = 56	Egypt	Age 40–50 years Cavity < 10 cm	LNG-IUS vs TCRE	Amenorrhoea PBAC score		Not as yet

continued

TABLE 1c Characteristics of available trials (Mirena versus ablation) *(continued)*

Study reference, number randomised	Country	Eligibility criteria	Randomised comparison	Outcome measures	Measure of outcome	Response
Shaw, 2007 *n*=66	England	Age 25–49 years Failed medical treatment Normal biopsy PBAC <120	TBEA vs LNG-IUS	PBAC score at 12 months	PBAC	Not as yet
Soysal, 2002 *n*=72	Turkey	Mean age 44 years	LNG IUS vs TBEA	Reduction in menstrual bleeding QoL		Not as yet
TALIS 2003		Age 25–50 years	LNG IUS vs TBEA	PBAC Satisfaction	PBAC	Not as yet
Tam, 2006 *n*=33	China	Premenopausal women >40 years Uterine cavity <10 cm	LNG IUS vs TBEA	SF-36	SF-36	Yes, willing to collaborate

TABLE 1d Characteristics of available trials (Mirena versus hysterectomy)

Study reference, number randomised	Country	Eligibility criteria	Randomised comparison	Outcome measures	Measure of outcome	Response
Hurskainen, 2001 *n*=236 5-year report published, Halmesmaki *et al.*, 2007	Finland	Menorrhagia Age 35–49 years	LNG IUS Mirena vs hysterectomy	EQ-5D Rand 36 MBL		Not as yet

Appendix B Search strategy for population

#1 Menorrhagia/all subheadings
#2 Hypermenorrhea/all subheadings
#3 Excessive NEAR ('menstrual bleeding' OR 'menstrual blood loss')
#4 Dysfunctional NEAR ('uterine bleeding' OR 'menstrual bleeding')
#5 Heavy NEAR ('menstrual bleeding' OR 'menstrual blood loss')
#6 'Iron deficient anaemia'
#7 (#3 OR #4 OR #5 OR #6) in TI, AB
#8 #1 OR #2 OR #7

Search strategy for interventions

Hysterectomy

#1 EXPLODE 'hysterectomy'/all subheadings
#2 'Vaginal hysterectomy'/all subheadings
#3 'Total abdominal hysterectomy'
#4 'Subtotal abdominal hysterectomy'
#5 'Laparoscopic hysterectomy'
#6 #1 OR #2 OR #3 OR #4 OR #5

Ablation

#1 EXPLODE 'hysteroscopy'/all subheadings

#2 ('Transcervical resection') NEAR 'endometrium'
#3 'TCRE'
#4 'Endometrial ablation'
#5 'Laser ablation'
#6 'Electrosurgery'
#7 'Rollerball'
#8 'Thermal balloon'
#9 'Hypertherm$'
#10 'Thermotherapy'
#11 'Photodynamic therapy'
#12 'Phototherapy'
#13 'Cryoablation'
#14 'Microwave ablation'
#15 'Radiofrequency'
#16 'Saline irrigation'
#17 'Laser interstitial'
#18 'Thermachoice'
#19 'Cavaterm'
#20 'ELITT'
#21 'Vesta'
#22 'Novasure'
#23 'Microsulis'
#24 'Cryogen'

Mirena

#1 EXPLODE 'contraceptive'/all subheadings
#2 'Mirena coil'/all subheadings
#3 'Levonorgestrel'
#4 'Intra uterine device'
#5 #1 OR #2 OR #3 OR #4

Search strategy for randomised controlled trials

#1 Randomized controlled trial IN PT
#2 Controlled clinical trial IN PT
#3 Randomized controlled trials IN SH
#4 Random allocation IN SH.
#5 Double blind method IN SH
#6 Single blind method IN SH
#7 (#1 OR #2 OR #3 OR #4 OR #5 OR #6)
#8 Animal in SH NOT human in SH
#9 #7 not # 8
#10 Clinical trial IN PT.
#11 EXPLODE clinical trials/all subheadings
#12 (clin NEAR trial) IN TI, AB
#13 [(Single OR double OR treble OR triple) NEAR (blind OR mask)] IN TI, AB
#14 Placebos IN SH
#15 Placebos IN TI, AB
#16 Random IN TI, AB
#17 Research Design IN SH
#18 #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17
#19 #18 NOT #8
#20 #19 NOT #9

