Chronic pelvic pain and menorrhagia: Assessing treatment effectiveness

Daniels, J.P.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 9

Hysterectomy, endometrial destruction and Mirena® for heavy menstrual bleeding:
A systematic review and individual patient data meta-analysis
Chapter 9

Abstract

Objective: Heavy menstrual bleeding is a common condition that does not always respond to oral medication. We undertook an individual patient data (IPD) meta-analysis to evaluate the relative effectiveness of hysterectomy, endometrial destruction (both ‘first generation’ hysteroscopic and ‘second generation’ non-hysteroscopic techniques) and Mirena® (levonorgestrel releasing intrauterine system).

Methods: Individual patient data on 2814 women were available from 17 of the 30 randomised controlled trials identified (14 trials including 2448 women for first versus second generation endometrial destruction; 7 trials including 1127 women for hysterectomy versus first generation endometrial destruction; 5 trials including 304 women for second generation endometrial destruction versus Mirena®; 3 trials including 190 women for first generation endometrial destruction versus Mirena®; 1 trial including 236 women for hysterectomy versus Mirena®). Direct and indirect comparisons were made where appropriate to assess the effect of interventions on the primary outcome measure of patient dissatisfaction.

Results: At around 12 months, 7.3% (12.6% v 5.3%) more women were dissatisfied with outcome of first generation hysteroscopic techniques than hysterectomy (OR [Odds Ratio]:2.46, 95%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, p<0.00001) and time to resumption of normal activities (WMD:5.2 days, 95%CI 4.7 to 5.7, p<0.00001) were longer for hysterectomy. Unsatisfactory outcomes were comparable with 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. Indirect comparison suggested more unsatisfactory outcomes with second generation techniques than hysterectomy (10.6% v 5.3%; OR:2.32, 95%CI 1.27 to 4.24, p=0.006).

Conclusion: More women are dissatisfied following endometrial destruction than hysterectomy. However, dissatisfaction rates are low after all treatments and hysterectomy is associated with increased hospital stay and recovery period. Definitive evidence on effectiveness of Mirena® compared to more invasive procedures is lacking.

Keywords: Endometrial destruction, hysterectomy, heavy menstrual bleeding, menorrhagia, Mirena®
Introduction

Heavy menstrual bleeding (HMB) is a common problem among women of reproductive age. It is often incapacitating, expensive to treat and significantly impacts on women's quality of life. Many women are not happy with medical treatment and end up undergoing surgery. Hysterectomy was once the only surgical option for HMB, and indeed, almost half of the hysterectomies currently performed worldwide are for the treatment of HMB. Endometrial destruction (ED) techniques, which aim to destroy or remove the endometrial tissue, have become increasingly popular alternatives and, as a result, the number of hysterectomies in the UK has declined by 64% between 1995 and 2002. ED techniques were introduced in the 1980's, with rollerball ablation and transcervical resection emerging as the predominant approaches under direct hysteroscopic vision. Subsequently, there have been a development of a second generation of non-hysteroscopic techniques which are easier to perform. Here, endometrial destruction is achieved through a variety of modalities, including high temperature fluids, bipolar electrical, or microwave energy. Intrauterine coil devices were initially introduced as contraceptives, but the addition of progestogen resulted in reduced menstrual bleeding. Mirena, a levonorgesterol releasing intrauterine system provides a non-surgical alternative, which is reversible and fertility sparing.

Women and clinicians now have greater choice of treatment, although evidence to support decision making is inadequate. In the UK, guidelines from the National Institute of Healthcare and Clinical Excellence recommend the use of Mirena in the first instance for women with benign HMB, followed by ED if pharmaceutical treatments fail to resolve symptoms. Syntheses of evidence from randomised controlled trials (RCTs) comparing these treatments have been limited, partly due to scarcity of head-to-head comparisons and variation in outcome measurements used to evaluate effectiveness. We undertook a meta-analysis of individual patient data (IPD) from all relevant trials, to address previous deficiencies in evidence synthesis. IPD meta-analysis has a number of advantages over traditional published data reviews, including the ability to carry out data checks, standardise analytical methods and undertake subgroup analyses.

