A randomised controlled trial of the clinical effectiveness and cost-effectiveness of the levonorgestrel-releasing intrauterine system in primary care against standard treatment for menorrhagia: the ECLIPSE trial

Janesh K Gupta, Jane P Daniels, Lee J Middleton, Helen M Pattison, Gail Prileszky, Tracy E Roberts, Sabina Sanghera, Pelham Barton, Richard Gray and Joe Kai on behalf of the ECLIPSE Collaborative Group

NHS

National Institute for Health Research

A randomised controlled trial of the clinical effectiveness and cost-effectiveness of the levonorgestrel-releasing intrauterine system in primary care against standard treatment for menorrhagia: the ECLIPSE trial

Janesh K Gupta,[1,2] Jane P Daniels,[3]* Lee J Middleton,[3] Helen M Pattison,[4] Gail Prileszky,[5] Tracy E Roberts,[6] Sabina Sanghera,[6] Pelham Barton,[6] Richard Gray[7] and Joe Kai[5] on behalf of the ECLIPSE Collaborative Group

[1]School of Clinical and Experimental Medicine, University of Birmingham, Birmingham, UK
[2]Birmingham Women's Hospital NHS Foundation Trust, Edgbaston, Birmingham, UK
[3]Birmingham Clinical Trials Unit, University of Birmingham, Birmingham, UK
[4]School of Health and Life Sciences, Aston University, Birmingham, UK
[5]Division of Primary Care, University of Nottingham, Nottingham, UK
[6]Health Economics Unit, University of Birmingham, Birmingham, UK
[7]Clinical Trials Service Unit, University of Oxford, Oxford, UK

*Corresponding author

Health Technology Assessment

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

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

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

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

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

This report

Abstract

A randomised controlled trial of the clinical effectiveness and cost-effectiveness of the levonorgestrel-releasing intrauterine system in primary care against standard treatment for menorrhagia: the ECLIPSE trial

Janesh K Gupta,[1,2] Jane P Daniels,[3]* Lee J Middleton,[3]
Helen M Pattison,[4] Gail Prileszky,[5] Tracy E Roberts,[6] Sabina Sanghera,[6]
Pelham Barton,[6] Richard Gray[7] and Joe Kai[5] on behalf of the ECLIPSE
Collaborative Group

[1]School of Clinical and Experimental Medicine, University of Birmingham, Birmingham, UK
[2]Birmingham Women's Hospital NHS Foundation Trust, Edgbaston, Birmingham, UK
[3]Birmingham Clinical Trials Unit, University of Birmingham, Birmingham, UK
[4]School of Health and Life Sciences, Aston University, Birmingham, UK
[5]Division of Primary Care, University of Nottingham, Nottingham, UK
[6]Health Economics Unit, University of Birmingham, Birmingham, UK
[7]Clinical Trials Service Unit, University of Oxford, Oxford, UK

*Corresponding author j.p.daniels@bham.ac.uk

Background: Heavy menstrual bleeding (HMB) is a common problem, yet evidence to inform decisions about initial medical treatment is limited.

Objectives: To assess the clinical effectiveness and cost-effectiveness of the levonorgestrel-releasing intrauterine system (LNG-IUS) (Mirena®, Bayer) compared with usual medical treatment, with exploration of women's perspectives on treatment.

Design: A pragmatic, multicentre randomised trial with an economic evaluation and a longitudinal qualitative study.

Setting: Women who presented in primary care.

Participants: A total of 571 women with HMB. A purposeful sample of 27 women who were randomised or ineligible owing to treatment preference participated in semistructured face-to-face interviews around 2 and 12 months after commencing treatment.

Interventions: LNG-IUS or usual medical treatment (tranexamic acid, mefenamic acid, combined oestrogen–progestogen or progesterone alone). Women could subsequently swap or cease their allocated treatment.

Outcome measures: The primary outcome was the patient-reported score on the Menorrhagia Multi-Attribute Scale (MMAS) assessed over a 2-year period and then again at 5 years. Secondary outcomes included general quality of life (QoL), sexual activity, surgical intervention and safety. Data were analysed using iterative constant comparison. A state transition model-based cost–utility analysis was undertaken alongside the randomised trial. Quality-adjusted life-years (QALYs) were derived from the European Quality

of Life-5 Dimensions (EQ-5D) and the Short Form questionnaire-6 Dimensions (SF-6D). The intention-to-treat analyses were reported as cost per QALY gained. Uncertainty was explored by conducting both deterministic and probabilistic sensitivity analyses.

Results: The MMAS total scores improved significantly in both groups at all time points, but were significantly greater for the LNG-IUS than for usual treatment [mean difference over 2 years was 13.4 points, 95% confidence interval (CI) 9.9 to 16.9 points; $p < 0.001$]. However, this difference between groups was reduced and no longer significant by 5 years (mean difference in scores 3.9 points, 95% CI –0.6 to 8.3 points; $p = 0.09$). By 5 years, only 47% of women had a LNG-IUS in place and 15% were still taking usual medical treatment. Five-year surgery rates were low, at 20%, and were similar, irrespective of initial treatments. There were no significant differences in serious adverse events between groups. Using the EQ-5D, at 2 years, the relative cost-effectiveness of the LNG-IUS compared with usual medical treatment was £1600 per QALY, which by 5 years was reduced to £114 per QALY. Using the SF-6D, usual medical treatment dominates the LNG-IUS. The qualitative findings show that women's experiences and expectations of medical treatments for HMB vary considerably and change over time. Women had high expectations of a prompt effect from medical treatments.

Conclusions: The LNG-IUS, compared with usual medical therapies, resulted in greater improvement over 2 years in women's assessments of the effect of HMB on their daily routine, including work, social and family life, and psychological and physical well-being. At 5 years, the differences were no longer significant. A similar low proportion of women required surgical intervention in both groups. The LNG-IUS is cost-effective in both the short and medium term, using the method generally recommended by the National Institute for Health and Care Excellence. Using the alternative measures to value QoL will have a considerable impact on cost-effectiveness decisions. It will be important to explore the clinical and health-care trajectories of the ECLIPSE (clinical effectiveness and cost-effectiveness of levonorgestrel-releasing intrauterine system in primary care against standard treatment for menorrhagia) trial participants to 10 years, by which time half of the cohort will have reached menopause.

Contents

List of tables

List of figures

List of abbreviations

BMI	body mass index	NICE	National Institute for Health and Care Excellence
CEAC	cost-effectiveness acceptability curve	NSAID	non-steroidal anti-inflammatory drug
CI	confidence interval	PSA	probabilistic sensitivity analysis
COC	combined oral contraceptive	QALY	quality-adjusted life-year
CUA	cost–utility analysis	QoL	quality of life
ECLIPSE	clinical effectiveness and cost-effectiveness of levonorgestrel-releasing intrauterine system in primary care against standard treatment for menorrhagia trial	RCOG	Royal College of Obstetricians and Gynaecologists
		RCT	randomised controlled trial
EQ-5D	European Quality of Life-5 Dimensions	SAQ	Sexual Activity Questionnaire
		SD	standard deviation
GP	general practitioner	SF-36	Short Form questionnaire-36 items
HMB	heavy menstrual bleeding	SF-6D	Short Form questionnaire-6 Dimensions
ICER	incremental cost-effectiveness ratio	VAS	visual analogue scale
IUS	intrauterine system	WTP	willingness to pay
LNG-IUS	levonorgestrel-releasing intrauterine system		
MMAS	Menorrhagia Multi-Attribute Scale		

Plain English summary

Heavy menstrual bleeding (HMB; or heavy periods) is a common problem in women presenting to their general practitioners. Initial treatment should usually be medical, using non-hormonal pills, the contraceptive pill or the levonorgestrel contraceptive coil. However, it is unclear whether or not the coil is more effective than pills and if it is more acceptable to women, particularly in the medium term.

We randomly assigned 571 women with HMB to treatment with the coil or pills. We asked them to assess the impact that their periods had on their daily life, at intervals over a 5-year period. We also explored their experiences and expectations through interviews.

All participants found that their symptoms and quality of life significantly improved, regardless of which treatment they started with, although those who had the coil fitted showed more improvement within 2 years. Not all women continued with the treatment: by 5 years, just over half had the coil removed because of unpredictable bleeding, and 85% of those taking pills had changed treatments, mainly to the coil, or stopped treatment. One in five women in each group went on to have surgery, for example a hysterectomy. We could not conclude that the coil was definitely a cost-effective treatment because the results differed depending on the calculation used.

Our trial shows that medical treatments should be considered first in women seeking treatment in primary care. Both the coil and pills are effective treatments over a 5-year period, with the coil more effective over the first 2 years only.

Scientific summary

Background

Heavy menstrual bleeding (HMB), also called menorrhagia, is a common problem that can significantly affect women's lives, burdening individuals and health-care systems. Initial management should usually be medical, using tranexamic acid or mefenamic acid, combined oral contraceptives or the levonorgestrel-releasing intrauterine system (LNG-IUS) (Mirena®, Bayer). Short-term reduction in menstrual blood loss may not translate into an improvement in a woman's overall menstrual experience or quality of life (QoL) or her need to seek further treatment, particularly as LNG-IUS discontinuation rates can be high.

Objectives

Primary objective
To assess the clinical effectiveness of the LNG-IUS compared with usual medical treatment for women seeking treatment for HMB in primary care in the short term (2 years following randomisation).

Secondary objective

- To assess the clinical effectiveness of initial treatment with the LNG-IUS compared with usual medical treatment for women seeking treatment for HMB in primary care in the medium term (5 years following randomisation).
- To determine the cost-effectiveness of the LNG-IUS compared with usual medical treatment in the short and medium term.
- To explore the perspectives of trial participants or women with HMB who declined to be randomised in a longitudinal qualitative study.
- To measure the reliability and validity properties of the Menorrhagia Multi-Attribute Scale (MMAS).

Methods

Study randomised trial
Women between 25 and 50 years who presented to their general practitioner (GP) with HMB over at least three consecutive cycles were eligible to participate. Women were excluded if they intended to become pregnant over the next 5 years, were taking hormone replacement therapy or tamoxifen (Soltamox™, Rosemount Pharmaceuticals), had intermenstrual or post-coital bleeding or findings suggestive of fibroids or other disorders, or had contraindications to, or a preference for, either the LNG-IUS or usual medical treatments. All patients provided written informed consent.

Patients were allocated to a treatment group in a 1 : 1 ratio via a central randomisation service, with minimisation used to achieve balance between the groups with respect to age (< 35 years or ≥ 35 years), body mass index (≤ 25 kg/m^2 or > 25 kg/m^2), duration of symptoms (< 1 year or ≥ 1 year), need for contraception and presence or absence of menstrual pain.

Participants were randomly assigned to either the LNG-IUS or usual medical treatment. Usual treatment options included mefenamic acid, tranexamic acid, norethisterone, a combined oestrogen–progestogen or progesterone-only oral contraceptive pill (any formulation), or medroxyprogesterone acetate (Depo-Provera™, Pfizer Ltd) injection and were chosen by the physician and patient on the basis of contraceptive needs or the desire to avoid hormonal treatment.

The primary outcome measure was the condition-specific MMAS, which is designed to measure the effect of HMB on six domains of daily life (practical difficulties, social life, psychological health, physical health, work and daily routine, and family life and relationships).

Secondary outcome measures included general health-related QoL [using the Short Form questionnaire-36 items (SF-36) and the European Quality of Life-5 Dimensions (EQ-5D)] and sexual activity. Outcomes were obtained before randomisation and at 6 months, 1, 2 and 5 years after randomisation. Data were collected regarding all serious adverse events, hospitalisations and reasons for discontinuation of the study drug.

The study was designed for 90% power (at $p < 0.05$) to detect small to moderate (0.3 standard deviations) differences in the mean MMAS score between groups at any one time point. This required an enrolment of 470 patients; we increased the sample size to 570 to allow for up to 20% loss to follow-up.

Analysis was by intention to treat with continuous measures compared using multilevel repeated-measures models, including parameters allowing for participant, treatment, time and baseline score. A range of sensitivity analyses was performed on the primary outcome to test the robustness of the results.

Longitudinal qualitative study

A purposeful sample of women randomised into the ECLIPSE (clinical effectiveness and cost-effectiveness of levonorgestrel-releasing intrauterine system in primary care against standard treatment) trial for menorrhagia, or who declined participation owing to strong preferences for a particular treatment, was selected. Fifty-one semistructured face-to-face interviews were conducted in the respondents' homes by a female researcher.

Initial interviews were conducted within 1–3 months of commencing medical treatments and between 12 and 14 months later. Interviews followed broad topic areas including menstrual history and experience, social factors influencing experience, relationships and family life, and emotional well-being as well as treatment preference and experience of treatments. Interviews were transcribed verbatim. Data were analysed using iterative constant comparison to identify emerging themes until saturation and respondent validation.

Economic evaluation

A state transition model-based cost–utility analysis was undertaken alongside the randomised trial. The model structure was informed by clinical input and the pathways followed by women in the trial, and all parameters used were based on trial data.

The evaluation was based on an outcome of cost per quality-adjusted life-year (QALY), derived from the EQ-5D and SF-36 values [using a Short Form questionnaire-6 Dimensions (SF-6D) conversion algorithm], obtained up to 5 years. Costs were collected from a UK NHS perspective in a primary care setting, with data on health-care resource use collected from women. Two analyses based on 2- and 5-year time horizons were conducted.

The intention-to-treat analyses were reported, in terms of incremental cost-effectiveness ratios (ICERs), as cost per QALY gained. The base-case analysis and three deterministic sensitivity analyses were carried out using the EQ-5D and then repeated using the SF-36 as an additional sensitivity analysis. Uncertainty in the model was explored by conducting both deterministic and probabilistic sensitivity analysis (PSA); the latter used to construct cost-effectiveness acceptability curves.

Results

Clinical effectiveness

Between February 2005 and July 2009, a total of 571 women with HMB from 63 UK centres were randomly assigned to either the LNG-IUS (285 women) or usual medical treatment (286 women). In the usual medical treatment group, 75% were initially treated with mefenamic acid or tranexamic acid or a combination of these two drugs. Study questionnaire booklets were returned by 84% and 74% of the participants at 2 and 5 years, respectively.

Women in the LNG-IUS group were almost twice as likely as those in the usual treatment group to still be receiving their assigned treatment at 2 years (64% vs. 38%; $p < 0.001$), while by 5 years only 47% had a LNG-IUS in place and 15% were still taking usual medical treatment. The most common reasons cited for discontinuation of the LNG-IUS were lack of effectiveness and irregular or prolonged bleeding. Of the 228 women who discontinued usual medical treatment by 5 years, 43% switched to the LNG-IUS. The most common reason for discontinuation of usual medical therapy was lack of effectiveness.

At 6 months, total scores on the MMAS had improved significantly, compared with baseline scores, and these improvements were maintained through to 2 years. Improvements in these scores were significantly greater among women assigned to the LNG-IUS than among those assigned to usual treatment over the course of 2 years [mean difference in scores 13.4 points, 95% confidence interval (CI) 9.9 to 16.9 points; $p < 0.001$]. However, this difference between groups was reduced and no longer significant by 5 years (mean difference in scores 3.9 points, 95% CI –0.6 to 8.3 points; $p = 0.09$).

In a sensitivity analysis that excluded women who crossed over from the assigned treatment to the other study treatments, improvement with the LNG-IUS compared with usual medical treatment increased (mean difference in scores over 2 years was 17.8 points, 95% CI 14.1 to 21.5 points; $p < 0.001$). Other sensitivity analyses yielded results that were not materially different from the results of the primary analysis.