#21 Comparative study IN SH
#22 EXPLORE evaluation studies/all subheadings
#23 Follow-up studies IN SH
#24 Prospective studies IN SH
#25 (Control OR prospective OR volunteer) IN TI, AB
#26 #21 OR #22 OR #23 OR #24 OR #25
#27 #26 NOT #8
#28 #27 NOT (#9 OR #20)
#29 #9 OR #20 OR #28

Version 4
5 February 2010

Health Technology Assessment programme

HTA Commissioning Board *(continued)*

Professor James Raftery,
Chair of NETSCC and Director of the Wessex Institute, University of Southampton

Professor Barney Reeves,
Professorial Research Fellow in Health Services Research, Department of Clinical Science, University of Bristol

Professor Martin Underwood,
Warwick Medical School, University of Warwick

Professor Marion Walker,
Professor in Stroke Rehabilitation, Associate Director UK Stroke Research Network, University of Nottingham

Dr Duncan Young,
Senior Clinical Lecturer and Consultant, Nuffield Department of Anaesthetics, University of Oxford

Observers

Dr Morven Roberts,
Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council

HTA Clinical Evaluation and Trials Board

Chair,
Professor Sallie Lamb,
Director,
Warwick Clinical Trials Unit,
Warwick Medical School,
University of Warwick and Professor of Rehabilitation,
Nuffield Department of Orthopaedic, Rheumatology and Musculoskeletal Sciences,
University of Oxford

Deputy Chair,
Professor Jenny Hewison,
Professor of the Psychology of Health Care,
Leeds Institute of Health Sciences,
University of Leeds

Programme Director,
Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Members

Professor Keith Abrams,
Professor of Medical Statistics, Department of Health Sciences, University of Leicester

Professor Martin Bland,
Professor of Health Statistics, Department of Health Sciences, University of York

Professor Jane Blazeby,
Professor of Surgery and Consultant Upper GI Surgeon, Department of Social Medicine, University of Bristol

Professor Julia M Brown,
Director, Clinical Trials Research Unit, University of Leeds

Professor Alistair Burns,
Professor of Old Age Psychiatry, Psychiatry Research Group, School of Community-Based Medicine, The University of Manchester & National Clinical Director for Dementia, Department of Health

Dr Jennifer Burr,
Director, Centre for Healthcare Randomised trials (CHART), University of Aberdeen

Professor Linda Davies,
Professor of Health Economics, Health Sciences Research Group, University of Manchester

Professor Simon Gilbody,
Prof of Psych Medicine and Health Services Research, Department of Health Sciences, University of York

Professor Steven Goodacre,
Professor and Consultant in Emergency Medicine, School of Health and Related Research, University of Sheffield

Professor Dyfrig Hughes,
Professor of Pharmacoeconomics, Centre for Economics and Policy in Health, Institute of Medical and Social Care Research, Bangor University

Professor Paul Jones,
Professor of Respiratory Medicine, Department of Cardiac and Vascular Science, St George's Hospital Medical School, University of London

Professor Khalid Khan,
Professor of Women's Health and Clinical Epidemiology, Barts and the London School of Medicine, Queen Mary, University of London

Professor Richard J McManus,
Professor of Primary Care Cardiovascular Research, Primary Care Clinical Sciences Building, University of Birmingham

Professor Helen Rodgers,
Professor of Stroke Care, Institute for Ageing and Health, Newcastle University

Professor Ken Stein,
Professor of Public Health, Peninsula Technology Assessment Group, Peninsula College of Medicine and Dentistry, Universities of Exeter and Plymouth

Professor Jonathan Sterne,
Professor of Medical Statistics and Epidemiology, Department of Social Medicine, University of Bristol

Mr Andy Vail,
Senior Lecturer, Health Sciences Research Group, University of Manchester

Professor Clare Wilkinson,
Professor of General Practice and Director of Research North Wales Clinical School, Department of Primary Care and Public Health, Cardiff University