Methods

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

Literature search and study selection

The Cochrane Library, Medline (1966-2008), Embase (1980 to July 2008) and CINAHL databases (1982 to July 2008) were searched using relevant terms and word variants for population
and interventions. 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 ISRCTN 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 predetermined 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: Hysterectomy (performed abdominally, vaginally or laparoscopically); ‘first generation’ ED techniques (using operative hysteroscopy, including endometrial laser ablation, transcervical resection of the endometrium (TCRE) and rollerball endometrial ablation); ‘second generation’ ED techniques (including thermal balloon (Cavaterm®, Thermachoice® and Vesta®), microwave (Microsulis®), laser (ELITT®), bipolar radio frequency (NovaSure®), cryoablation and hydrothermal ablation); and levonorgesteral releasing intrauterine system (Mirena®). Our intention was to compare these categories of treatment against each other; studies were not included in the meta-analysis if a comparison between relevant categories did not exist. Data from studies making a comparison within these categories 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 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 pro-formas 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 IPD meta-analysis. The data were cleaned and results cross-checked against published reports of the trials. Where discrepancies existed authors were contacted for clarification.

Patient dissatisfaction was used as the primary outcome measure. Level of satisfaction was the most frequently measured outcome, but dissatisfaction was used here to simplify interpretation of statistical output. Responses were dichotomised, with an “unsure” or “dissatisfied” response interpreted as indicative of an unsatisfactory rating of treatment. Perception of treatment
(“moderate”/”poor”), recommendation to a friend (“not sure”/”not recommended”), or resolution of the major presenting problem (“no”) were used as substitutes if dissatisfaction was not reported explicitly, however not all trials could contribute to this end-point as a suitable measure was not always available. Ideally, a more disease specific quality of life tool would have been used\textsuperscript{18}, but the majority of studies did not employ such a measure.

Other outcome measures were bleeding scores (ranging from a minimum of 0 with no upper limit)\textsuperscript{19}, amenorrhoea rate (converted from a bleeding score of 0 where data existed, otherwise as reported), HMB rate (converted from bleeding scores of >100\textsuperscript{19} where data existed, otherwise as reported), EQ-5D utility score\textsuperscript{20}, SF-36 scores\textsuperscript{21}, duration of surgery/hospital stay, general anaesthesia rates, post-operative pain score (standardised 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\textsuperscript{8}. Pre-defined subgroups were age at randomisation (≤40 v >40 years), parity (nulliparous v parous), uterine cavity length (≤8 v >8 cm), presence or absence of fibroids/polyps and, where available, severity of bleeding at baseline (bleeding scores≤350 or >350). Time points up to 2 years were considered; results from the limited number of studies with longer term follow-up are given in the Data Supplement but are not referred to in detail. Data at less than 12 months were combined, and are described as results at 6 months.

All selected trials were assessed for their methodological quality, using received datasets where available in addition to the reported information. Quality was scrutinised by checking 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 sub-questions 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\textsuperscript{22}. IPD were reduced to AD to allow studies with AD only to be combined with those where IPD were obtained. Point estimates and 95% confidence intervals (CI) were calculated for individual studies at each time point. Differences in effect estimates between trials and the pre-defined subgroups of patients are displayed using odds ratio plots, with heterogeneity investigated using Cochran’s Q Statistic\textsuperscript{23}. Sensitivity 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 odds ratios (OR) using the method of Peto24, and, for continuous variables, weighted mean differences (WMD) were calculated25 at each time point.

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 6 months. If it was not possible to make a direct comparison between treatments (e.g. hysterectomy vs second generation ED), indirect estimates were made26 using a logistic regression model27 allowing for trial and treatment28. 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)29.

Access to IPD also allowed the inclusion of patient-level covariates to examine possible predictors of dissatisfaction. First, covariates were considered individually, whilst 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 ED (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. Revman v5.0 (Cochrane Collaboration, Denmark) and SAS v9.2 (SAS Institute, Cary, 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 techniques30-35. The seventh study used a combination of first and second generation in equal pro-
portions\textsuperscript{36} and was included here as a first generation comparison, with a sensitivity analysis performed without the trial. One study compared hysterectomy\textsuperscript{37} with Mirena\textsuperscript{®}. Fourteen studies compared first generation ED techniques with those of second generation\textsuperscript{38-51} and eight studies compared Mirena\textsuperscript{®} with ED, three of which were first generation\textsuperscript{62-64} and five second generation\textsuperscript{55-59}. Characteristics of these studies are shown in the Data Supplement. Data from a further five studies\textsuperscript{60-64}, which involved comparisons within first and second generation ED, were also received.