Domains from the SF-36 were generally significantly improved from baseline in both groups at all time points up to 2 years, although the scores were better for women in the LNG-IUS group than for those in the usual treatment group in seven of the eight domains in the analysis over all time points; mental health was the only domain for which there were no significant between-group differences. By 5 years, the only significant difference between groups was seen in the general health perception domain of the SF-36 and favoured the LNG-IUS. No significant differences were seen between treatments with respect to the EQ-5D instrument. The treatments did not differ significantly with respect to the scores for the pleasure, discomfort and frequency domains of the Sexual Activity Questionnaire in either the short or medium term. There were no significant differences in serious adverse events between groups.

Five-year surgery-free survival rates were comparable, irrespective of initial treatments, estimated to be 80% (95% CI 74% to 84%) in the LNG-IUS group versus 77% (95% CI 71% to 82%) in the usual medical treatment group (hazard ratio 0.90, 95% CI 0.62 to 1.31; $p = 0.6$). In total, there were 24 ablations in the LNG-IUS group versus 31 in the usual treatment group, as well as 30 hysterectomies in both groups.

Economic evaluation

Using the EQ-5D, at 2 years, the relative cost-effectiveness of the LNG-IUS compared with usual medical treatment was £1600 per QALY, which by 5 years was reduced to £114 per QALY in the base-case analyses. This increase in cost-effectiveness is caused by both an increase in the mean difference in QALYs generated (of 0.023) between 2 and 5 years and a reduction in the mean difference in the costs of treatment over the same period, from a difference of £100 at 2 years to a difference of £17 at 5 years. All deterministic sensitivity analyses supported the base-case results. PSAs showed that from £2000 per QALY at 2 years, and £500 at 5 years, the LNG-IUS has a greater probability of being the more cost-effective intervention.

However, using the SF-36 to produce utility values, usual medical treatment was assessed as being less costly, by £100, and generated 0.002 more QALYs than the LNG-IUS at both 2 and 5 years, showing that usual medical treatment dominates treatment using the LNG-IUS.

Longitudinal qualitative study

The findings show that women's experiences and expectations of medical treatments for HMB vary considerably and change over time. Practitioners, in addition to considering a range of practical issues when discussing treatments with women, should consider that rate of menstrual blood flow, pain and well-being, alongside wider functioning socially or at work, may be as or more important to women as volume of menstrual blood loss experienced. Women had high expectations of a prompt effect from medical treatments. Unpredictable irregular bleeding with the LNG-IUS was more problematic than the volume of blood loss for women in this study, owing to its impact on their established behavioural coping mechanisms for HMB. Women viewed standard treatment options, such as oral tranexamic acid, as a short-term measure, and one that was more within their personal control.

Conclusions

The results of the ECLIPSE trial show that the LNG-IUS, compared with usual medical therapies, leads to greater improvement over 2 years in women's assessments of the effect of HMB on their daily routine, including work, social and family life, and psychological and physical well-being. These findings were no longer significantly different at 5 years.

The higher rate of discontinuation in the usual treatment group than in the LNG-IUS group could reflect greater symptom relief with the LNG-IUS, but another possible explanation is that discontinuation of usual medical treatment does not require consultation. Nonetheless, by 5 years, 53% of women in the LNG-IUS group had had the system removed, generally owing to lack of effectiveness or to irregular or prolonged bleeding.

In the primary care setting, treating HMB using the LNG-IUS costs more, but is also more effective, using the EQ-5D as the outcome, than usual medical treatment in both the short and medium term. As the National Institute for Health and Care Excellence (NICE) guidelines recommend interventions if the ICER is below £20,000 per QALY, the LNG-IUS would be considered cost-effective. The difference in the cost-effectiveness results derived by using the alternative measures to value QoL will have a considerable impact on cost-effectiveness decisions. The different measures did not just change the strength of cost-effectiveness of the same treatment, but the most cost-effective treatment itself changed.

Implications for health care

The results provide unique and valuable practical information for women and primary care practitioners when considering choice of, and what to expect from, medical treatments for HMB. Our clinical results support the NICE guidelines that recommend that medical treatment options should be considered when women initially present with HMB in primary care and that the LNG-IUS should be considered first. Women with HMB, a relatively normal-sized uterus on examination and no other risk factors (e.g. intermenstrual bleeding, post-coital bleeding, irregularity of the menstrual cycle which has been assessed by an endometrial biopsy) can be successfully treated and expect a significant reduction in the range of negative impacts of HMB on QoL. Although the LNG-IUS is more effective than other usual medical options within the first 2 years of treatment, both the LNG-IUS and usual medical treatments are shown to be helpful initial choices 5 years later. The low overall rates of surgery underline the importance and feasibility of initial medical management of women with HMB in primary care, and the avoidance of referral to secondary care.

The use of the LNG-IUS is cost-effective in both the short and medium term, using the method generally recommended by NICE. However, when the SF-6D is used to generate QALYs, the main base-case results are reversed and usual medical treatment is shown to be more cost-effective at both 2- and 5-year time points. Therefore, the recommendation to decision-makers differs depending on the outcome measure used.

Qualitative data indicate that clinicians in primary care should also focus on the wider life context for those presenting with HMB rather than solely on the amount of blood loss experienced by women. They support the trial evidence that if women persevere with their LNG-IUS, they will probably experience significant improvement in menstrual symptoms and related QoL. Our data suggest other treatment options, while less effective, may suit individual circumstances by remaining helpful and more within their control.

Recommendations for further research

It will be important to continue assessment of the ECLIPSE trial participants to 10 years, as we expect that over half of our cohort will have reached menopause at that stage. This can examine the clinical and health-service trajectories that women initially treated for HMB in primary care may follow in the longer term. Relevant data are lacking in these contexts. Such further research may helpfully identify the nature of continuing use of treatments, particularly the LNG-IUS, and any differential impact on frequency of surgical interventions in the longer term.

The difference in the results of the economic outcomes arising from the different instruments used to derive utilities may be explained by the limited recall periods used in these measures and their focus only on health-related QoL. These conflicting findings suggest that these measures may be capturing different aspects of QoL, which clearly has an impact on the results. Further consideration of alternative measures to assess cost-effectiveness, such as willingness to pay, which provides a broader assessment of well-being, is also needed.

Trial registration

This trial is registered as ISRCTN86566246.

Funding

Chapter 1 Introduction

Heavy menstrual bleeding

Definition and diagnosis

Heavy menstrual bleeding (HMB), also called menorrhagia, is a common problem that can significantly impact on women's lives and burden individuals and health-care systems. The National Institute for Health and Care Excellence (NICE) states that 'for clinical purposes, HMB should be defined as excessive menstrual blood loss which interferes with the woman's physical, emotional, social and material quality of life (QoL), and which can occur alone or in combination with other symptoms'.[1] In many previous studies, attempts were made to define HMB using an objective measure of blood loss, by weighing sanitary products or using the alkaline haematin method, with a loss of more than 80 ml of blood per cycle being considered excessive. Menstruation changes over a woman's reproductive lifespan, with cycle length decreasing and the duration and heaviness of the period increasing with age. Heaviness is thus subjective, defined relative to what a woman considers normal and is influenced by sociocultural and psychological well-being factors. QoL and actual menstrual blood are not closely linked,[2,3] so any interventions should aim to reduce the burden of the symptoms for the woman.

Heavy menstrual bleeding is generally considered to be heavy cyclical menstrual blood loss over a minimum of three consecutive cycles without any intermenstrual or post-coital bleeding. Where no organic pathology is present, the term dysfunctional uterine bleeding is used, whereas abnormal uterine bleeding includes that caused by uterine pathology, for example fibroids or genetic factors (e.g. von Willebrand disease).[4–6] The majority of women with HMB have no pathology that can be attributed as the cause of their symptoms, while approximately 30% will have uterine fibroids and 10% uterine polyps, with malignant causes being rare in the pre-menopausal population.[7]

With a subjective definition, diagnosis of HMB is made through careful history taking, with any suggestion of pathology such as intermenstrual or post-coital bleeding prompting further investigation, including physical examination, ultrasound and biopsy if endometrial hyperplasia or cancer is suspected. Objective measures are inconvenient and impractical, and indirect measures of menstrual blood loss, such as pictorial diaries, are not consistently accurate and these too fail to reflect women's experience of what is burdensome for them.[8] A full blood count is recommended by the NICE guidelines,[1] but serum ferritin and oestradiol measurement are not considered worthwhile. Further testing for coagulation or thyroid disorders are dependent on symptoms and personal history.

Epidemiology

In various cohorts studies, reported prevalence has ranged from 4% to 51%, although it is acknowledged that differences in definition, measurement (objective vs. subjective), clinical and cultural setting will undoubtedly influence reporting.[1] A UK survey of women aged 18–54 years indicated an annual community incidence HMB of 25% and prevalence of 52%, with a slight increase with increasing age.[9]

Heavy menstrual bleeding accounts for 19% of gynaecologist office consultations in the USA[10] and 20% in the UK,[11] and each year around 6% of women in the USA consult their general practitioner (GP).[12] Another survey found that while 22% of menstruating women over 35 years considered their periods heavy and interfering with their lives, only 7% had consulted their GP in the preceding 6 months.[13] Despite the many factors influencing women's decisions not to seek help,[14] the number and cost of consultations and treatments impose substantial demands on the NHS. In 2004–5, there were 7179 hysterectomies and 9701 endometrial ablations performed; compared with 1989–90 this represents a 69% decrease and an 11-fold increase in the number of procedures for HMB, respectively.[15] Rates of surgical procedures for HMB are 17.8 per 10,000 women aged 25–44 years in the USA,[16] and 14.3 per 10,000 women in the UK.[17]

Management of heavy menstrual bleeding

There are many treatments used for HMB, including medical therapies and surgical procedures. At the commencement of the ECLIPSE (clinical effectiveness and cost-effectiveness of levonorgestrel-releasing intrauterine system in primary care against standard treatment for menorrhagia) trial, the Royal College of Obstetricians and Gynaecologists (RCOG) guidelines[18] stated that the initial management should be usually medical, using the combined oral contraceptive (COC), tranexamic acid or mefenamic acid or the levonorgestrel-releasing intrauterine system (LNG-IUS) (Mirena®, Bayer) for women not specifically requiring contraception, but prepared to accept hormonal treatments. For those women requiring contraception, the options are COCs and the LNG-IUS.

In early 2007, NICE produced clinical guidelines for the management of HMB.[1] In these guidelines, pharmaceutical treatments were recommended as the first line of therapy, regardless of whether the women presents in primary or secondary care. Endometrial ablation may be considered as initial therapy, although only after full discussion of the risks and benefits of this and other treatments. Hysterectomy should not be offered as first-line treatment. The RCOG guidelines[18] at the time considered the evidence for the effectiveness of other drug therapies as limited, and advised that high-dose norethisterone and long-acting injectable progestogens be used only if all other medical treatments are unsuitable or unacceptable. NICE, however, included these hormonal options in the guidelines to be considered after the LNG-IUS, tranexamic acid, mefenamic acid and COCs.

Tranexamic acid

Tranexamic acid is a plasminogen activator inhibitor that exerts an antifibrinolytic effect on the endometrium, reducing menstrual blood loss by 50%.[19] It is taken in high doses, of up to 1 g four times per day, during the first 5 days of the menstrual cycle. Initial concerns regarding a potential increase in adverse thrombogenic events has been refuted by long-term studies showing no excess risk.[20]

Mefenamic acid

Mefenamic acid is a non-steroidal anti-inflammatory drug (NSAID), which can reduce the formation of prostaglandins, implicated in the abnormal clotting in the endometrium by inhibiting the enzyme cyclo-oxygenase. It is also taken on days 1–5 days of the menstrual cycle, at a dose of 500 mg three times per day, but it is contraindicated in women with bleeding disorders. In terms of reducing menstrual blood loss, mefenamic acid substantially reduces objectively measured loss, but not to the same extent as tranexamic acid.[21]

Combined oral contraceptive pill

Combined oral contraceptives of any formulation reduce the thickness of the endometrium and inhibit ovulation, with a resultant reduction in menstrual blood loss of 43%.[22] There is no evidence of any relative difference with other medical treatments. COCs are not suitable for everyone. They are contraindicated for those with a personal history or risk factors for thromboembolism, a body mass index (BMI) of $\geq 35\,kg/m^2$ or migraine with aura, and are advised to be used with caution in women over 35 years who smoke or who are overweight or obese.[23]

Norethisterone

Norethisterone is a synthetic progesterone which can reduce menstrual blood loss in ovulatory women if 10–15 mg is taken daily from day 5 to day 26 of the cycle, but is ineffective if taken only during the late luteal phase, from day 19 to day 26 of the menstrual cycle. Compared with tranexamic acid or the LNG-IUS, norethisterone is not as effective at reducing objectively measured menstrual blood loss.[21,24]

Levonorgestrel-releasing intrauterine system

The LNG-IUS is a highly effective long-acting reversal contraceptive. Since 2009 in the USA, and earlier in Europe, the LNG-IUS has been available for HMB.[23] The LNG-IUS is a T-shaped plastic frame containing 52 mg of levonorgestrel, which is released at a rate of approximately 20 µg per day and is licensed for

5 years' use, after which it should be removed and replaced if desired. Fertility is quickly restored after removal. The LNG-IUS is contraindicated in women with a congenital uterine anomaly or who have fibroids that distort the uterine cavity, in women with a uterine or pelvic infection, and in those with a history of gynaecological cancers. The adverse events of interest fall into two categories: those related to an intrauterine device, such as dysmenorrhoea, irregular bleeding, ectopic pregnancy and expulsion of the device; and those related to progestogens, such as bloating, weight gain and breast tenderness. The main disadvantage of the LNG-IUS is the disruption to the menstrual cycle, particularly in the first 6 months, with the number of bleeding and spotting days increasing to a median of 18 days and the cycle becoming irregular in approximately 20% of women, before reducing in duration and severity.[25] The proportion of women who are amenorrhoeic a year following insertion has been reported at around 12%.[26,27]

Effectiveness of the levonorgestrel-releasing intrauterine system for heavy menstrual bleeding

Clinical effectiveness of the levonorgestrel-releasing intrauterine system compared with medical treatments

Prior to the initiation of the ECLIPSE trial, the most comprehensive summary of the data for the LNG-IUS came from a systematic review which identified 10 studies of the LNG-IUS for women with heavy menstrual blood loss (≥ 80 ml per cycle).[28] Of these, only five studies were randomised controlled trials (RCTs)[20,24,29–31] and only four reported menstrual blood loss reduction in those with the LNG-IUS in place for 3–12 months, in whom the reduction in blood loss ranged from 79% to 96%.

In 2007, the NICE guidance[1] considered all pharmaceutical treatments for HMB. In comparing the LNG-IUS with non-surgical options, NICE referred to the Cochrane review,[27] which cited three small randomised trials comparing the LNG-IUS with norethisterone,[24,32] mefenamic acid[32,33] or danazol[32] (Danol®, Sanofi-aventis), although the trial reported by Cameron et al.[32] used a now discontinued formulation of the LNG-IUS. From this review, the odds ratio for amenorrhoea (> 3 months) was 8.67 [95% confidence interval (CI) 1.52 to 49.35] in favour of the LNG-IUS. The odds ratio for the proportion of women unwilling to continue with treatment was 0.27 (95% CI 0.10 to 0.67) and for the proportion of women satisfied with treatment the odds ratio was 2.13 (95% CI 0.62 to 7.33), both in favour of the LNG-IUS.

For COCs, only one small RCT of four treatments existed at the time of the NICE guidelines,[34] suggesting that COCs reduce menstrual blood loss to a greater extent than the NSAID naproxen, but less than mefenamic acid or the gonadotropin-releasing hormone antagonist danazol. However, the COC used was a first-generation formulation and no data exist for the lower oestrogen doses now available.