Dr Ian B Wilkinson,
Senior Lecturer and Honorary Consultant, Clinical Pharmacology Unit, Department of Medicine, University of Cambridge

Observers

Ms Kate Law,
Director of Clinical Trials,
Cancer Research UK

Dr Morven Roberts,
Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council

Diagnostic Technologies and Screening Panel

Members

Chair,
Professor Lindsay Wilson Turnbull,
Scientific Director of the Centre for Magnetic Resonance Investigations and YCR Professor of Radiology, Hull Royal Infirmary

Professor Judith E Adams,
Consultant Radiologist, Manchester Royal Infirmary, Central Manchester & Manchester Children's University Hospitals NHS Trust, and Professor of Diagnostic Radiology, University of Manchester

Mr Angus S Arunkalaivanan,
Honorary Senior Lecturer, University of Birmingham and Consultant Urogynaecologist and Obstetrician, City Hospital, Birmingham

Dr Stephanie Dancer,
Consultant Microbiologist, Hairmyres Hospital, East Kilbride

Dr Diane Eccles,
Professor of Cancer Genetics, Wessex Clinical Genetics Service, Princess Anne Hospital

Dr Trevor Friedman,
Consultant Liason Psychiatrist, Brandon Unit, Leicester General Hospital

Dr Ron Gray,
Consultant, National Perinatal Epidemiology Unit, Institute of Health Sciences, University of Oxford

Professor Paul D Griffiths,
Professor of Radiology, Academic Unit of Radiology, University of Sheffield

Mr Martin Hooper,
Public contributor

Professor Anthony Robert Kendrick,
Associate Dean for Clinical Research and Professor of Primary Medical Care, University of Southampton

Dr Anne Mackie,
Director of Programmes, UK National Screening Committee, London

Mr David Mathew,
Public contributor

Dr Michael Millar,
Consultant Senior Lecturer in Microbiology, Department of Pathology & Microbiology, Barts and The London NHS Trust, Royal London Hospital

Mrs Una Rennard,
Public contributor

Dr Stuart Smellie,
Consultant in Clinical Pathology, Bishop Auckland General Hospital

Ms Jane Smith,
Consultant Ultrasound Practitioner, Leeds Teaching Hospital NHS Trust, Leeds

Dr Allison Streetly,
Programme Director, NHS Sickle Cell and Thalassaemia Screening Programme, King's College School of Medicine

Dr Alan J Williams,
Consultant Physician, General and Respiratory Medicine, The Royal Bournemouth Hospital

Observers

Dr Tim Elliott,
Team Leader, Cancer Screening, Department of Health

Dr Catherine Moody,
Programme Manager, Medical Research Council

Professor Julietta Patrick,
Director, NHS Cancer Screening Programme, Sheffield

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

Disease Prevention Panel

Members

Chair,
Professor Margaret Thorogood,
Professor of Epidemiology, University of Warwick Medical School, Coventry

Dr Robert Cook,
Clinical Programmes Director, Bazian Ltd, London

Dr Colin Greaves,
Senior Research Fellow, Peninsula Medical School (Primary Care)

Mr Michael Head,
Public contributor

Professor Cathy Jackson,
Professor of Primary Care Medicine, Bute Medical School, University of St Andrews

Dr Russell Jago,
Senior Lecturer in Exercise, Nutrition and Health, Centre for Sport, Exercise and Health, University of Bristol

Dr Julie Mytton,
Consultant in Child Public Health, NHS Bristol

Professor Irwin Nazareth,
Professor of Primary Care and Director, Department of Primary Care and Population Sciences, University College London

Dr Richard Richards,
Assistant Director of Public Health, Derbyshire Country Primary Care Trust

Professor Ian Roberts,
Professor of Epidemiology and Public Health, London School of Hygiene & Tropical Medicine

Dr Kenneth Robertson,
Consultant Paediatrician, Royal Hospital for Sick Children, Glasgow

Dr Catherine Swann,
Associate Director, Centre for Public Health Excellence, NICE