Trials comparing hysterectomy with ED and those comparing first and second generation ED involved women of a similar age, with averages of 40.6 years (95\%CI 40.3 to 40.9) and 41.0 years (95\%CI 40.8 to 41.2), respectively. Women involved in trials comparing Mirena\textsuperscript{®} with ED were slightly older, with an average age of 43.6 years (95\%CI 43.3 to 44.0). 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 (7/8 studies, 1278 of 1363 women), less for trials of ED techniques (7/14 studies, 1359 of 2448 women) and those involving Mirena\textsuperscript{®} (3/8 studies, 177 of 494 women). Overall, we received some IPD from 65\% (2814/4305) of women involved in the trials, although only 8 studies were able to provide all requested variables\textsuperscript{31,32,38,39,42,45,51,59}. The remaining studies had some missing information, with limited details on patient follow-up covering subsequent operations (e.g. hysterectomy following Mirena\textsuperscript{®}). See section on statistical analysis for details on how data from studies providing IPD was utilised.

**Study quality**

The methodological quality of the published data from the studies was variable (Figure 2, ). 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 had been performed yet only analysing those women that had received treatment. For 4 studies that reported non-ITT analyses\textsuperscript{39,42,45,57} 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\textsuperscript{®}. In the nine trials involving Mirena\textsuperscript{®}, only four had greater than 80\% of women with Mirena\textsuperscript{®} in situ at 12 months post-randomisation.
Figure 1  Study selection process for systematic review and individual patient data (IPD) meta-analysis of randomised trials comparing hysterectomy, ED techniques and Mirena® for HMB

<table>
<thead>
<tr>
<th>Total citations identified (n=557)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Articles retrieved for detailed evaluation (n=111)</td>
</tr>
<tr>
<td>Articles excluded (n=72):</td>
</tr>
<tr>
<td>Abstract/letter (n=6)</td>
</tr>
<tr>
<td>Data not extractable (n=9)</td>
</tr>
<tr>
<td>Duplicates (n=57)</td>
</tr>
<tr>
<td>Articles included (n=39)</td>
</tr>
<tr>
<td>Unpublished studies (n=1*)</td>
</tr>
<tr>
<td>Articles included in systematic review and meta-analysis (n=40)</td>
</tr>
<tr>
<td>Trials unable to contribute to meta-analysis (n=10):</td>
</tr>
<tr>
<td>First generation v first generation (n=4)</td>
</tr>
<tr>
<td>Second generation v second generation (n=4)</td>
</tr>
<tr>
<td>Hysterectomy v hysterectomy (n=2)</td>
</tr>
<tr>
<td>Hysterectomy v Mirena</td>
</tr>
<tr>
<td>(n=7; 1127 women)</td>
</tr>
<tr>
<td>IPD obtained from</td>
</tr>
<tr>
<td>6 trials (1042 women)</td>
</tr>
<tr>
<td>3 trials (570 women)</td>
</tr>
<tr>
<td>Hysterectomy v Mirena</td>
</tr>
<tr>
<td>(n=1; 236 women)</td>
</tr>
<tr>
<td>IPD obtained from</td>
</tr>
<tr>
<td>1 trial (236 women)</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>First v second generation</td>
</tr>
<tr>
<td>endometrial destruction</td>
</tr>
<tr>
<td>(n=14; 2448 women)</td>
</tr>
<tr>
<td>IPD obtained from</td>
</tr>
<tr>
<td>7 trials (1359 women)</td>
</tr>
<tr>
<td>6 trials (1083 women)</td>
</tr>
<tr>
<td>Mirena v first generation</td>
</tr>
<tr>
<td>endometrial destruction</td>
</tr>
<tr>
<td>(n=3; 190 women)</td>
</tr>
<tr>
<td>No IPD obtained</td>
</tr>
<tr>
<td>(0 women)</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>Mirena v second</td>
</tr>
<tr>
<td>generation endometrial destruction</td>
</tr>
<tr>
<td>(n=5; 304 women)</td>
</tr>
<tr>
<td>IPD obtained from</td>
</tr>
<tr>
<td>3 trials (177 women)</td>
</tr>
<tr>
<td>1 trial (83 women)</td>
</tr>
<tr>
<td>Satisfaction as an outcome within 2 years (with IPD)</td>
</tr>
</tbody>
</table>
**Figure 2** Quality of studies included in systematic review and individual patient data (IPD) meta-analysis of randomised trials comparing hysterectomy, ED techniques and Mirena® for HMB. Numbers inside bars are numbers of studies.