There was sufficient evidence in three systematic reviews for NICE to conclude that a tranexamic acid causes a clinically significant reduction in menstrual blood loss for women with HMB, ranging from 29% to 58% in studies lasting up to 1 year, although long-term data are missing.[11,21,27,35] NSAIDs (mefenamic acid or naproxen) were also found to produce clinically important reductions in menstrual blood loss, ranging from 20% to 49%, but were inferior to danazol and tranexamic acid, although NSAIDs have fewer adverse effects than danazol.[36,37]

The NICE guidelines development group stated that, in its interpretation of the evidence for pharmaceutical treatments for HMB, a high value was placed on reduction of menstrual blood loss and minimising adverse effects. The guidelines development group based its assessment first on the clinical effectiveness of treatments and, second, on the cost-effectiveness of treatments. The results of the systematic review showed that the LNG-IUS, mefenamic acid, tranexamic acid and COCs could be considered equivalent in terms of clinical effectiveness. However, the guidelines went on to state that, if both hormonal and non-hormonal treatments are acceptable, then the order in which treatment should be considered should be the LNG-IUS, then tranexamic acid, a NSAID or COC, then norethisterone on days 5–26 of the cycle.

An updated review identified seven further randomised trials,[38–44] involving only 718 women in total, which compared the LNG-IUS with non-hormonal and hormonal treatments. In an industry-sponsored review,[45] reduction in menstrual blood loss was achieved in only 230 women, from five studies,[33,38,43,44,46] in which the LNG-IUS was administered. Reported median blood loss had decreased by 85% by 3 months, and had further decreased, by a total of 94%, by 12 months, remaining constant thereafter to 5 years.

The emphasis of all the previous trials has been on the reduction in menstrual blood loss. Very few trials collected data on QoL or impact of the symptoms. A small study of the LNG-IUS compared with norethisterone found no difference in the proportion of women in whom menstrual blood loss interfered with their QoL after treatment.[24]

The trial by Lahteenmaki *et al.*[31] did report on QoL, measured using visual analogue scales (VASs) rather than any validated scale. The study selected women awaiting a hysterectomy and offered them the option of randomisation between continuing their existing treatment and the LNG-IUS. The lack of blinding and the control group remaining on an unsatisfactory treatment is likely to have influenced the women's attitude to treatment and subjective assessment of QoL, therefore, is inherently biased against existing therapy.[31]

Clinical effectiveness of the levonorgestrel-releasing intrauterine system compared with surgery

Prior to the start of the ECLIPSE trial, there were six RCTs comparing the effectiveness of the LNG-IUS with surgery, two using first-generation ablation methods,[29,30] three comparing the LNG-IUS with second-generation thermal balloon ablation[47–49] and one comparing the LNG-IUS with hysterectomy.[50] No difference in the rates of amenorrhoea or satisfaction rates were seen compared with endometrial ablation, although menstrual blood loss was reduced to a greater extent with ablation.[29,48]

In a subsequent Cochrane review of medical versus surgical treatments for HMB,[51] an additional six trials, four[49,52–54] comparing the LNG-IUS with thermal balloon ablation, one[55] comparing the LNG-IUS with hysterectomy and two[38,56] comparing the LNG-IUS with first-generation ablation, were included. There was no statistically significant difference in satisfaction rates up to 2 years or in most domains of the Short Form questionnaire-36 items (SF-36) measure of QoL. As before, endometrial ablation was significantly more effective than the LNG-IUS in controlling bleeding at 1 year (relative risk 1.19, 95% CI 1.07 to 1.32; $p = 0.001$). Hysterectomy stopped all bleeding and provided greater satisfaction, but caused serious complications for some women. This review considered the long-term data to be weak.

This review was extended by obtaining individual patient data in order to be able to standardise outcome data, such as satisfaction or, the corollary, dissatisfaction.[57] Rates of dissatisfaction with the LNG-IUS and non-hysteroscopic endometrial destruction were similar (18% vs. 23%; odds ratio 0.8, 95% CI 0.4 to 1.5; $p = 0.4$). Lack of data from individual patients prohibited any further investigation of subgroups or first-generation ablative methods. There was weak evidence to suggest hysterectomy is preferable to the LNG-IUS (dissatisfaction 5% vs. 17%; odds ratio 2.2, 95% CI 0.9 to 5.3; $p = 0.07$).

As might be anticipated, women having surgical treatment were significantly less likely to require further surgery within 1 year than women who had a LNG-IUS (relative risk 0.13, 95% CI 0.05 to 0.33; $p = 0.00002$). None of the hysterectomy group required additional surgery, while about 3% of women who had endometrial ablation required extra surgery during the first year, compared with about 15% of women who were allocated the LNG-IUS. One study[38] reported that 14% (4/29) of the surgical group underwent repeat ablation within 3 years (with two having subsequent hysterectomy), while 30% (9/30) of women allocated to the LNG-IUS group discontinued its use and 'were offered' endometrial resection. In a Finnish trial[58] of 117 women with a LNG-IUS, 60 women no longer had it in situ 5 years after randomisation; 50 of these women had had a hysterectomy (42% of those randomised to the LNG-IUS) and one had had endometrial ablation.

The NICE guidelines[1] highlighted the high level of subsequent surgery associated with the LNG-IUS and recommend that endometrial ablation may be offered as an initial treatment for HMB after full discussion with the woman of the risks and benefits and of other treatment options, while hysterectomy should not be used as a first-line treatment solely for HMB.

Cost-effectiveness of the levonorgestrel-releasing intrauterine system compared with other treatments

The first trial to incorporate a cost-effectiveness analysis was the Finnish trial cited above.[50,58] There was no statistically significant difference in QoL scores between the two treatment groups at 5 years, as measured by the European Quality of Life-5 Dimensions (EQ-5D) instrument. Mean direct costs remained significantly lower in the LNG-IUS arm (US$1892) than in the hysterectomy arm (US$2787), despite 42% of women in the LNG-IUS arm going on to have a hysterectomy. This trial, however, compared the LNG-IUS with hysterectomy in women referred to hospital. No economic analysis relevant to the use of the LNG-IUS in a UK primary care setting is available, nor had the relative cost-effectiveness of the LNG-IUS relative to medical treatment been assessed at the start of the ECLIPSE trial.

The health economic assessment performed in the process of developing the NICE guidelines populated a decision model with data predominantly from the Finnish study, and made significant assumptions regarding the effectiveness and discontinuation rates of COCs, tranexamic acid and mefenamic acid. The results showed that the LNG-IUS is the more effective treatment option when long-term use of a treatment is required, as the LNG-IUS generated more quality-adjusted life-years (QALYs), at a lower cost, than any other pharmaceutical treatment strategy. The base-case result was £840 per QALY gained.[1] The uncertainty in this result was not explored at all and so the result could be highly misleading.

Rationale for the ECLIPSE trial

An early systematic review highlighted the lack of evidence on the relative benefits of the LNG-IUS compared with medical treatment and recommended large pragmatic trials.[28] This was reiterated by the NICE guidelines.[1] The LNG-IUS is an effective, relatively safe, treatment for HMB, at least in the short term, although it is less effective than endometrial ablation. However, women with HMB often have additional concerns, which may be altered by treatment. These include the presence or absence of pain, risk of sexually transmitted disease and reproductive function. Short-term reduction in menstrual blood loss may not translate into an improvement in a woman's overall QoL or her need to seek further treatment, particularly as the LNG-IUS discontinuation rates are high.

Moreover, the consequences of HMB and its treatment extend for many years, and so a treatment that appears to be effective at 1 year may merely delay, not prevent, a definitive solution such as surgical intervention. None of the early trials measured the effect of HMB on women's lives or followed women for longer than 1 year. In relation to the LNG-IUS, there may be 'phase shifting' of the patient's journey. For example, once the device is removed, some patients' symptoms may recur, resulting in later surgery. If a woman discontinues any medical treatment, it could be because the HMB has resolved, or it could be that the treatment does not produce sustained benefit. This means that it is essential to determine and compare the long-term consequences of different treatments over a prolonged period of time appropriate to the long natural history of HMB.

It is also unclear whether or not women presenting in primary care with HMB are best treated with a LNG-IUS. The LNG-IUS is slightly more difficult to insert than standard contraceptive coils and specific training for GPs is required.[1] A LNG-IUS is occasionally associated with troublesome menstrual irregularities, especially in the first few months, and is often removed for this reason. Given the potential complications, better evidence to establish the effectiveness of the LNG-IUS as a first-line therapy is needed before it becomes widely used in primary care.

It is apparent that a RCT of the LNG-IUS should take into account a range of patient needs and preferences. These include:

- women's preferences for contraception: some desire the maintenance of fertility, some require contraception contemporaneously with relief from HMB, while some others may want to be sterilised
- long-term assessment of the treatment 'pathway', with different initial policies
- self-reported outcome measures that identify the impact of treatment on overall QoL and further treatment decisions
- the initial management of HMB in primary care
- the clinical effectiveness and cost-effectiveness, and acceptability of treatment policies.

The choice of comparator needs to reflect current practice. The treatment objective in HMB is to alleviate heavy menstrual flow and, consequently, to improve QoL. Iron-deficiency anaemia must also be prevented. Both the superseded RCOG[18] and NICE guidelines supported the use of COCs, tranexamic acid, mefenamic acid, injectable medroxyprogesterone acetate (Depo-Provera™, Pfizer Ltd) and norethisterone, with the choice dependent on the woman's preferences for hormonal or non-hormonal treatments and contraceptive needs. The comparison group was defined as usual medical treatment, with the choice specific drug being a joint decision between the GP and patient.

Objectives

Primary objective
To assess the clinical effectiveness of the LNG-IUS compared with standard medical treatment for women seeking treatment for HMB in primary care in the short term (2 years following randomisation).

Secondary objective

- To assess the clinical effectiveness of initial treatment with the LNG-IUS compared with standard medical treatment for women seeking treatment for HMB in primary care in the medium term (5 years following randomisation).
- To determine the cost-effectiveness of the LNG-IUS compared with standard medical treatment in the short and medium term.
- To explore the perspectives of trial participants or women with HMB who declined to be randomised in a longitudinal qualitative study.
- To measure the reliability and validity properties of the Menorrhagia Multi-Attribute Scale (MMAS).

Chapter 2 Outcome measures for the evaluation of treatment of heavy menstrual bleeding

Introduction

Heavy menstrual bleeding is a common gynaecological condition in the UK that has a significant impact on many women's well-being. The principal driver for treatment of HMB is women's experience of its impact on their lives. Women are mainly treated on the basis of symptoms, and therefore assessment of the condition and of the effectiveness of treatment has a large subjective component. There is poor correlation between quasi-objective measures of HMB, for example self-reports of blood loss, so a woman's subjective assessment of the perceived impact on her QoL is increasingly used to assess treatment success. As the aim of treatment is to improve women's well-being and QoL, it is necessary to have valid and reliable instruments to measure this.

A systematic review of research published before the ECLIPSE trial,[59] which assessed the quality of QoL instruments from 19 studies of HMB, was used to select the most appropriate instruments to use for the trial. Studies for the review were identified through MEDLINE, EMBASE, Science Citation Index, Social Science Citation Index, Cumulative Index to Nursing and Allied Health Literature and PsycLIT, and the references of primary and review articles. Search terms included menorrhagia, questionnaires, psychometry, psychometrics, psychological(al) tests and QoL. Studies were selected if they measured QoL in women with HMB as either an outcome measure or as part of the development of the QoL instrument itself. They were assessed using a checklist of items for clinical face validity (i.e. issues relevant to patients' expectations and concerns) and items for measurement properties (including reliability, responsiveness, criterion or construct validity and acceptability).[60]

In terms of quality, although 90% of studies complied with more than half of the criteria for measurement properties, only 37% of studies complied with more than half the criteria for face validity. Only two studies[61,62] used a condition-specific QoL instrument for HMB, which considered the most appropriate measure to assess the effect of treatment; these were instrument development studies rather than studies using the instrument as an outcome measure. The generic SF-36 was the most commonly used (63% of studies). This generic instrument has been validated in HMB[60] and seems to be reliable and responsive;[60,63] however, it is designed to assess QoL on a continuum from full health to death over the previous 4 weeks. Therefore, the SF-36 is not entirely appropriate for women with HMB because the symptoms of HMB are cyclical and the condition is distressing, but not life-threatening. This makes using the SF-36 on its own an inappropriate instrument for a patient-based outcome measure in HMB.

This review concluded that there is a need for methodologically sound condition-specific QoL instruments in HMB with clinical face validity to assess treatment outcomes. One condition-specific instrument identified as having high face validity is the MMAS, developed by Shaw et al.[61] This was developed in collaboration with women with HMB using a multi-attribute utility method.[64] It was decided that this scale should be the principal outcome measure for the ECLIPSE trial. However, although the MMAS was designed to have good face and content validity, other psychometric properties of this instrument, such as reliability, convergent and discriminant validity, and floor and ceiling effects, could not be evaluated in the original study. One study reported a small advantage for the MMAS over the SF-36 in predicting management outcome for HMB,[65] but otherwise the measurement properties of the questionnaire when used in a treatment or observational study had not been reported before. These properties of the MMAS were assessed at the beginning of the ECLIPSE trial.

The disadvantage of the MMAS is that, as a condition-specific measure, it is anchored by full health and the worst possible state for the condition, rather than full health and death, as in generic measures. This is important in the economic evaluation of treatments, which is recommended by NICE.[66] Since economic evaluation is used by decision-making bodies such as NICE to help with decisions about resource allocation, it is important in the context of treatment trials. NICE endorses the use of generic instruments in economic evaluation of treatments in order to calculate the QALY.[66] The QALY reflects changes in both the quantity of life and the QoL, where quality is measured on a scale from 0 (death) to 1 (full health).

In order to undertake an economic evaluation in the ECLIPSE trial, it was therefore necessary to include generic measures in addition to the condition-specific MMAS. In the systematic review cited above,[59] the SF-36 health survey questionnaire was the most commonly used generic instrument. However, NICE recommend the use of the EQ-5D,[66] which can more straightforwardly be converted to QALYs. Clark *et al.*[59] suggest that, in order to ensure that a trial has an appropriate set of outcome measures, a condition-specific instrument could be used in combination with a generic instrument. It was therefore decided to combine the MMAS with a generic instrument. There was better evidence for the use of the SF-36 in HMB, but given the recommendation from NICE to use the EQ-5D and its relative brevity it was decided to use both generic instruments in the ECLIPSE trial. In addition to the QoL instruments, it was decided to include a measure of sexual well-being. Clinical experience has shown that sexual activity is an important dimension of women's lives which may be specifically affected by HMB. The outcome measures used are described below.[61]

Outcome measures used in the ECLIPSE trial

Menorrhagia Multi-Attribute Scale
The MMAS questionnaire captures the subjective consequences of HMB on six domains: practical difficulties, social life, psychological well-being, physical health, work routine and family life. Each of the six domains has four statements which represent four levels of response. Respondents indicate the statement which best matches their feelings for each domain. The statement scores derive from a weighting of the domains and a weighting of the statements in level of severity by women in the origin study. Scores range from 0 (worst possible state in all domains) to 100 (best possible state in all domains).

Medical outcomes study: Short Form questionnaire-36 items
The SF-36 version 2 is a 36-item short-form survey that measures general health-related QoL. The SF-36 is a practical and reliable way to obtain important health outcomes data in a variety of settings, measuring eight domains of health: physical functioning, role limitations owing to physical health, bodily pain, general health perceptions, vitality, social functioning, role limitations owing to emotional problems and mental health. Respondents are asked to recall how they have felt over the previous 4 weeks. It is commonly used in studies of HMB.[65]

European Quality of Life-5 Dimensions
The EQ-5D is a standardised instrument for use as a measure of health outcome and is widely used in economic evaluations of medical interventions. Applicable to a wide range of health conditions and treatments, it provides a simple descriptive profile and also a single-index value for health status measured on a VAS.