Professor Carol Tannahill,
Glasgow Centre for Population Health

Mrs Jean Thurston,
Public contributor

Professor David Weller,
Head, School of Clinical Science and Community Health, University of Edinburgh

Observers

Ms Christine McGuire,
Research & Development, Department of Health

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

External Devices and Physical Therapies Panel

Members

Chair,
Dr John Pounsford,
Consultant Physician North Bristol NHS Trust

Deputy Chair,
Professor E Andrea Nelson,
Reader in Wound Healing and Director of Research, University of Leeds

Professor Bipin Bhakta,
Charterhouse Professor in Rehabilitation Medicine, University of Leeds

Mrs Penny Calder,
Public contributor

Dr Dawn Carnes,
Senior Research Fellow, Barts and the London School of Medicine and Dentistry

Dr Emma Clark,
Clinician Scientist Fellow & Cons. Rheumatologist, University of Bristol

Mrs Anthea De Barton-Watson,
Public contributor

Professor Nadine Foster,
Professor of Musculoskeletal Health in Primary Care Arthritis Research, Keele University

Dr Shaheen Hamdy,
Clinical Senior Lecturer and Consultant Physician, University of Manchester

Professor Christine Norton,
Professor of Clinical Nursing Innovation, Bucks New University and Imperial College Healthcare NHS Trust

Dr Lorraine Pinnigton,
Associate Professor in Rehabilitation, University of Nottingham

Dr Kate Radford,
Senior Lecturer (Research), University of Central Lancashire

Mr Jim Reece,
Public contributor

Professor Maria Stokes,
Professor of Neuromusculoskeletal Rehabilitation, University of Southampton

Dr Pippa Tyrrell,
Senior Lecturer/Consultant, Salford Royal Foundation Hospitals' Trust and University of Manchester

Dr Sarah Tyson,
Senior Research Fellow & Associate Head of School, University of Salford

Dr Nefyn Williams,
Clinical Senior Lecturer, Cardiff University

Observers

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

Interventional Procedures Panel

Members

Chair,
Professor Jonathan Michaels,
Professor of Vascular Surgery, University of Sheffield

Deputy Chair,
Mr Michael Thomas,
Consultant Colorectal Surgeon, Bristol Royal Infirmary

Mrs Isabel Boyer,
Public contributor

Mr David P Britt,
Public contributor

Mr Sankaran Chandra Sekharan,
Consultant Surgeon, Breast Surgery, Colchester Hospital University NHS Foundation Trust

Professor Nicholas Clarke,
Consultant Orthopaedic Surgeon, Southampton University Hospitals NHS Trust

Ms Leonie Cooke,
Public contributor

Mr Seumas Eckford,
Consultant in Obstetrics & Gynaecology, North Devon District Hospital

Professor Sam Eljamel,
Consultant Neurosurgeon, Ninewells Hospital and Medical School, Dundee

Dr Adele Fielding,
Senior Lecturer and Honorary Consultant in Haematology, University College London Medical School

Dr Matthew Hatton,
Consultant in Clinical Oncology, Sheffield Teaching Hospital Foundation Trust

Dr John Holden,
General Practitioner, Garswood Surgery, Wigan

Professor Nicholas James,
Professor of Clinical Oncology, School of Cancer Sciences, University of Birmingham

Dr Fiona Lecky,
Senior Lecturer/Honorary Consultant in Emergency Medicine, University of Manchester/Salford Royal Hospitals NHS Foundation Trust

Dr Nadim Malik,
Consultant Cardiologist/Honorary Lecturer, University of Manchester

Mr Hisham Mehanna,
Consultant & Honorary Associate Professor, University Hospitals Coventry & Warwickshire NHS Trust

Dr Jane Montgomery,
Consultant in Anaesthetics and Critical Care, South Devon Healthcare NHS Foundation Trust

Professor Jon Moss,
Consultant Interventional Radiologist, North Glasgow Hospitals University NHS Trust

Dr Simon Padley,
Consultant Radiologist, Chelsea & Westminster Hospital

Dr Ashish Paul,
Medical Director, Bedfordshire PCT

Dr Sarah Purdy,
Consultant Senior Lecturer, University of Bristol

Professor Yit Chiun Yang,
Consultant Ophthalmologist, Royal Wolverhampton Hospitals NHS Trust