### Effectiveness in reducing dissatisfaction with treatment

**Hysterectomy versus first generation endometrial destruction**

More women were dissatisfied at 12 months following first generation ED than hysterectomy (12.6% vs 5.3%; 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). 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 2.18-6.46, p<0.0001) and an analysis using dissatisfaction at any time point (OR:3.37, 95%CI 2.14-5.31, p<0.0001). There was no evidence of any differences be-
between subgroups (see data collection and study quality assessment section), including between studies providing IPD or AD (test for heterogeneity: p=0.9).

**Figure 3** Dissatisfaction at 12 months - Hysterectomy vs First Generation ED

<table>
<thead>
<tr>
<th>Study</th>
<th>1st Gen Events</th>
<th>Hysterectomy Events</th>
<th>Total</th>
<th>Weight</th>
<th>Peto Odds Ratio Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dickersin et al</td>
<td>13</td>
<td>107</td>
<td>7</td>
<td>103</td>
<td>26.1% 1.88 [0.96, 3.68]</td>
</tr>
<tr>
<td>Crolignani PG et al</td>
<td>5</td>
<td>38</td>
<td>2</td>
<td>36</td>
<td>9.3% 2.61 [0.94, 7.93]</td>
</tr>
<tr>
<td>O’Connor H et al</td>
<td>7</td>
<td>106</td>
<td>2</td>
<td>50</td>
<td>10.7% 1.61 [0.7, 3.55]</td>
</tr>
<tr>
<td>Pinion S et al</td>
<td>13</td>
<td>104</td>
<td>4</td>
<td>93</td>
<td>22.4% 2.81 [0.76, 10.39]</td>
</tr>
<tr>
<td>Dwyer N et al</td>
<td>19</td>
<td>99</td>
<td>6</td>
<td>97</td>
<td>31.5% 3.21 [0.96, 10.61]</td>
</tr>
</tbody>
</table>

**Total (95% CI)**: 454 events, 382 total, 100.0% weight, 2.46 [1.84, 3.68]

**Total events**: 57 total, 21 Hysterectomy

Heterogeneity: Ch2 = 1.15, df = 4 (P = 0.89); I² = 0%

Test for overall effect: Z = 2.15 (P = 0.032)

---

**First versus second generation endometrial destruction techniques**

Similar dissatisfaction rates were seen with second and first generation ED (10.6% vs 12.2%; OR:1.20, 95%CI 0.88 to 1.62, p=0.2; test for heterogeneity: p=0.7) (Figure 4). Similar 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 onto receive 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).