The Sexual Activity Questionnaire
The Sexual Activity Questionnaire (SAQ) was developed as a self-report questionnaire, for use in gynaecological clinical trials, which would be quick to complete and acceptable to the majority of women.[67] Three dimensions of perceptions of sexual activity are measured: pleasure, discomfort and habit.

Measurement of reliability and validity properties of the menorrhagia multi-attribute quality-of-life scale

Method

Participants
Data from the first 431 women enrolled in the ECLIPSE trial were analysed for this study.

Procedure
Participants completed the MMAS, SF-36 version 2, EQ-5D and the SAQ as part of the baseline measures for the trial before randomisation, but before treatment. Fifty women were also invited to participate in a test–retest study of the MMAS and completed the questionnaires again within 2 weeks of the baseline measurement, but before treatment commenced.

Psychometric analysis
The measurement properties of the MMAS were assessed using traditional psychometric procedures. Estimates of reliability and validity were made using intra- and inter-test correlations, with additional comparison of means for the subscales of the questionnaire. Floor and ceiling effects were also assessed using frequency data on high and low scores.

Results
The mean age of the participants was 41.25 years [standard deviation (SD) 5.3 years]. The ethnic mix of the sample was representative of the patient population for the area and the trial as a whole, with 93% of women categorising themselves as from the two main ethnic groups, white British (83%) and South Asian (10%).

Reliability
Of the 50 women asked to take the questionnaire again within 2 weeks, 49 returned completed questionnaires. The test–retest correlation for the total MMAS score was acceptable ($r = 0.836$), suggesting that it is stable over time. Cronbach's alpha was calculated as a measure of internal consistency of the six items of the MMAS. The Cronbach's alpha value was 0.82 and lies within the range 0.7 and 0.9, suggesting that the items are all measuring aspects of the same construct.[68] Evidence of the homogeneity of the MMAS was also provided by the item total correlations (the correlation between each item and the total score minus that item) which range between 0.44 and 0.68. Only one such correlation fell below 0.6, that for 'practical difficulties' caused by HMB.

Validity
The MMAS has already been shown to have good face validity, so this analysis concentrated on construct validity, that is the extent to which the instrument measures what it is intended to measure. An important method of demonstrating construct validity is by showing that an instrument correlates with other instruments measuring similar constructs, for example generic QoL scales (convergent validity), and does not correlate with measures of different constructs, for example age and BMI (discriminant validity).

Convergent validity
To show convergent validity, correlations between test scores should be moderate to high, that is 0.4 or higher. *Table 1* shows the correlations between the MMAS total score, the subscales of the SF-36 and the EQ-5D. These show moderate correlations for most components of the SF-36, but weaker correlations for physical functioning, general health perception and for the EQ-5D standardised score.

The individual domain scores of the MMAS were compared with the SF-36 subscales. With the exception of the practical difficulties domain on the MMAS, which did not significantly correlate with any SF-36 subscale, the highest correlations were between each of the MMAS domains and the SF-36 social functioning scale. The correlation between the psychological health scales of the two instruments is 0.39,

TABLE 1 Correlations between the MMAS total score, the EQ-5D summary and the SF-36 subscales

Questionnaire and domains	MMAS total score	
	r	*n*
SF-36 subscales		
Physical functioning	0.33	420
Physical role limitations	0.45	422
Emotional role limitations	0.49	421
Social functioning	0.54	427
Mental health	0.40	426
Energy/vitality	0.42	426
Pain	0.38	428
General health	0.32	419
EQ-5D summary score	0.32	428

and that between the MMAS relationships scale and the SF-36 emotional role limitations is 0.47. However, these correlations of subscales should be interpreted with caution as there are only 4 points on the MMAS subscales. In order to investigate these relationships further, respondents were divided into four groups for each MMAS domain and their scores on the SF-36 were compared using analysis of variance. The results for most subscales and domains were highly significant for each comparison, with high scores on the MMAS corresponding to high scores on the SF-36. Again, the exception was for comparisons in the practical difficulties domain. Here, three comparisons were not significant: mental health ($F_{3,422} = 1.66$; $p = 0.175$), pain ($F_{3,424} = 0.62$; $p = 0.60$) and general health ($F_{3,415} = 2.21$; $p = 0.09$).

There was a small non-significant difference in the total MMAS scores of women reporting that they were not engaging in sexual activity on the SAQ and those who were (38.86 and 41.49, respectively). It may be that the MMAS relationship domain reflects a particular problem with sexual relationships. To test this, comparisons were made between the scores on the pleasure, discomfort and habit scales of the SAQ for women scoring at the four different levels of the relationship domain. This revealed a statistically significant effect of level on the discomfort scale ($F_{3,321} = 4.05$; $p = 0.008$), but non-significant effects of pleasure ($F_{3,321} = 1.71$; $p = 0.165$) and habit ($F_{3,326} = 2.54$; $p = 0.057$).

Discriminant validity
The MMAS total score would not be expected to correlate with age or BMI. As expected, these correlations are very low for age ($r = 0.086$) and for BMI ($r = -0.057$). As well as providing evidence for the validity of the MMAS, it suggests that it can be used with all age groups.

Acceptability
All items in the MMAS were completed by all participants, indicating that the instrument was acceptable to them.

Floor and ceiling effects
In psychometric instruments, floor and ceiling effects have implications for the precision of the instrument and its responsiveness to change, as they reduce the likelihood that the instrument will measure further improvement or deterioration. In this sample, five (1.2%) women's scores were the lowest possible and three (0.7%) were the highest possible. The overall distribution of scores was slightly positively skewed (skew = 0.41, standard error = 0.12), which is what might be expected in a sample of women just about to embark on treatment for HMB. These statistics indicate that there are no floor or ceiling effects.

Discussion

This study investigated further the psychometric properties of the MMAS. It has shown that, in addition to having high face validity, the MMAS has good convergent and discriminant validity, and test–retest reliability. It also has high internal consistency, with a Cronbach's alpha score of 0.82, which is further evidence of reliability and indicates that all the items are measuring different aspects of the same construct.

The overall MMAS score showed moderate correlation with most of the subscales of the SF-36. Lower correlations were found for physical functioning and general health perception. Rather than being a negative feature, these low correlations may reflect the inappropriateness of the SF-36 as the sole measure of QoL in women with HMB. The physical functioning subscale is heavily weighted to mobility and self-care, not likely to be affected by HMB. The general health subscale and the EQ-5D, which was also relatively poorly correlated with the MMAS, may not be precise enough to measure differences in this patient group, especially when women are asked to consider their general condition rather than that pertaining to their menstrual cycle.

Conversely, scores on the practical difficulties subscale of the MMAS were not related to other QoL measures. This domain of the MMAS is very specific to HMB, relating to sanitary protection and flooding and may pick up specific important issues not measured by generic scales.

All women completed the whole of MMAS, which suggests that it is acceptable to women and seen to be relevant. It could be argued that this sample was more highly motivated than women with HMB more generally because they had already agreed to take part in the ECLIPSE trial. However, there were missing data in other questionnaires administered at the same time. These results suggest that the MMAS has good measurement properties and is, therefore, an appropriate condition-specific instrument to measure the outcome of treatment for HMB.

General discussion

In this chapter we have discussed some of the challenges in assessing the outcome of treatment trials for HMB. We have demonstrated that the MMAS is a reliable and valid instrument to measure the impact of HMB, and its treatment, on women. However, although the MMAS appears to be the most suitable measure, owing to its condition-specific nature it cannot produce QALYs and is therefore not appropriate for economic evaluation.

A more recent systematic review of the QoL instruments used in HMB by Sanghera *et al.*[69] found no consensus on the most appropriate economic measure to use when valuing outcomes in HMB, with both the SF-36 and EQ-5D lacking face validity. Evidence suggests that women do not consider HMB to be solely a health-related condition, so purely generic health-related QoL measures are unsuitable. This is supported by the analysis, which showed that the practical difficulties question in the MMAS did not correlate with any of the components of the health-related generic instruments.

A different approach that Sanghera *et al.*[69] explored in their review is willingness to pay (WTP). WTP considers a broader range of QoL than that related to health. Many studies have successfully elicited WTP values in other disease areas,[70] so perhaps WTP could prove to be the most suitable economic measure in HMB in future trials. However, even at the time of this later review there was insufficient evidence for its use, with only one study reviewed assessing the reliability of WTP.[69]

Conclusions

At the beginning of the ECLIPSE trial the evidence suggested that while the primary outcome of the trial, that is, the impact of treatment on women with HMB, should be measured by a condition-specific instrument, economic evaluation of the trial required generic instruments. In addition, clinical experience suggested that sexual well-being was an aspect of women's lives particularly affect by HMB, so it was decided to investigate this in more detail using the SAQ. Consequently, four self-report instruments were chosen to measure the outcomes of the ECLIPSE trial, MMAS as the primary outcome measure with the SF-36 version 2, the EQ-5D (both the profile and VAS scores) and the SAQ as secondary outcome measures.

Chapter 3 Methods and results of the randomised controlled trial up to 2 years' follow-up

Objectives

The study rationale is presented in *Chapter 1*. The objectives of the RCT presented in this chapter were to assess the clinical effectiveness of the LNG-IUS compared with standard medical treatment for women seeking treatment for HMB in primary care, with up to 2 years' follow-up.

Methods

Population

Women between 25 and 50 years of age who presented to their GP with HMB involving at least three consecutive menstrual cycles were eligible to participate. Women were excluded if they intended to become pregnant over the next 5 years, were taking hormone replacement therapy or tamoxifen (Soltamox™, Rosemount Pharmaceuticals), had intermenstrual bleeding (between expected periods) or post-coital bleeding or findings suggestive of fibroids (abdominally palpable uterus equivalent in size to that at 10–12 weeks' gestation) or other disorders, or had contraindications to or a preference for either the LNG-IUS or usual medical treatments. Women with heavy, irregular bleeding were ineligible unless the results of endometrial biopsy were reported to be normal; no further investigations were mandated by the protocol. All patients provided written informed consent.

Randomisation

Patients were assigned to a study group by telephone or a web-based central randomisation service at the University of Birmingham Clinical Trials Unit. A computerised, minimised randomisation procedure was used to achieve balance between the groups with respect to age (< 35 years or ≥ 35 years), BMI (≤ 25 kg/m^2 or > 25 kg/m^2), duration of symptoms (< 1 year or ≥ 1 year), need for contraception (yes or no) and heavy HMB alone or HMB accompanied by menstrual pain. GP practice was not included as a minimisation variable to avoid any chance of the allocation becoming too predictable.

Study interventions and compliance

Eligible women who provided written informed consent were randomly assigned to either the LNG-IUS or usual medical treatment. Usual treatment options included mefenamic acid, tranexamic acid, norethisterone, a combined oestrogen–progestogen or progesterone-only oral contraceptive pill (any formulation), or medroxyprogesterone acetate injection and were chosen by the physician and patient on the basis of contraceptive needs or the desire to avoid hormonal treatment.[1,18] The particular medical treatment to be used was specified before randomisation. Subsequently, treatments could be changed (from one medical treatment to another, from the LNG-IUS to medical treatment, or from medical treatment to the LNG-IUS) or could be discontinued because of a perceived lack of benefit, side effects, a change in the need for contraception, referral for endometrial ablation or hysterectomy, or other reasons, according to usual practice.[1,18] Treatment changes reported by patients were confirmed with the GP.

Outcome measures and follow-up

The primary outcome measure was the condition-specific MMAS,[61,71] which is designed to measure the effect of HMB on six domains of daily life (practical difficulties, social life, psychological health, physical health, work and daily routine, and family life and relationships). Summary scores, which range from 0 (severely affected) to 100 (not affected), were assessed at 6 months, 1 year and 2 years after randomisation. Details on how summary scores are calculated can be found elsewhere.[61] The MMAS has a

high degree of reliability and internal consistency,[71] has good content and construct validity,[61,71] is responsive[65,72] and is acceptable to respondents.[61,65,71,72] Secondary outcome measures included general health-related QoL and sexual activity. To assess QoL, we used three instruments: the SF-36 version 2 [with scores ranging from 0 (severely affected) to 100 (not affected)], the EQ-5D [with scores ranging from −0.59 (health state worse than death) to 100 (perfect health state)] and the EQ-5D VAS [with scores ranging from 0 (worst health state imaginable) to 100 (most perfect health state imaginable)]. The validated SAQ measures pleasure [with scores ranging from 0 (lowest level) to 18 (highest level)], discomfort [with scores ranging from 0 (greatest) to 6 (none)] and frequency (assessed relative to perceived usual activity as an ordinal response).[73] Scores were obtained before randomisation and by mail at 6 months, 1 year and 2 years after randomisation. Data were collected from participating clinicians regarding all serious adverse events, defined as adverse events that resulted in death, disability or hospitalisation. Patients were also asked to report any hospitalisations and adverse events leading to discontinuation of the study drug.

Study oversight

The study sponsor was the University of Birmingham. Study oversight was provided by an independent steering committee and an independent data and safety monitoring committee, whose three reviews of interim data provided no reason to modify the trial protocol on the basis of pragmatic stopping criteria,[74] other than to support revision of the recruitment target, which is described in *Statistical considerations*. The study was conducted in accordance with the protocol, which is available at www.nets.nihr.ac.uk/ projects/hta/020602. Approval of the study was obtained from the South-West England Multicentre Research Ethics Committee, and clinical trial authorisation was received from the Medicines and Healthcare products Regulatory Authority. All medications and devices were prescribed by providers through the NHS. The manufacturers of the LNG-IUS and other therapeutic agents used in the study were not involved in any aspect of the trial.

Statistical analysis

The study was originally designed to have 95% power at $p < 0.01$ to detect small to moderate (0.3 SD) differences in mean MMAS score between groups at any one time point.[75] Following interim review of recruitment rates, in consultation with the independent data monitoring committee, the power and type 1 error parameters were revised to more conventional levels of 90% and $p = 0.05$, to allow timely completion of the study. This required an enrolment of 470 patients; we increased the sample size to 570 to allow for up to 20% loss to follow-up.

Analyses was by intention to treat with continuous measures compared using multilevel repeated-measures models,[76] including parameters allowing for participant, treatment, time and baseline score. Responses including all three time points were considered in the primary analysis. Differences between slopes were assessed by including treatment-by-time interaction parameters in the models. If these were not significant ($p > 0.05$), then they were dropped from models and constant treatment differences over time were assumed. In a similar fashion, treatment-by-subgroup interaction parameters were included in the above model to test for the importance of the prespecified subgroup parameters (same as the minimisation variables, see *Randomisation*). Treatment effect estimates within subgroups are presented where interaction parameters were statistically important. Differences between groups at each assessment time point were also examined by analysis of covariance (adjusting for baseline score). Changes from baseline score within groups were examined using paired *t*-tests.

A number of sensitivity analyses were performed on the primary outcome to test the robustness of the results.[77,78] Analyses were performed excluding data received subsequent to a crossover to the alternative treatment group and also once study treatment was ceased altogether. Further sensitivity analyses included assigning the best possible score (100) to those women who indicated that they were no longer experiencing bleeding and did not complete the MMAS, and a carried forward (pre-intervention) score for those undergoing hysterectomy or endometrial ablation. Further analyses included the exclusion of MMAS data from late returned booklets (defined as more than 3 months past the due date) and an analysis assuming all missing responses were the worst possible score (0). Missing MMAS responses were also simulated using a multiple imputation approach with analysis performed on the resulting 20 simulated data sets.[78]

Kaplan–Meier plots were constructed for time to first treatment change with women censored at date of last follow-up or, if appropriate, date of death, withdrawal or loss to follow-up. A Cox proportional hazards model was used to construct hazard ratios. Standard statistical methods were used to test the statistical significance of other responses (chi-squared tests for dichotomous data, Cochrane–Armitage test for trend and Wilcoxon signed-rank test for ordinal data). Effect sizes are presented with 95% CIs and two-sided p-values for all. SAS version 9.2 (SAS Institute Inc., Cary, NC, USA) was used for analyses.