Observers

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Dr Morven Roberts,
Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

Pharmaceuticals Panel

Members

Chair,
Professor Imti Choonara,
Professor in Child Health, University of Nottingham

Deputy Chair,
Dr Yoon K Loke,
Senior Lecturer in Clinical Pharmacology, University of East Anglia

Dr Martin Ashton-Key,
Medical Advisor, National Commissioning Group, NHS London

Mr John Chapman,
Public contributor

Dr Peter Elton,
Director of Public Health, Bury Primary Care Trust

Dr Ben Goldacre,
Research Fellow, Division of Psychological Medicine and Psychiatry, King's College London

Dr James Gray,
Consultant Microbiologist, Department of Microbiology, Birmingham Children's Hospital NHS Foundation Trust

Ms Kylie Gyertson,
Oncology and Haematology Clinical Trials Manager, Guy's and St Thomas' NHS Foundation Trust London

Dr Jurjees Hasan,
Consultant in Medical Oncology, The Christie, Manchester

Dr Carl Heneghan,
Deputy Director Centre for Evidence-Based Medicine and Clinical Lecturer, Department of Primary Health Care, University of Oxford

Dr Dyfrig Hughes,
Reader in Pharmacoeconomics and Deputy Director, Centre for Economics and Policy in Health, IMSCaR, Bangor University

Dr Maria Kouimtzi,
Pharmacy and Informatics Director, Global Clinical Solutions, Wiley-Blackwell

Professor Femi Oyebode,
Consultant Psychiatrist and Head of Department, University of Birmingham

Dr Andrew Prentice,
Senior Lecturer and Consultant Obstetrician and Gynaecologist, The Rosie Hospital, University of Cambridge

Ms Amanda Roberts,
Public contributor

Dr Martin Shelly,
General Practitioner, Silver Lane Surgery, Leeds

Dr Gillian Shepherd,
Director, Health and Clinical Excellence, Merck Serono Ltd

Mrs Katrina Simister,
Assistant Director New Medicines, National Prescribing Centre, Liverpool

Professor Donald Singer,
Professor of Clinical Pharmacology and Therapeutics, Clinical Sciences Research Institute, CSB, University of Warwick Medical School

Mr David Symes,
Public contributor

Dr Arnold Zermansky,
General Practitioner, Senior Research Fellow, Pharmacy Practice and Medicines Management Group, Leeds University

Observers

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Mr Simon Reeve,
Head of Clinical and Cost-Effectiveness, Medicines, Pharmacy and Industry Group, Department of Health

Dr Heike Weber,
Programme Manager, Medical Research Council

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

Psychological and Community Therapies Panel

Members

Chair,
Professor Scott Weich,
Professor of Psychiatry, University of Warwick, Coventry

Deputy Chair,
Dr Howard Ring,
Consultant & University Lecturer in Psychiatry, University of Cambridge

Professor Jane Barlow,
Professor of Public Health in the Early Years, Health Sciences Research Institute, Warwick Medical School

Dr Sabyasachi Bhaumik,
Consultant Psychiatrist, Leicestershire Partnership NHS Trust

Mrs Val Carlill,
Public contributor

Dr Steve Cunningham,
Consultant Respiratory Paediatrician, Lothian Health Board

Dr Anne Hesketh,
Senior Clinical Lecturer in Speech and Language Therapy, University of Manchester

Dr Peter Langdon,
Senior Clinical Lecturer, School of Medicine, Health Policy and Practice, University of East Anglia

Dr Yann Lefeuvre,
GP Partner, Burrage Road Surgery, London

Dr Jeremy J Murphy,
Consultant Physician and Cardiologist, County Durham and Darlington Foundation Trust

Dr Richard Neal,
Clinical Senior Lecturer in General Practice, Cardiff University

Mr John Needham,
Public contributor

Ms Mary Nettle,
Mental Health User Consultant

Professor John Potter,
Professor of Ageing and Stroke Medicine, University of East Anglia

Dr Greta Rait,
Senior Clinical Lecturer and General Practitioner, University College London