**Figure 4** Dissatisfaction at 12 months - First vs Second Generation ED

<table>
<thead>
<tr>
<th>Study</th>
<th>1st Gen Events</th>
<th>2nd Gen Events</th>
<th>Total</th>
<th>Weight</th>
<th>Peto Odds Ratio Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brun</td>
<td>3</td>
<td>16</td>
<td>2</td>
<td>27</td>
<td>2.5% 2.95 [0.24, 36.00]</td>
</tr>
<tr>
<td>Cooper J 04</td>
<td>1</td>
<td>101</td>
<td>3</td>
<td>201</td>
<td>2.1% 0.68 [0.04, 10.59]</td>
</tr>
<tr>
<td>Perino</td>
<td>5</td>
<td>55</td>
<td>3</td>
<td>56</td>
<td>4.5% 1.74 [0.26, 11.42]</td>
</tr>
<tr>
<td>Duleba</td>
<td>10</td>
<td>72</td>
<td>16</td>
<td>156</td>
<td>12.0% 1.43 [0.45, 4.53]</td>
</tr>
<tr>
<td>Hawe</td>
<td>3</td>
<td>33</td>
<td>4</td>
<td>37</td>
<td>3.8% 0.83 [0.11, 6.38]</td>
</tr>
<tr>
<td>Van Zon-Rabelink</td>
<td>13</td>
<td>58</td>
<td>15</td>
<td>75</td>
<td>13.1% 1.16 [0.38, 3.47]</td>
</tr>
<tr>
<td>Cooper J 02</td>
<td>5</td>
<td>82</td>
<td>11</td>
<td>154</td>
<td>8.1% 0.85 [0.21, 3.43]</td>
</tr>
<tr>
<td>Pellacano</td>
<td>14</td>
<td>36</td>
<td>7</td>
<td>37</td>
<td>9.2% 2.40 [0.64, 8.96]</td>
</tr>
<tr>
<td>Souyal</td>
<td>19</td>
<td>48</td>
<td>15</td>
<td>45</td>
<td>13.1% 1.31 [0.43, 3.94]</td>
</tr>
<tr>
<td>Cooper K</td>
<td>33</td>
<td>128</td>
<td>29</td>
<td>121</td>
<td>28.0% 1.10 [0.52, 2.34]</td>
</tr>
<tr>
<td>Meyer</td>
<td>1</td>
<td>120</td>
<td>5</td>
<td>125</td>
<td>3.5% 0.27 [0.03, 2.24]</td>
</tr>
</tbody>
</table>

**Total (95% CI)**: 751 events, 1034 total, 100.0% weight, 1.20 [0.88, 1.62]

**Total events**: 107 total, 110 Hysterectomy

Heterogeneity: Ch2 = 7.47, df = 10 (P = 0.68); I² = 0%

Test for overall effect: Z = 1.17 (P = 0.24)
Mirena® versus endometrial destruction techniques

Rates of dissatisfaction with Mirena® and second generation ED were similar (18.1% vs 22.5%; OR:0.76, 95%CI 0.38 to 1.53, p=0.4) (Figure 5). The combined estimate of this and the one study which compared Mirena® with first generation ED[4] (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; test for heterogeneity over all studies: p=0.1). Overall rates of dissatisfaction were 17.2% for Mirena® and 18.2% for both first and second generation ED. Lack of IPD prohibited any further investigation of subgroups or repeated measures.

Indirect comparisons of hysterectomy with second generation endometrial destruction techniques and Mirena®

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. This is confirmed by the repeated measures analysis over all 3 time points, which perforce 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 (OR:2.22, 95%CI:0.94-5.29, p=0.07), but given the lack of precision from Mirena® comparisons this was not a surprising result.

Figure 5  Dissatisfaction at 12 months – First and Second generation ED v Mirena®
Predictors of dissatisfaction

For second generation ED, uterine cavity length was the strongest predictor of dissatisfaction (p=0.02), with shorter uterine cavity length (<=8cm v >8cm) associated with reduced rates (OR: 0.59, 95%CI 0.38-0.93; p=0.02) (Table 1). 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 3 studies had data on fibroids/polyps. There were no convincing associations with any of the variables for hysterectomy or first generation ED.

Effectiveness in improving other outcomes

Hysterectomy versus endometrial destruction and Mirena®

These comparisons focused on recovery times and quality of life, as estimates of post operative menstrual blood loss are redundant after hysterectomy. ED offered quicker surgery (WMD:32 minutes, 95%CI 30 to 34, p<0.0001), shorter hospital stay (WMD:3.0 days, 95%CI 2.9 to 3.1, p<0.0001), faster recovery periods (time to return to normal activities: WMD:5.2 days, 95%CI 4.7 to 5.7, p<0.00001) and less pain post-operatively (WMD:2.5 points, 95%CI 2.2 to 2.9, p<0.0001), although estimates of differences for some of these parameters should be used with caution given high variability between studies. One study suggested no obvious difference in EQ-5D utility score, whilst another suggested differences in favour of hysterectomy in the general health (WMD:9.6 points, 95%CI 5.7 to 13.5, p<0.0001), social functioning (WMD:24 points, 95%CI 21 to 27, p<0.0001) and vitality (WMD:13 points, 95%CI 9.3 to 16, p<0.0001) domains.
of the SF-36 questionnaire. Peri-operative adverse events associated with hysterectomy were relatively few (each 0.5% to 2.0%) but urinary tract infections were more common with hysterectomy (43/530, 8.1%) than ED (9/585, 1.5%). Of the women who were initially treated with ED, 15% had undergone a hysterectomy by two years.