Results

Patients and follow-up

Between February 2005 and July 2009, a total of 571 women with HMB from 63 UK centres (median centre recruitment was three, interquartile range one to seven) were randomly assigned to either the LNG-IUS (285 women) or usual medical treatment (286 women). Baseline characteristics were similar between the two treatment groups (*Table 2*).

TABLE 2 Baseline characteristics of the patients

Characteristics	Usual medical treatment ($n = 286$)	LNG-IUS ($n = 285$)
Age (years)		
≥ 35[a]	255 (89%)	257 (90%)
Mean (SD)	41.8 (5.5)	42.1 (5.0)
BMI (kg/m^2)		
≥ 25[a]	200 (70%)	200 (70%)
Mean (SD)	29.3 (6.7)	29.1 (6.1)
Ethnic group[b]		
White	246 (86%)	225 (79%)
Asian	23 (8%)	28 (10%)
Black	12 (4%)	18 (6%)
Mixed	4 (1%)	9 (3%)
Other	1 (< 1%)	4 (1%)
Duration of HMB ≥ 1 year[a]	229 (80%)	231 (81%)
Presence of menstrual pain[a]	211 (74%)	213 (75%)
Contraceptive requirement[a]	55 (19%)	55 (19%)
Copper or non-hormonal coil in place	10 (3%)	9 (3%)

a This characteristic was a minimisation variable and was assessed in the predefined subgroup analyses.
b Self-identified ethnic grouping. One 'not given' response in the LNG-IUS group.

For 215 (75%) of the women assigned to usual medical treatment, the initial prescription, decided prior to randomisation, was for mefenamic acid, tranexamic acid or a combination of the two drugs (*Table 3*); 55 (19%) of the women in the usual treatment group required contraception. Study questionnaire booklets were returned by 478 (84%) of the patients at the 2-year time point (see *Figure 1*);[38] of these women, 45 (9%) could not complete the MMAS appropriately because their menstrual bleeding had ceased, but they completed other parts of the booklet and this information was used to inform the sensitivity analysis.

Of the 285 women randomly assigned to the LNG-IUS, 24 (8%) did not have the intrauterine system (IUS) inserted: 10 chose usual medical treatment, six chose no treatment and, in eight, insertion of the system was unsuccessful and, therefore, usual medical treatment was subsequently instituted (*Figure 1*).

Women in the LNG-IUS group were almost twice as likely as those in the usual treatment group to still be receiving their assigned treatment at 2 years (64% vs. 38%; $p < 0.001$) (*Figure 2*). The most common reasons cited for discontinuation of the LNG-IUS were lack of effectiveness (37%) and irregular or prolonged bleeding (28%). Of the 163 women who discontinued usual medical treatment, 80 (49%) switched to the LNG-IUS. The most common reason for discontinuation of usual medical therapy was lack of effectiveness (53%). Reasons for discontinuing therapy are summarised in *Appendices 1* and *2*.

Primary outcome: Menorrhagia Multi-Attribute Scale
In both groups, total scores on the MMAS were significantly improved, compared with baseline scores, at 6 months, and this improvement was maintained at 1 and 2 years (*Figure 3* and *Table 4*) However, improvements in scores were significantly greater among women assigned to the LNG-IUS than among those assigned to usual treatment over the course of the 2 years (mean difference in scores 13.4 points, 95% CI 9.9 to 16.9 points; $p < 0.001$). This difference was apparent by 6 months and maintained thereafter.

TABLE 3 Intended prescriptions if allocated to the usual medical treatment group

Intended prescriptions	Usual medical treatment ($n = 286$)	LNG-IUS ($n = 285$)
Mefenamic acid and tranexamic acid	134 (47%)	123 (43%)
Tranexamic acid alone	50 (17%)	66 (23%)
Mefenamic acid alone	31 (11%)	23 (8%)
High-dose norethisterone alone	20 (7%)	23 (8%)
COC alone	16 (6%)	22 (8%)
Medroxyprogesterone acetate injection only	14 (5%)	1 (< 1%)
Mefenamic acid, tranexamic acid and COC	9 (3%)	14 (5%)
Other combination[a,b]	12 (4%)	13 (5%)

a Usual medical treatment arm: mefenamic acid/tranexamic acid/desogestrel ($n = 2$); mefenamic acid/COC ($n = 2$); tranexamic acid/COC ($n = 2$); tranexamic acid/norethisterone ($n = 2$); mefenamic acid/tranexamic acid/COC/ medroxyprogesterone acetate injection ($n = 1$); mefenamic acid/tranexamic acid/medroxyprogesterone acetate injection/ norethisterone ($n = 1$); mefenamic acid/tranexamic acid/COC/norethisterone ($n = 1$); and medroxyprogesterone acetate injection/norethisterone ($n = 1$).
b LNG-IUS arm: mefenamic acid/tranexamic acid/desogestrel ($n = 2$); tranexamic acid/COC ($n = 2$); tranexamic acid/ norethisterone ($n = 2$); mefenamic acid/tranexamic acid/norethisterone ($n = 2$); mefenamic acid/COC ($n = 1$); mefenamic acid/tranexamic acid/COC/medroxyprogesterone acetate injection ($n = 1$); mefenamic acid/tranexamic acid/ medroxyprogesterone acetate injection ($n = 1$); tranexamic acid/desogestrel ($n = 1$); and norethisterone/ desogestrel ($n = 1$).
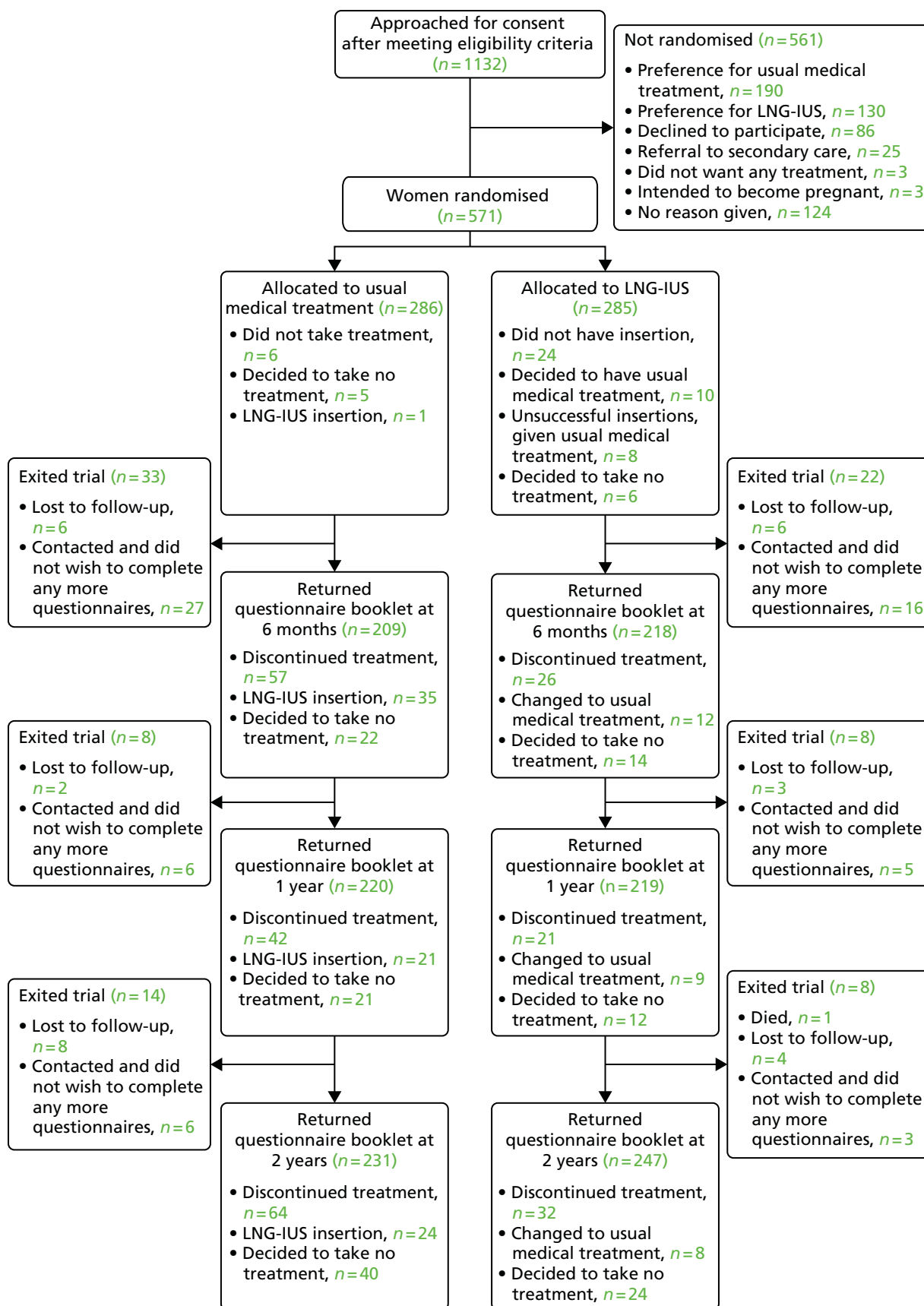
FIGURE 1 Enrolment, randomisation and follow-up to 2 years of the study patients. See *Appendices 1* and *2* for reasons for discontinuation of treatment.
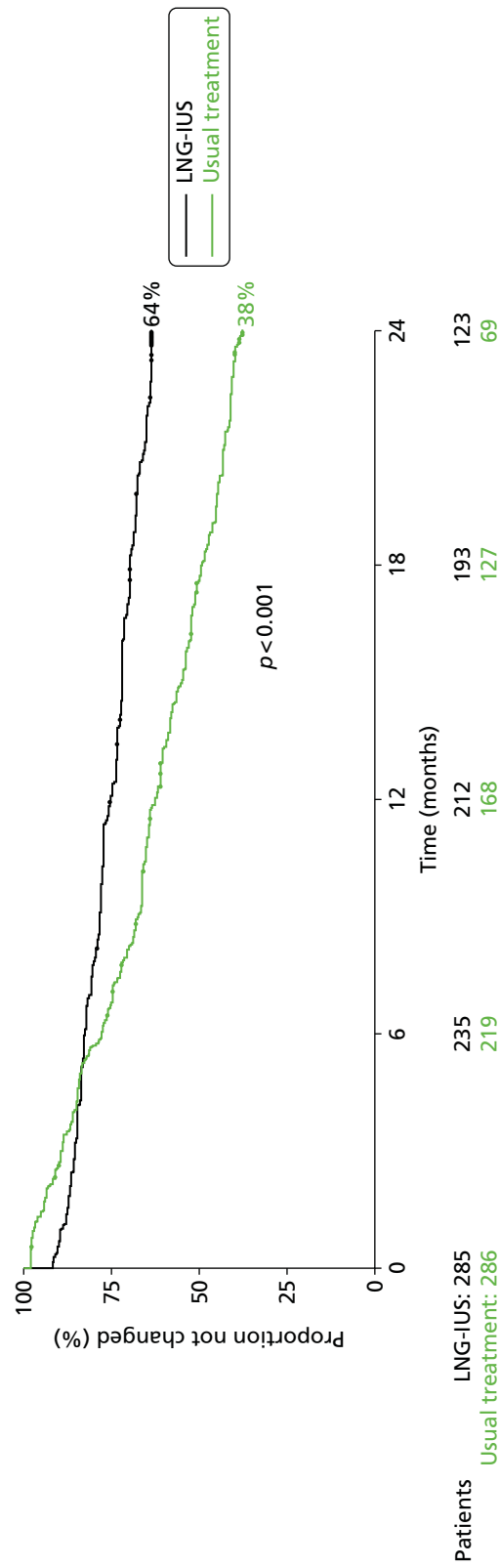
FIGURE 2 Time to first treatment change during 2 years' follow-up. Data are for women who crossed over from the assigned study treatment to the other study treatment and for those who discontinued treatment.
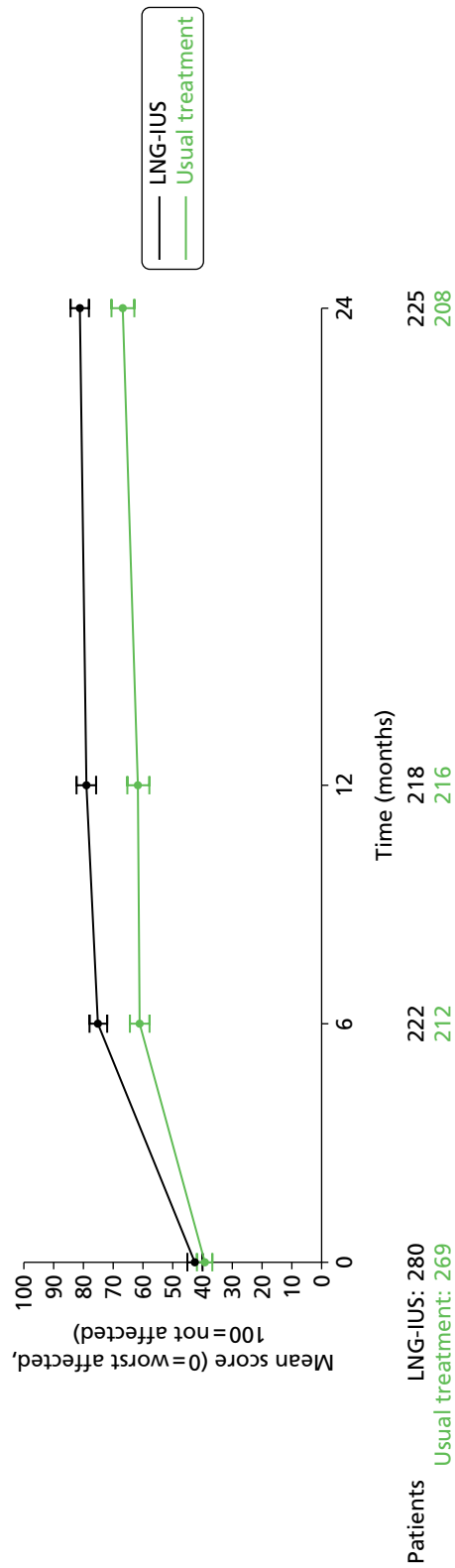
FIGURE 3 Primary outcome in the two treatment groups up to 2 years' follow-up. The primary outcome was the score on the MMAS (scores range from 0 to 100, with lower scores indicating greater severity). Mean MMAS scores are shown for the two groups at 6, 12, and 24 months. Error bars indicate 95% CIs. Overall, the average difference in scores between the women treated with the LNG-IUS and those treated with the usual medical therapy was 13.4 points (95% CI 9.9 to 16.9 points; *p* < 0.001).