Dr Paul Ramchandani,
Senior Research Fellow/Cons. Child Psychiatrist, University of Oxford

Dr Karen Roberts,
Nurse/Consultant, Dunston Hill Hospital, Tyne and Wear

Dr Karim Saad,
Consultant in Old Age Psychiatry, Coventry and Warwickshire Partnership Trust

Dr Lesley Stockton,
Lecturer, School of Health Sciences, University of Liverpool

Dr Simon Wright,
GP Partner, Walkden Medical Centre, Manchester

Observers

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Dr Morven Roberts,
Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

Expert Advisory Network

Members

Professor Douglas Altman,
Professor of Statistics in Medicine,
Centre for Statistics in Medicine,
University of Oxford

Professor John Bond,
Professor of Social Gerontology
& Health Services Research,
University of Newcastle upon Tyne

Professor Andrew Bradbury,
Professor of Vascular Surgery,
Solihull Hospital, Birmingham

Mr Shaun Brogan,
Chief Executive, Ridgeway
Primary Care Group, Aylesbury

Mrs Stella Burnside OBE,
Chief Executive, Regulation and
Improvement Authority, Belfast

Ms Tracy Bury,
Project Manager, World
Confederation of Physical Therapy,
London

Professor Iain T Cameron,
Professor of Obstetrics and
Gynaecology and Head of the
School of Medicine, University of
Southampton

Professor Bruce Campbell,
Consultant Vascular & General
Surgeon, Royal Devon & Exeter
Hospital, Wonford

Dr Christine Clark,
Medical Writer and Consultant
Pharmacist, Rossendale

Professor Collette Clifford,
Professor of Nursing and Head
of Research, The Medical School,
University of Birmingham

Professor Barry Cookson,
Director, Laboratory of Hospital
Infection, Public Health
Laboratory Service, London

Dr Carl Counsell,
Clinical Senior Lecturer in
Neurology, University of Aberdeen

Professor Howard Cuckle,
Professor of Reproductive
Epidemiology, Department
of Paediatrics, Obstetrics &
Gynaecology, University of Leeds

Professor Carol Dezateux,
Professor of Paediatric
Epidemiology, Institute of Child
Health, London

Mr John Dunning,
Consultant Cardiothoracic
Surgeon, Papworth Hospital NHS
Trust, Cambridge

Mr Jonothan Earnshaw,
Consultant Vascular Surgeon,
Gloucestershire Royal Hospital,
Gloucester

Professor Martin Eccles,
Professor of Clinical Effectiveness,
Centre for Health Services
Research, University of Newcastle
upon Tyne

Professor Pam Enderby,
Dean of Faculty of Medicine,
Institute of General Practice
and Primary Care, University of
Sheffield

Professor Gene Feder,
Professor of Primary Care
Research & Development, Centre
for Health Sciences, Barts and The
London School of Medicine and
Dentistry

Mr Leonard R Fenwick,
Chief Executive, Freeman
Hospital, Newcastle upon Tyne

Mrs Gillian Fletcher,
Antenatal Teacher and Tutor and
President, National Childbirth
Trust, Henfield

Professor Jayne Franklyn,
Professor of Medicine, University
of Birmingham

Mr Tam Fry,
Honorary Chairman, Child
Growth Foundation, London

Professor Fiona Gilbert,
Consultant Radiologist and NCRN
Member, University of Aberdeen

Professor Paul Gregg,
Professor of Orthopaedic Surgical
Science, South Tees Hospital NHS
Trust

Bec Hanley,
Co-director, TwoCan Associates,
West Sussex

Dr Maryann L Hardy,
Senior Lecturer, University of
Bradford

Mrs Sharon Hart,
Healthcare Management
Consultant, Reading

Professor Robert E Hawkins,
CRC Professor and Director of
Medical Oncology, Christie CRC
Research Centre, Christie Hospital
NHS Trust, Manchester

Professor Richard Hobbs,
Head of Department of Primary
Care & General Practice,
University of Birmingham

Professor Alan Horwich,
Dean and Section Chairman,
The Institute of Cancer Research,
London