Table 1  Results from logistic regression analysis with dissatisfaction at 12 months as the outcome

<table>
<thead>
<tr>
<th></th>
<th>Hysterectomy</th>
<th>1st generation ED</th>
<th>2nd generation ED</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (95% CI)*</td>
<td>P-value</td>
<td>OR (95% CI)*</td>
</tr>
<tr>
<td>Individual estimates*:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uterine cavity length</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(&lt;=8 v &gt;8)</td>
<td>-</td>
<td>0.97 (0.38-2.44)</td>
<td>0.9</td>
</tr>
<tr>
<td>Age (&lt;=40 v &gt;40)</td>
<td>2.28 (0.66-7.89)</td>
<td>0.2</td>
<td>1.21 (0.81-1.81)</td>
</tr>
<tr>
<td>Fibroids/polyps (absence v presence)</td>
<td>0.51 (0.14-1.93)</td>
<td>0.3</td>
<td>1.15 (0.55-2.38)</td>
</tr>
<tr>
<td>Parity (nullparous v parous)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline bleeding score (&lt;=350 v 350)</td>
<td>-</td>
<td></td>
<td>0.73 (0.27-1.97)</td>
</tr>
</tbody>
</table>

# Estimates <1 favour the first subgroup listed, i.e. have reduced dissatisfaction
* 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®, whilst the only statistically significant effect observed in the SF-36 questionnaire was in the pain domain, favouring hysterectomy (WMD:9.6 points, 95%CI 2.7 to 16.6, p=0.007). All results were consistent over subgroups.

First versus second generation endometrial destruction techniques
The proportion of women with amenorrhoea or still experiencing HMB was similar in both groups at all time points apart from at 2 years, where there was a borderline significant difference in favour of 2nd generation techniques ([amenorrhea] 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). Change from baseline analysis of bleeding scores showed no evidence of a difference at any of the time points. Two studies45,47 using the SF-36 questionnaire and one small study38 using the EQ-5D questionnaire showed no consistent difference between first and second generation techniques.

Second generation ED was quicker (WMD:15 minutes; 95%CI 14 to 15, p<0.0001), and less likely to need general anaesthesia (OR:0.16, 95%CI 0.12 to 0.20, p<0.0001), although highly significant heterogeneity makes estimates difficult to interpret. Lower use of general anaesthesia
with second generation ED translated to a slightly quicker time to normal activities (WMD:0.48 days, 95%CI 0.20 to 0.75, p=0.0008) and time to return to work (WMD:1.36 days, 95%CI 0.69 to 2.03, p<0.0001). Post-operative pain was similar following either method of ED. Adverse events were relatively low in both groups (each <2%), but peri-operative complications such as uterine perforation, excessive bleeding, fluid overload and cervical laceration were higher with first generation ED. The number of women requiring a subsequent hysterectomy was lower for second generation ED, but these differences were not large enough to be statistically significant within the first two years ([1 year] 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% (74/2265) and 8% (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 destruction 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 not at 12 months (OR:0.74, 95%CI 0.34 to 1.61, p=0.5), where most trials had made this measurement. Amenorrhea rates were similar at all time points. Change in bleeding scores favoured ED at 12 months only (WMD:38 points; 95%CI 15 to 60, p=0.0009). Other outcome measures could not separate the two treatments. Two studies[57,59] provided SF-36 change from baseline scores and no differences were found in any of the domains. The number of women subsequently undergoing a hysterectomy were similar at each time point; rates at 12 months were 2% (2/86) for ED and 7% (6/89) for Mirena®. A high proportion of women originally prescribed Mirena® discontinued use of this treatment: 30/191(16%) at 12 months, rising to 29/105 (28%) by 2 years. Reported adverse events were low with Mirena®; only around 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 ED.

Discussion

Main 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. Based on direct and indirect comparisons using all available data, the review found that both first and second generation ED 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 ED were quicker, had faster recovery times, were associated with fewer
adverse procedural events and could be offered under local anaesthetic. Less women subsequently converted to hysterectomy after second generation ED compared with first generation ED, but this difference was not statistically significant. Shorter uterine cavity length showed was associated with lower levels of dissatisfaction for second generation ED. Comparisons of ED with Mirena® suggest comparable efficacy, although trials were generally small and suffered from poor compliance in the Mirena® arm.