TABLE 4 Results of primary outcome: MMAS summary scores up to 2 years' follow-up

Treatment group and comparisons	Baseline	6 months	12 months	2 years	Overall[a]
Usual medical treatment, mean (SD, n)	39.2 (21.3, 269)	61.0 (25.1, 212)	61.5 (26.3, 216)	66.8 (28.5, 208)	–
Change within group (95% CI)	–	21.4 (18.1 to 24.7; p < 0.001)	21.1 (17.5 to 24.6; p < 0.001)	26.8 (22.8 to 30.8; p < 0.001)	–
LNG-IUS, mean (SD, n)	42.5 (20.5, 280)	74.9 (22.5, 222)	78.8 (25.0, 218)	81.0 (23.2, 225)	–
Change within group (95% CI)	–	32.7 (29.3 to 36.0; p < 0.001)	35.0 (31.2 to 38.7; p < 0.001)	39.0 (35.4 to 42.6; p < 0.001)	–
Difference between groups (95% CI)	**–**	**12.8 (8.6 to 17.0; p < 0.001)**	**16.0 (11.4 to 20.7; p < 0.001)**	**13.3 (8.6 to 18.1; p < 0.001)**	**13.4 (9.9 to 16.9; p < 0.001)**

a Based on a repeated-measures model. See *Chapter 3, Statistical analysis* for details.
Estimates > 0 from differences between groups favour LNG-IUS. Results at each time point adjusted for baseline score. Change within group compared with baseline.
0 = worst affected, 100 = not affected.

The effect in favour of the LNG-IUS was apparent in all six individual domains of the MMAS at every time point through examination of the frequencies of responses (*Table 5*).

In a sensitivity analysis that excluded women who crossed over from the assigned treatment to the other study treatments, improvement with the LNG-IUS, as compared with usual medical treatment, increased (mean difference in scores over the course of 2 years, 17.8 points, 95% CI 14.1 to 21.5 points; $p < 0.001$). Other sensitivity analyses yielded results that were not materially different from the results of the primary analysis ($p < 0.001$ for all comparisons) (see *Appendix 3*).

In subgroup analyses, there was a significant interaction between treatment and BMI ($p = 0.004$). The benefit of the LNG-IUS was greater in women with a BMI > 25 kg/m^2 (16.7 MMAS points, 95% CI 12.6 to 20.9 MMAS points; $p < 0.001$) than in those with a BMI of $\geq$ 25 kg/m^2 (5.4 MMAS points, 95% CI −1.0 to 11.8 MMAS points; $p = 0.10$). This finding appeared to be attributable to the superior outcome with usual medical treatment in leaner women (see *Appendix 4*). Improvements with the LNG-IUS were similar in both subgroups. None of the other tests for subgroup interaction was significant ($p > 0.10$).

Generic quality of life and sexual activity

The SF-36 domains were generally significantly improved from baseline in both groups at all time points, although the scores for women in the LNG-IUS group were better than for those in the usual treatment group in seven of the eight domains in the analyses over all time points (*Table 6*); mental health was the only domain for which there were no significant between-group differences. The improvements appeared to be greatest at 6 months, but had lessened by the 2-year follow-up assessment (see *Appendix 5*). No significant differences were seen between treatments with respect to the EQ-5D instrument; scores were significantly improved from baseline in both groups at 2 years, but not at earlier assessments (see *Appendix 6*). The treatments did not differ significantly with respect to the scores for the pleasure, discomfort and frequency domains of the SAQ (see *Appendix 7*).

Adverse events

There were no serious adverse reactions attributable to study treatments. There was no significant difference between the two groups in the frequency of unrelated serious adverse events (58 in the usual treatment group and 49 in the LNG-IUS group; $p = 0.59$). There was one death in the LNG-IUS group; the cause of death was recorded by the coroner as inconclusive and the LNG-IUS was not in situ. Unrelated serious adverse events are summarised in *Appendix 8*.

Surgical interventions

The frequency of surgical interventions for HMB within 2 years did not differ significantly between the two groups. Hysterectomy was performed in 6% of the women in each group; endometrial ablations were performed in 4% of women in the LNG-IUS group and in 6% of those in the usual treatment group ($p = 0.44$).

TABLE 5 Results of primary outcome: MMAS individual domain response up to 2 years' follow-up

Domain and response	Baseline		6 months		12 months		2 years	
	Usual medical treatment (n = 269)	LNG-IUS (n = 280)	Usual medical treatment (n = 212)	LNG-IUS (n = 222)	Usual medical treatment (n = 216)	LNG-IUS (n = 218)	Usual medical treatment (n = 208)	LNG-IUS (n = 225)
Practical difficulties								
1. None	6 (2%)	7 (3%)	48 (23%)	131 (59%)	61 (28%)	142 (65%)	82 (39%)	152 (68%)
2. Extra sanitary protection only	86 (32%)	100 (36%)	115 (54%)	71 (32%)	100 (46%)	50 (23%)	71 (34%)	54 (24%)
3. Extra sanitary protection and clothes	92 (34%)	92 (33%)	29 (14%)	9 (4%)	30 (14%)	14 (6%)	35 (17%)	10 (4%)
4. Severe problems	85 (32%)	81 (29%)	20 (9%)	11 (5%)	25 (12%)	12 (6%)	20 (10%)	9 (4%)
Social life during cycle								
1. Unaffected	16 (6%)	26 (9%)	57 (27%)	126 (57%)	68 (31%)	145 (67%)	86 (41%)	157 (70%)
2. Slightly affected	122 (45%)	139 (50%)	106 (50%)	78 (35%)	98 (45%)	53 (24%)	77 (37%)	51 (23%)
3. Limited	106 (39%)	92 (33%)	42 (20%)	16 (7%)	38 (18%)	11 (5%)	35 (17%)	14 (6%)
4. Devastated	25 (9%)	23 (8%)	7 (3%)	2 (1%)	12 (6%)	9 (4%)	10 (5%)	3 (1%)
Psychological health during cycle								
1. No worries	23 (9%)	29 (10%)	56 (26%)	99 (45%)	64 (30%)	127 (58%)	85 (41%)	133 (59%)
2. Some anxiety and worry	119 (44%)	129 (46%)	107 (50%)	99 (45%)	105 (49%)	62 (28%)	79 (38%)	66 (29%)
3. Often feel down and worry	100 (37%)	100 (36%)	36 (17%)	19 (9%)	37 (17%)	20 (9%)	35 (17%)	18 (8%)
4. Depressed and cannot cope	27 (10%)	22 (8%)	13 (6%)	5 (2%)	10 (5%)	9 (4%)	9 (4%)	8 (4%)

Domain and response	Baseline		6 months		12 months		2 years	
	Usual medical treatment (n = 269)	LNG-IUS (n = 280)	Usual medical treatment (n = 212)	LNG-IUS (n = 222)	Usual medical treatment (n = 216)	LNG-IUS (n = 218)	Usual medical treatment (n = 208)	LNG-IUS (n = 225)
Physical health and well-being during cycle								
1. Well and relaxed	7 (3%)	11 (4%)	48 (23%)	75 (34%)	50 (23%)	101 (46%)	76 (37%)	113 (50%)
2. Well and relaxed most of the time	55 (20%)	55 (20%)	72 (34%)	84 (38%)	75 (35%)	73 (33%)	57 (27%)	75 (33%)
3. Often tired and not especially well	165 (61%)	178 (64%)	84 (40%)	57 (26%)	79 (37%)	36 (17%)	67 (32%)	29 (13%)
4. Very tired and not well at all	42 (16%)	36 (13%)	8 (4%)	6 (3%)	12 (6%)	8 (4%)	8 (4%)	8 (4%)
Work/daily routine during cycle								
1. No disruptions	22 (8%)	19 (7%)	57 (27%)	121 (55%)	67 (31%)	140 (64%)	81 (39%)	146 (65%)
2. Occasional disruptions	92 (34%)	127 (45%)	107 (50%)	80 (36%)	90 (42%)	54 (25%)	79 (38%)	55 (24%)
3. Frequent disruptions	125 (46%)	104 (37%)	39 (18%)	16 (7%)	50 (23%)	17 (8%)	41 (20%)	16 (7%)
4. Severe disruptions	30 (11%)	30 (11%)	9 (4%)	5 (2%)	9 (4%)	7 (3%)	7 (3%)	8 (4%)
Family life/relationships during cycle								
1. Unaffected	33 (12%)	41 (15%)	61 (29%)	95 (43%)	65 (30%)	117 (54%)	84 (40%)	139 (62%)
2. Suffer some strain	120 (45%)	140 (50%)	105 (50%)	108 (49%)	104 (48%)	83 (38%)	84 (40%)	67 (30%)
3. Suffers quite a lot	83 (31%)	88 (31%)	38 (18%)	15 (7%)	37 (17%)	14 (6%)	35 (17%)	16 (7%)
4. Severely disrupted	33 (12%)	11 (4%)	8 (4%)	4 (2%)	10 (5%)	4 (2%)	5 (2%)	3 (1%)

TABLE 6 Scores on the QoL questionnaires and the SAQ at baseline, and mean between-group difference over 2 years

Questionnaire and domain	Baseline score (SD)		Difference between groups over 2 years (95% CI)[a]	p-value
	Usual medical treatment	LNG-IUS		
SF-36[b]				
Physical functioning	77.8 (24.7)	80.0 (20.4)	2.7 (0.0 to 5.4)	0.05
Physical role	68.9 (26.2)	72.1 (24.7)	5.9 (2.6 to 9.1)	< 0.001
Emotional role	69.8 (26.8)	71.9 (25.1)	4.6 (1.3 to 8.0)	0.007
Social functioning	62.4 (25.9)	64.3 (24.5)	5.1 (2.0 to 8.1)	0.001
Mental health	59.0 (19.8)	60.3 (19.3)	1.5 (–1.0 to 3.9)	0.23
Energy/vitality	40.8 (21.7)	40.7 (20.9)	5.3 (2.5 to 8.2)	< 0.001
Pain	49.5 (24.9)	54.2 (24.9)	7.8 (4.5 to 11.0)[c]	< 0.001
General health perception	60.3 (21.9)	61.8 (21.4)	2.9 (0.3 to 5.4)	0.03
EQ-5D				
EQ-5D descriptive system[d]	0.714 (0.276)	0.756 (0.243)	0.013 (–0.016 to 0.042)	0.38
EQ-5D VAS[e]	69.7 (19.8)	70.3 (19.1)	2.0 (–0.5 to 4.6)[c]	0.12
SAQ[f]				
Pleasure	10.9 (4.9)	10.8 (4.9)	0.4 (–0.3 to 1.1)	0.26
Discomfort	4.62 (1.69)	4.65 (1.48)	–0.07 (–0.30 to 0.16)	0.55

a Estimated values > 0 favour LNG-IUS.
b The SF-36 has scores in each of the eight domains that range from 0 (severely affected) to 100 (not affected).
c Some evidence here of time by treatment effect ($p \leq 0.05$). Refer to *Appendix 5* for full details of estimates at each separate time point.
d The EQ-5D descriptive system has scores that range from –0.59 (state of health worse than death) to 1.00 (perfect state of health).
e The EQ-5D VAS ranges from 0 (worst health state imaginable) to 100 (most perfect health state imaginable).
f The SAQ has scores for pleasure that range from 0 (lowest level) to 18 (highest level), and scores for discomfort that range from 0 (greatest) to 6 (none).

Discussion

The results of this trial show that, compared with usual medical therapies for HMB, the LNG-IUS leads to greater improvement in women's assessments of the effect of HMB on their daily routine, including work, social and family life, and psychological and physical well-being.

At baseline, the women were substantially affected by HMB, as assessed with the use of condition-specific (MMAS) and general (SF-36) health-related scales. The scores improved significantly over a period of 2 years in both the LNG-IUS group and the usual treatment group. However, improvements in average MMAS scores were greater in the LNG-IUS group than in the usual medical treatment group over the first 2 years, by an average of 13.4 points, in an intention-to-treat analysis. The greater improvement in the LNG-IUS group than in the usual treatment group was both statistically significant and clinically meaningful. The between-group difference was more than 0.5 SD, which is the minimum clinically important difference identified in a systematic review of studies reporting such data for health-related QoL measures.[79] A 13.4-point difference represents a change in two or three MMAS domains: from being substantially to minimally affected by HMB (e.g. from frequent to occasional disruptions of work and daily routine) or from being minimally affected to being unaffected (e.g. from experiencing some strain in family

life to experiencing no strain in family life). The between-group difference reported here is also greater than that reported in an observational study comparing women who did and those who did not undergo surgery for HMB.[65] The differences seen in favour of the LNG-IUS on the SF-36 domains were generally highly statistically significant, but were on average small in magnitude (SD 0.2–0.3).[79] It is therefore debatable whether or not they are clinically important.

The strengths of our randomised trial include its size (larger than prior trials of treatments for HMB), the multicentre design, the inclusion of patients ethnically representative of the UK population, the relatively low rates of loss to follow-up, and the assessment of outcomes over a period of 2 years rather than 6 or 12 months, as in previous studies.[28,45] In addition, previous trials have focused on the reduction of menstrual blood loss, which does not reflect the full effect of HMB on women's lives.[28,45] In contrast, our primary outcome measure was the patient-reported, psychometrically valid, condition-specific MMAS, which better reflects women's personal experience of the burden of HMB. Interference with the QoL, rather than perceptions of HMB itself, appears to be the primary factor in women's decision to seek treatment.[80]

Some limitations of our study should be noted. The range of options available for medical treatment complicates any efforts to compare the LNG-IUS with individual agents. However, the choice among the various agents is representative of current clinical practice. In addition, substantial numbers of patients switched treatments over the course of the study; however, these crossovers would be expected to result in an underestimation of the benefits that might be achieved with perfect compliance. A range of sensitivity analyses did not change the conclusions. Although the interventions studied in this trial represent options available in primary care settings in the UK, insertion of intrauterine devices is not part of primary care in all health-care settings, and in some circumstances it requires consultation with a gynaecologist.

The 21.4-point improvement from baseline in the average MMAS score at 6 months in the usual treatment group, which was sustained throughout the 2 years of follow-up, was not explained by a switch in treatment, since similar improvements were noted when crossovers to the LNG-IUS were excluded from the analyses. The higher rate of discontinuation in the usual treatment group than in the LNG-IUS group could reflect greater symptom relief with the LNG-IUS, but another possible explanation is that discontinuation of usual medical treatment does not require consultation. Nonetheless, at 2 years, 36% of women in the LNG-IUS group had had the system removed, generally because of lack of effectiveness or irregular or prolonged bleeding, which are well-recognised reasons for discontinuing the LNG-IUS.[81,82] This proportion is consistent with the proportions of women who discontinued the LNG-IUS treatment in smaller trials that compared it with hysterectomy[50] (31% of 117 women at 12 months) or with endometrial ablation (28% of 105 women at 2 years).[30,49,81] In subgroup analyses, the LNG-IUS appeared to be less beneficial in women with a BMI of 25 kg/m^2 or less than in those with a BMI of more than 25 kg/m^2, an observation that was explained by an apparently greater efficacy of usual medical treatments in the leaner women. This analysis was one of several subgroup analyses and should be interpreted with caution, since the findings may be explained by chance and require confirmation.

We expected fewer surgical interventions in the LNG-IUS group, but rates were similarly low in the two groups. This finding may reflect the eligibility criteria for the trial, since women who had fibroids or other disorders were excluded.

Finally, given the long natural history of HMB, study outcomes need to be assessed over a period that is longer than 2 years; additional intention-to-treat analyses are reported for 5 years in *Chapter 4* and planned for 10 years.

In conclusion, our study showed that both the LNG-IUS and usual medical treatments reduced the adverse effect of HMB on women's lives over the course of 2 years, but the LNG-IUS was the more effective first choice, as assessed by the impact of bleeding on the women's QoL.

Chapter 4 Methods and results of randomised controlled trial at 5 years' follow-up

Objectives

The objectives of the RCT presented in this chapter were to assess the clinical effectiveness of the LNG-IUS compared with standard medical treatment for women seeking treatment for HMB in primary care in the medium to long term (5 years' follow-up).

Methods

Trial methods

For details on the population, randomisation and outcome measures the reader should refer to *Chapter 3*.