Professor Allen Hutchinson,
Director of Public Health and
Deputy Dean of ScHARR,
University of Sheffield

Professor Peter Jones,
Professor of Psychiatry, University
of Cambridge, Cambridge

Professor Stan Kaye,
Cancer Research UK Professor of
Medical Oncology, Royal Marsden
Hospital and Institute of Cancer
Research, Surrey

Dr Duncan Keeley,
General Practitioner (Dr Burch &
Ptnrs), The Health Centre, Thame

Dr Donna Lamping,
Research Degrees Programme
Director and Reader in
Psychology, Health Services
Research Unit, London School of
Hygiene and Tropical Medicine,
London

Professor James Lindesay,
Professor of Psychiatry for the
Elderly, University of Leicester

Professor Julian Little,
Professor of Human Genome
Epidemiology, University of
Ottawa

Professor Alistaire McGuire,
Professor of Health Economics,
London School of Economics

Professor Neill McIntosh,
Edward Clark Professor of Child
Life and Health, University of
Edinburgh

Professor Rajan Madhok,
Consultant in Public Health, South
Manchester Primary Care Trust

Professor Sir Alexander Markham,
Director, Molecular Medicine
Unit, St James's University
Hospital, Leeds

Dr Peter Moore,
Freelance Science Writer, Ashtead

Dr Andrew Mortimore,
Public Health Director,
Southampton City Primary Care
Trust

Dr Sue Moss,
Associate Director, Cancer
Screening Evaluation Unit,
Institute of Cancer Research,
Sutton

Professor Miranda Mugford,
Professor of Health Economics
and Group Co-ordinator,
University of East Anglia

Professor Jim Neilson,
Head of School of Reproductive
& Developmental Medicine
and Professor of Obstetrics
and Gynaecology, University of
Liverpool

Mrs Julietta Patnick,
Director, NHS Cancer Screening
Programmes, Sheffield

Professor Robert Peveler,
Professor of Liaison Psychiatry,
Royal South Hants Hospital,
Southampton

Professor Chris Price,
Director of Clinical Research,
Bayer Diagnostics Europe, Stoke
Poges

Professor William Rosenberg,
Professor of Hepatology and
Consultant Physician, University
of Southampton

Professor Peter Sandercock,
Professor of Medical Neurology,
Department of Clinical
Neurosciences, University of
Edinburgh

Dr Philip Shackley,
Senior Lecturer in Health
Economics, Sheffield Vascular
Institute, University of Sheffield

Dr Eamonn Sheridan,
Consultant in Clinical Genetics, St
James's University Hospital, Leeds

Dr Margaret Somerville,
Director of Public Health
Learning, Peninsula Medical
School, University of Plymouth

Professor Sarah Stewart-Brown,
Professor of Public Health,
Division of Health in the
Community, University of
Warwick, Coventry

Dr Nick Summerton,
GP Appraiser and Codirector,
Research Network, Yorkshire
Clinical Consultant, Primary Care
and Public Health, University of
Oxford

Professor Ala Szczepura,
Professor of Health Service
Research, Centre for Health
Services Studies, University of
Warwick, Coventry

Dr Ross Taylor,
Senior Lecturer, University of
Aberdeen

Dr Richard Tiner,
Medical Director, Medical
Department, Association of the
British Pharmaceutical Industry

Mrs Joan Webster,
Consumer Member, Southern
Derbyshire Community Health
Council

Professor Martin Whittle,
Clinical Co-director, National
Co-ordinating Centre for Women's
and Children's Health, Lymington

Current and past membership details of all HTA programme 'committees' are available from the HTA website (www.hta.ac.uk)

Feedback

The HTA programme and the authors would like to know your views about this report.

The Correspondence Page on the HTA website (www.hta.ac.uk) is a convenient way to publish your comments. If you prefer, you can send your comments to the address below, telling us whether you would like us to transfer them to the website.

We look forward to hearing from you.

NETSCC, Health Technology Assessment
Alpha House
University of Southampton Science Park
Southampton SO16 7NS, UK
Email: hta@hta.ac.uk
www.hta.ac.uk

ISSN 1366-5278