Strengths and limitations of the review

We used optimal methodology, complying with guidelines on reporting of systematic reviews and meta-analyses. 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, to improve the assessment of study quality, to 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, have not previously 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 trialists 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 ED and Mirena® focused on comparing reduction in bleeding, whilst hysterectomy trials focused on patient satisfaction, quality of life and resource usage.

Interpretation

In this review we found that more women were dissatisfied following ED when compared to hysterectomy, although operating time, total hospital stay and recovery period are much less for ED. Nevertheless, rates of dissatisfaction are relatively low for ED and it is an effective alternative to hysterectomy for women with abnormal uterine bleeding who do not seek amenorrhoea. Whilst 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. In contrast, ED has low rates of complication. All these factors need to be taken into consideration when considering any potential benefit of hysterectomy. Indeed, nearly fourteen women need to be treated with hysterectomy to prevent a single woman being dissatisfied with ED (number needed to treat from 1st generation comparison: 13.7, 95%CI 9.1 to 28.4). Given the extra resources and recovery period needed for hysterectomy a benefit of this size may not be large enough to warrant its use. The cost-effectiveness of these treatments will be an important factor and this is to be examined separately in a concurrent study.
We found that second generation techniques, such as thermal balloon ablation (Thermachoice® and Cavaterm®) 42,67,68, the Novasure® device44, or microwave (Microsulis®)47,69, were at least as effective as first generation techniques. 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 ED 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 fails70. It may even be an alternative to drug treatment as a first line agent but we did not address this question in our review. However, research on Mirena® presents some difficulties in interpretation, as seen in the trial by Hurskainen et al37, which compared Mirena® with hysterectomy. Whilst the study was well conducted and reported, the lack of further investigation into the analysis of the primary outcome measure (EQ-5D quality of life of life measure) made the interpretation of 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 ED and this information should be used as part of consultation with women making a choice about treatment options when initial drug treatment fails to control heavy menstrual bleeding. ED is satisfactory for a very high proportion of women, but, if complete cessation of bleeding is sought, then hysterectomy may be offered. Although the evidence is not strong, our findings concur with a recent NICE recommendation that women should be offered Mirena® before more invasive procedures, particularly as this can be fitted in primary care10.

Implications for research
This review has shown that further investment in a randomised controlled trial comparing hysterectomy with second generation ED 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 (4/5 years or more) is limited to a handful of studies involving differing comparisons36,37,71,72. Mirena®, in particular, versus alternative forms of surgical treatment requires further research. Whilst 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 and provides the most definitive possible synthesis 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. 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.

Acknowledgements

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.

Supplementary material

Data supplements available at http://www.bmj.com/content/341/bmj.c3929

*Members of Collaborative group

The following authors provided us with individual patient data 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; M.Y. Bongers, Maxima Medical Centre, Veldhoven, The Netherlands; J-L Brun, Hopital Universitaria Pellegrin, Bordeaux, France; R Busfield, data supplied by M Sowter, Auckland Obstetric Centre, New Zealand; T.J. Clark, Birmingham Womens Hospital, Birmingham, UK; J Cooper (2004 trial), data supplied by Microsulis Medical Ltd, Hampshire, UK; K.G. Cooper, Aberdeen Royal Infirmary, Aberdeen, UK; S.L. Corson (2001 trial), data supplied by Boston Scientific Corporation, Marlborough, USA; K Dickersin, John 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; W.R. Meyer; data supplied by Ethicon, A Johnson & Johnson Company, New Jersey, USA; H. O’Connor, Coombe Women’s Hospital, Dublin 8, Ireland; S. Pinion, Aberdeen Royal Infirmary, Aberdeen, UK; A.M. Sambrook, Aberdeen Royal Infirmary, Aberdeen, UK; W.H. Tam, Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong, China; I.A.A. van Zon-Rabelink, Medical Spectrum Twente, Enschede, The Netherlands, E. Zupi, Tor Vergata University, Rome, Italy.
References


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 July;116(8):1033-7.