Statistical considerations

The sample size assumptions for MMAS scores were the same as previously detailed in *Chapter 3*, *Statistical consideration*. Four hundred and seventy patients were required to have 90% power ($p = 0.05$) to detect 0.3 SD in MMAS score. We ultimately received 424 responses at 5 years' follow-up. A post-hoc calculation suggests this many responses would provide 87% power ($p = 0.05$) to detect this same size of difference. For progression to surgical intervention (hysterectomy or endometrial ablation), using an assumed rate of 35% in the standard arm (a figure that was set out in the protocol), 424 women (the number followed up to 5 years; see *Figure 4* for details) would provide 80% power ($p = 0.05$) to detect an absolute reduction of 12%, that is, 35% down to 23%.

The statistical methods used to analyse the data were similar to those set out in *Chapter 3*; however, given the long length of time between this and the previous assessment point at 2 years, we felt a repeated-measures analysis incorporating all accumulated responses to no longer be clinically meaningful and therefore this has not been included.

Results

Patients and follow-up

See *Chapter 2* for details on the demographics and prescribed treatments of those randomised and *Figure 4* for losses to follow-up. Study questionnaire booklets were returned by 424 (74%) of the participants. One hundred and fifteen of the women (27%) declined to complete the MMAS, but indicated on the form that they were no longer bleeding; their score was assumed to be the maximum achievable (100, 'no problems'). These women completed other parts of the questionnaire booklet and a sensitivity analysis of MMAS responses without any assumption was also performed.
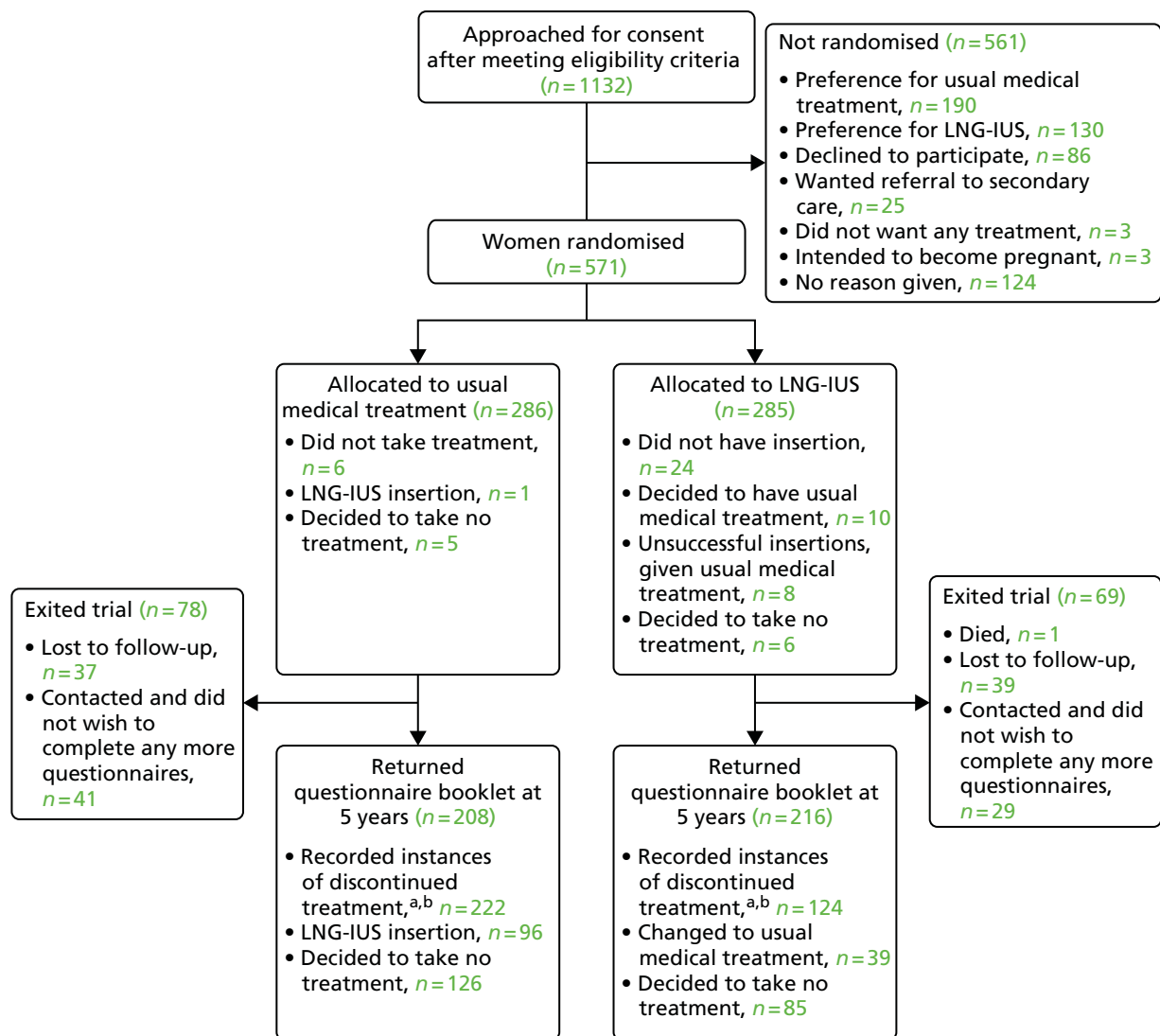
FIGURE 4 Enrolment, randomisation and follow-up up to 5 years of the study patients. a, See *Figure 5* for time to first treatment change survival estimates; b, see *Appendices 9* and *10* for reason reported for discontinuing treatment.

The 5-year allocated treatment retention rate was estimated to be 47% (95% CI 40% to 52%) in the LNG-IUS group and 15% (95% CI 11% to 20%) in the usual treatment group (*Figure 5*). Reported reasons for discontinuation of treatment are summarised in *Appendices 9* and *10*. Lack of efficacy was cited 41% (94/228) of the time in the usual treatment group and 24% (36/148) of the time in the LNG-IUS group. Of the 228 recorded instances of treatment change in women allocated usual medical treatment, 97 (43%) were to the LNG-IUS. In the LNG-IUS group, 39% (57/148) of treatment switches were to usual treatment. There was no significant difference between the groups in the total number of serious adverse events up to 5 years (see *Appendix 11*).

Menorrhagia Multi-Attribute Scale

After 5 years, scores on the MMAS were significantly improved compared with baseline in both groups (*Table 7*). This improvement was greater, on average, among women assigned to the LNG-IUS, but the difference was not statistically significant (3.9 points, 95% CI –0.6 to 8.3 points; $p = 0.09$). The same analysis without any assumption about MMAS scores where a blank form was returned and the woman indicated she was no longer bleeding returned a similar result (5.2 points difference in favour of the LNG-IUS, 95% CI –0.4 to 10.8 points; $p = 0.07$). Per-protocol analysis (excluding women who switched or stopped taking treatment) – limited to 28 responses in the usual treatment group and 88 in the LNG-IUS group – returned a 15.8-point difference in favour of the LNG-IUS (95% CI 7.4 to 24.2 points; $p = 0.0003$).
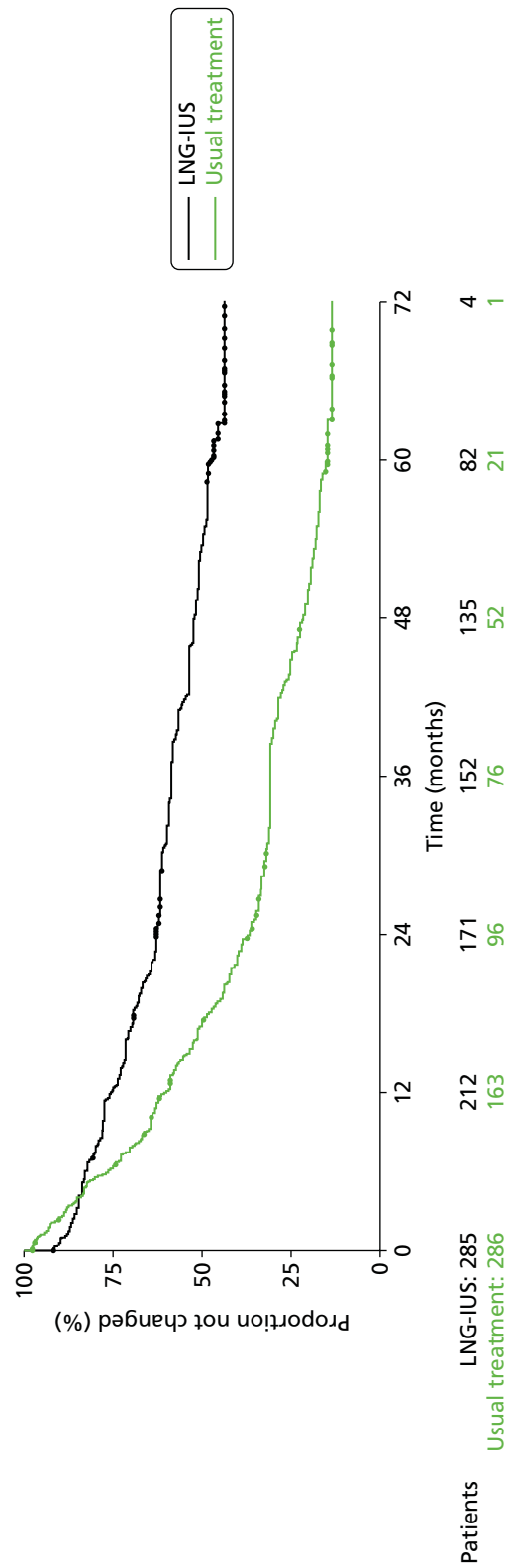
FIGURE 5 Time to first treatment change over 5 years' follow-up.

TABLE 7 Menorrhagia Multi-Attribute Scale summary scores at baseline and 5 years' follow-up

Treatment group and comparisons	Baseline score	5-year score
Usual medical treatment (SD, n)	39.2 (21.3, 269)	83.1 (24.4, 208)
Change within group (95% CI)	–	43.4 (39.1 to 47.7; $p < 0.0001$)
LNG-IUS (SD, n)	42.5 (20.5, 280)	87.1 (22.1, 216)
Change within group (95% CI)	–	44.9 (4.10 to 48.8; $p < 0.0001$)
Difference between groups (95% CI)	**–**	**3.9 (–0.6 to 8.3; $p = 0.09$)**

Surgical interventions

There were 53 surgical events (endometrial ablation or hysterectomy) in the LNG-IUS group, compared with 56 in the usual treatment group, that were included in the surgery-free survival analysis (109 events in total). This difference was not statistically significant (hazard ratio 0.90, 95% CI 0.62 to 1.31; $p = 0.6$) (Figure 6). Analysis excluding participants who crossed over from one group to another returned a similar result (HR 0.96, 95% CI 0.60 to 1.52; $p = 0.9$). Five-year surgery-free survival rates were estimated to be 80% (95% CI 74% to 84%) in the LNG-IUS group compared with 77% (95% CI 71% to 82%) in the usual treatment group. In total, there were 24 ablations in the LNG-IUS group compared with 31 in the usual treatment group as well as 30 hysterectomies in both groups (115 operations in total; the extra six events coming from patients who had both types of surgery).

Generic quality of life and sexual activity

Responses to the EQ-5D and SF-36 instruments were generally significantly improved from baseline in both groups (Table 8); the only statistically significant difference between groups was seen in the general health perception domain of the SF-36 and favoured the LNG-IUS (4.7 points, 95% CI 0.6 to 8.8 points; $p = 0.02$). The treatment groups did not differ significantly with respect to any of the domains of the SAQ.

Discussion

In Chapter 3, analysis of trial data showed significant benefit over 2 years after women sought and were treated for HMB in primary care, from both usual medical treatments (tranexamic acid, mefenamic acid, COCs, high-dose progestogens and medroxyprogesterone acetate or a combination of these methods) and the LNG-IUS. However, the superiority of the LNG-IUS in improving women's general QoL and reducing the effect of their HMB on their work, social and family lives was highlighted. This research remains the world's largest long-term multicentre RCT of non-surgical treatments for HMB. A high follow-up rate of 84% at 2 years has been sustained at 5 years at 74% (424/571 randomised patients) by using systematic methods of ongoing communication and feedback with participants.

Our results are encouraging in showing that the significant improvement in both HMB-specific and generic QoL was sustained 5 years after commencing either usual medical treatments or the LNG-IUS. However, the greater clinical efficacy of the LNG-IUS compared with usual medical treatment seen at 2 years has now diminished. Although MMAS scores and the majority of QoL domains at 5 years still favoured the LNG-IUS, these were no longer statistically significant, and contrast with the relatively large differences we noted at the 2-year follow-up. Even if we consider our finding for MMAS to be of borderline statistical significance (the p-value was 0.09), the difference between group means was estimated to be 3.9 points on average, which is less than 0.2 SD and unlikely to be clinically meaningful. This lack of persistence of
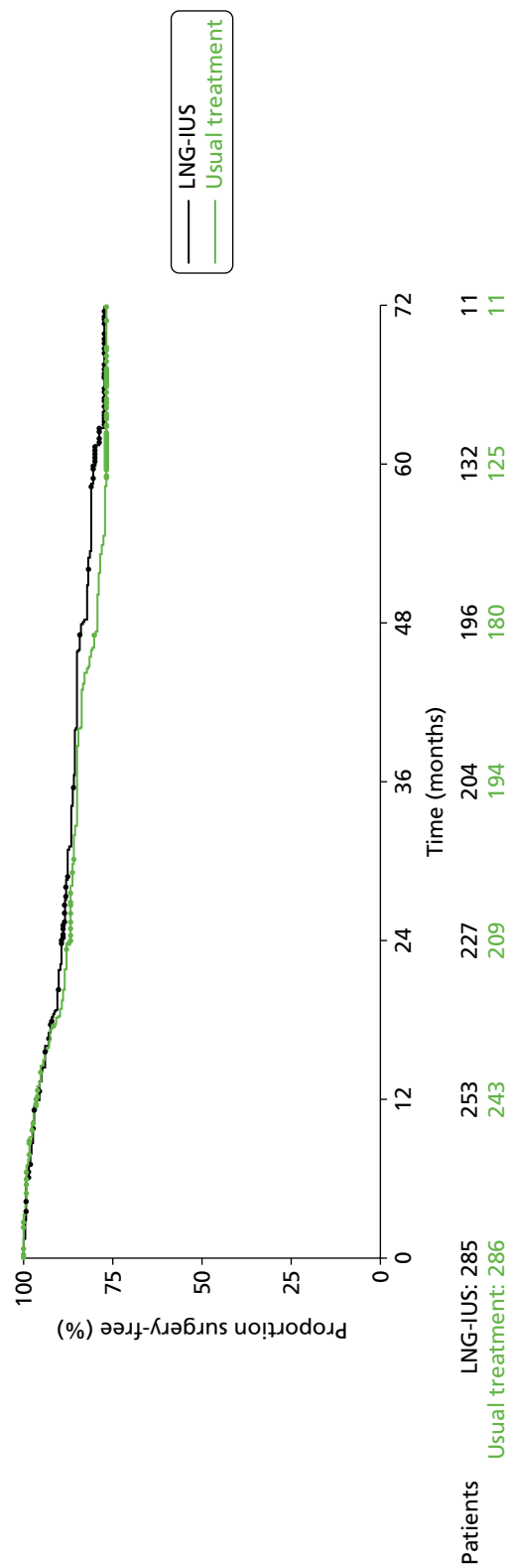
FIGURE 6 Surgery-free (hysterectomy/endometrial ablation) survival analysis over 5 years' follow-up.

TABLE 8 Scores on the QoL questionnaire and SAQ at baseline and 5 years' follow-up

Questionnaire and domain	Baseline score (SD, n) Usual medical treatment	LNG-IUS	5-year score (SD, n) Usual medical treatment	LNG-IUS	Difference between groups over 5 years (95% CI, p-value)[a]	Change within group (95% CI, p-value) Usual medical treatment	LNG-IUS
SF-36[b]							
Physical functioning	77.8 (24.7, 264)	80.0 (20.4, 272)	83.0 (26.3, 208)	85.4 (23.5, 216)	1.6 (−2.7 to 5.9; p = 0.5)	5.6 (1.8 to 9.5; p = 0.004)	5.9 (2.8 to 9.1; p = 0.0003)
Physical role	68.9 (26.2, 264)	72.1 (24.7, 276)	80.6 (27.2, 208)	83.9 (25.5, 217)	2.7 (−2.1 to 7.5; p = 0.3)	12.6 (7.9 to 17.2; p < 0.0001)	13.3 (9.5 to 17.1; p < 0.0001)
Emotional role	69.8 (26.8, 264)	71.9 (25.1, 276)	83.3 (25.3, 208)	81.4 (26.9, 217)	−2.0 (−6.8 to 2.9; p = 0.4)	13.1 (8.7 to 17.6; p < 0.0001)	10.9 (6.7 to 15.1; p < 0.0001)
Social functioning	62.4 (25.9, 268)	64.3 (24.5, 277)	76.9 (25.5, 207)	79.5 (25.1, 214)	2.2 (−2.5 to 6.9; p = 0.4)	14.0 (9.4 to 18.5; p < 0.0001)	15.6 (11.6 to 19.6; p < 0.0001)
Mental health	59.0 (19.8, 267)	60.3 (19.3, 277)	72.6 (19.4, 206)	71.4 (20.5, 214)	−1.6 (−5.2 to 2.0; p = 0.4)	13.6 (10.5 to 16.7; p < 0.0001)	11.4 (8.5 to 14.3; p < 0.0001)
Energy/vitality	40.8 (21.7, 268)	40.7 (20.9, 277)	56.6 (23.1, 206)	59.2 (21.0, 213)	2.8 (−1.2 to 6.9; p = 0.2)	15.7 (11.9, 19.4; p < 0.0001)	18.6 (15.2 to 22.0; p < 0.0001)
Pain	49.5 (24.9, 268)	54.2 (24.9, 278)	71.3 (28.6, 207)	76.1 (24.8, 214)	3.7 (−1.3 to 8.7; p = 0.1)	21.4 (16.9 to 25.8; p < 0.0001)	22.7 (18.6 to 26.9; p < 0.0001)
General health perception	60.3 (21.9, 264)	61.8 (21.4, 274)	61.9 (23.8, 207)	67.8 (22.1, 214)	4.7 (0.6 to 8.8; p = 0.02)	2.0 (−1.5 to 5.6; p = 0.26)	5.1 (2.0 to 8.3; p = 0.002)
EQ-5D							
EQ-5D descriptive system[c]	0.71 (0.28, 269)	0.76 (0.24, 277)	0.82 (0.25, 207)	0.81 (0.25, 214)	−0.02 (−0.06 to 0.02; p = 0.4)	0.10 (0.06, 0.14; p < 0.0001)	0.06 (0.02, 0.09; p = 0.003)
EQ-5D VAS[d]	69.7 (19.8, 246)	70.3 (19.1, 250)	76.1 (18.8, 198)	77.3 (19.9, 207)	0.6 (−3.2 to 4.5; p = 0.8)	6.3 (2.9 to 9.8; p = 0.0004)	6.5 (3.0 to 9.9; p = 0.0003)
SAQ[e]							
Pleasure	10.9 (4.9, 199)	10.8 (4.9, 210)	11.8 (4.6, 100)	11.2 (4.5, 111)	−0.4 (−1.7 to 0.9; p = 0.6)	0.6 (−0.6 to 1.7; p = 0.4)	0.1 (−1.0 to 1.2; p = 0.9)
Discomfort	4.6 (1.7, 201)	4.7 (1.5, 209)	4.8 (1.6, 101)	4.8 (1.6, 111)	0.0 (−0.4 to 0.4; p = 0.9)	0.1 (−0.2 to 0.4; p = 0.7)	0.1 (−0.3 to 0.4; p = 0.7)

a Estimated values > 0 favour the LNG-IUS.
b The SF-36 has scores in each of the eight domains that range from 0 (severely affected) to 100 (not affected).
c The EQ-5D descriptive system has scores that range from −0.59 (state of health worse than death) to 1.00 (perfect state of health).
d The EQ-5D VAS ranges from 0 (worst health state imaginable) to 100 (most perfect health state imaginable).
e The SAQ has scores for pleasure that range from 0 (lowest level) to 18 (highest level), and scores for discomfort that range from 0 (greatest) to 6 (none).
Results at each time point adjusted for baseline score. Change within group compared with baseline.

difference between trial treatment groups may be unsurprising given the high proportions of women who, by this 5-year stage, had changed to a treatment that worked for them or had ceased bleeding – either naturally or through surgical intervention (retention rates were 15% with usual medical therapy and 47% in the LNG-IUS group).

Similar reasons may explain why our 5-year data provided no evidence of any reduction in surgical interventions with the LNG-IUS, compared with usual medical therapy (even when we discounted the significant number of treatment crossovers), despite the large differences observed in symptom-related QoL at 2 years post randomisation. We recognise this could also possibly be because of a lack of power to detect a difference, although we had sufficient power to detect a relative reduction of approximately 34%. At 2 years post randomisation we reported that surgical interventions were similarly low at about 10% in both medical treatment groups and this has increased to approximately 20% in both treatment groups at 5 years. However, this rate remains considerably lower than the high surgical intervention rate identified in a previous Cochrane review of trials of women comparing oral medical therapy to surgical interventions. This showed 58% of women with medical treatment subsequently undergoing endometrial ablation or hysterectomy by 2 years.[51] Our low surgical intervention rates may possibly be explained by the exclusion of women with known uterine disease such as fibroids and enlarged uteri and who were deemed suitable for treatment in a community setting.

We are aware of few trials in this field reporting outcomes at 5 years.[38,58] The current data have allowed an accurate estimate of the 5-year retention rate of the LNG-IUS – at around half of those initially allocated to this treatment at randomisation. In addition, they offer a good estimate of the proportions of women progressing to surgical intervention after 5 years from initial presentation with HMB to their GP. Our results at 5 years provide unique and valuable practical information for women and clinicians considering what to expect from these medical treatments for HMB over the medium term. Moreover, the low surgical intervention rates observed may underline the importance and feasibility of emphasising initial medical management of women with HMB in primary care and avoiding referral to secondary care.

Our results suggests that women who are found to have a relatively normal-sized uterus with HMB and no other risk factors (i.e. intermenstrual bleeding, post-coital bleeding, irregularity of the menstrual cycle that has been assessed by an endometrial biopsy) can be successfully treated in the primary care setting with significant benefit from medical treatment in reducing the impact of HMB on their QoL that is sustained over 5 years. We intend to follow up patients to 10 years. This will be important as we expect that over half of our cohort will have reached menopause at that stage (mean age at randomisation was 42 years). With most women's menopause occurring at approximately 51–52 years, we thus expect to determine the final rates of surgical intervention in the majority of participants.

Chapter 5 Women's experiences of medical treatments for heavy menstrual bleeding: a longitudinal qualitative study

Introduction

Heavy menstrual bleeding and other changes in menstrual experience may affect many aspects of women's health, well-being and social functioning.[83] Although there is existing evidence comparing the effectiveness of various treatments in reducing menstrual blood loss, we know less about women's experiences or expectations of medical treatments. The ECLIPSE trial, therefore, provided an opportunity to generate qualitative insights in this context by exploring the perspectives of trial participants or women with HMB who declined to be randomised, but who nevertheless commenced medical treatment in primary care. A longitudinal qualitative study was undertaken to enable the ECLIPSE trial outcomes to be further contextualised in ways that may enhance the utility of the quantitative results and their application in clinical practice.

Women have discussed the management of menstruation including its concealment in order to conform to societal norms,[84] their reluctance to request treatment[85,86] and their experience of a change in menstruation requiring further explanation. Women with psychological distress appear more likely to report heavy periods[13,87] and the effect of menstrual bleeding on mood may influence a woman's ability to continue life as normal.[88] Similarly, other menstrual symptoms such as pain, mood changes and irregular bleeding may influence perceptions of the severity of HMB.[2,89]

Women who seek treatment for HMB can have predetermined expectations and preferences for a treatment influenced by diverse factors, such as employment status as well as severity of symptoms.[90,91] They may also reflect contraceptive preferences, attitudes to hormonal treatments, 'the pill' or 'coil' and the presence of other symptoms such as painful periods.

Clinical guidelines for HMB from NICE, issued towards the end of the current ECLIPSE trial recruitment,[1] introduced the option of the LNG-IUS on the basis of limited evidence to date, with other usual, and mainly oral, medical treatments as further choices. Women's expectations about treatments have obvious implications in discussing these options and concordance with treatment. There was also a need to better understand the impact of medical treatments on HMB-related QoL from a qualitative perspective,[1,92] alongside the trial's quantitative measures of the QoL. This study sought to explore women's experiences and expectations of medical treatments used in the ECLIPSE trial for HMB.

Methods

Sampling
A purposeful sample was selected from women approached for inclusion in the ECLIPSE trial of the effectiveness of medical treatments for HMB in primary care.[93] They had consulted their GP about HMB (lasting three or more consecutive cycles). Women participating in the trial had either agreed to be randomised to the LNG-IUS or to other standard treatment (including oral tranexamic acid, mefenamic acid, norethisterone, contraceptive pill or medroxyprogesterone acetate injection) or had declined to enter the trial because they had a strong preference for a particular medical treatment, which they had then commenced. We selected women from eight general practices across five counties of the UK of varying demographics (age, social and educational background, ethnicity) and parity, who experienced different medical treatments for HMB of differing treatment duration.

Recruitment and consent

Practitioners identified potential participants at the time that interest in participation in the ECLIPSE trial was discussed, and trial participation accepted or declined. Permission was sought to pass contact details to the research team, who sent study information and a request for written consent, translated where appropriate, to participants.

Data generation and analysis

Semistructured face-to-face interviews were conducted by a female researcher (G Prileszky), with a midwifery background, at respondents' convenience and in their own homes. The support of an interpreter was offered where appropriate (four cases), but all interviewees opted to be interviewed in English. Initial interviews were conducted within 1–3 months of commencing medical treatments. A second interview was conducted, if participants were willing, between 12 and 14 months following commencement of treatment to explore any changes in experience or treatment, or in views occurring over time.

Interviews were digitally audio-taped and transcribed verbatim, and contemporaneous field notes kept of the interview context. Verbatim quotations are annotated with participant number and suffixed with 'A' for the initial interview and 'B' for the follow-up interview. They followed broad topic areas including menstrual history and experience, social factors influencing experience, relationships and family life, and emotional well-being as well as treatment preference and experience of treatments. The topic prompts were modified and refined following early interviews, with participants encouraged to discuss their perceptions and experiences freely. A sample of participants was also asked to participate in further respondent validation.

Data were analysed using constant comparison[94] by the field researcher, with two researchers of different disciplinary backgrounds (academic midwifery and clinical primary care) contributing to development of the analysis. Coding was aided by application of NVivo software (QSR International, Warrington, UK) in identifying emerging themes from the data. Data generation and analysis were iterative, each informing the other, with further theoretical sampling and data collection to extend and challenge earlier data and interpretation, both within and between first and second interview data sets. This tested the integrity and credibility of the analysis until no new themes emerged, which suggested saturation.

Validation

Interim findings were also fed back and reviewed with a sample of six interviewees who were willing to be involved a third time. This was by audio-taped telephone interview following prior distribution of written summaries. Respondents were asked to consider and comment on the findings, enabling the research team to confirm and further refine data interpretation and analysis where appropriate.

Findings

Consent to contact forms were received from 52 women over an 18-month period of recruitment, of whom 27 were purposefully selected and interviewed within 3 months of commencing treatment for HMB. Of these women, 24 were interviewed again more than 12 months after commencing treatment. The 51 interviews lasted between 30 and 85 minutes. Characteristics of women are summarised in *Table 9*.

TABLE 9 Characteristics of women interviewed in the qualitative study

Characteristics	Number of women ($n = 27$)
Age range (years)	
20–29	2
30–39	9
40–49	15
> 50	1
Ethnicity	
White/British	23
Asian/Indian	2
White/mixed	1
Black/Caribbean	1
Parity	
No children	4
One child	5
Two or more children	18
Highest formal educational attainment	
No qualification	6
GCSE or equivalent	9
NVQ3/A-level or equivalent	7
Degree or higher	5
Experience of HMB	
≤ 1 year	9
1–5 years	6
> 5 years	12
Initial medical treatment for HMB	
Trial participants	17
LNG-IUS	8
Tranexamic acid	5
Medroxyprogesterone	2
Oral contraceptive pill	1
Norethisterone	1
Declined trial participation	10
LNG-IUS	6
Tranexamic acid	4
Changed treatment for HMB during study period	
No change	15
Changed once	10
Changed twice	2

A-level; advanced level; GCSE, General Certificate of Secondary Education; NVQ, National Vocational Qualification.

Concealment of menstruation

Framing women's experiences of treatments, women felt a strong societal pressure to remain silent about menstruation, which remained taboo and was amplified by experiencing HMB.

> *I think in society it's quite a taboo subject. I work in an all-female environment with one man who is obviously married, who has got children but if we were actually talking about periods we would stop if he came into the room.*
>
> *11B*

> *It makes me feel not so much like a normal woman really when I have a heavy period because you feel like you're the only one that's having a heavy period. Obviously there are hundreds of women out there that are having them but it's not common knowledge that a lot of women are having heavy periods like I did, I kept it quiet.*
>
> *15A*

The taboo existed within families and among friends. It was strongest in the workplace and could persist within consultation with a health professional.

> *A lot of women don't like to go [to the GP] because it's this taboo subject, so it has to start somewhere other than a GP. It's got to start somewhere that you know that it is acceptable to have a problem, there isn't a right and a wrong.*
>
> *27B*

Women could maintain the societal norm by denying their problems, lying to employers and withholding information from their doctors in order to stay silent about HMB. Not speaking about HMB (and menstruation in general) had a negative impact on most women's QoL.

> *I had gone to the doctors a few times and somebody had asked me and I wished somebody had asked me earlier and perhaps if you talk about it then somebody perhaps at work might actually well I am doing this, this and this and then you are more confident then to seek help.*
>
> *26B*

Women also expressed a strong sense that women should 'put up' with HMB and not complain about it by maintaining a stoic silence. The decision to seek treatment and therefore break the silence was seen as a weakness or failure and was often associated with reduced self-esteem.

> *And you just feel a bit, not worthless because worthless is the wrong word – but you think some people must think you are one of these women. That you have just got a heavy period – get on and sort of deal with it – but you can't physically get on unless you have actually been through a phase of having heavy periods – you can't. And I am the same because I didn't understand when people said they had really heavy periods because I never had them before.*
>
> *26B*

Expectations of treatments

Many women interviewed had high expectations of a prompt positive outcome from their treatments, both those prescribed standard treatments and those prescribed the LNG-IUS. Women recounted only positive information tended to be shared by friends about treatment for HMB, using upbeat language. The reported benefits of a particular treatment in the lay press or online underline this, in particular, expectations for the effectiveness for the LNG-IUS. Some women expected a very significant impact, not only on a range of menstrual issues, but also on other areas of life, such as life at home, relationship issues and career. This highlighted how HMB could dominate their lives.

I would never have to worry again, it might be the miracle cure that I am after.

<div align="right">

18A
</div>

Most women had other concerns in relation to HMB. These included pain (cramping period pain), fatigue, anaemia, irritability, low mood, joint pain, lack of energy and headaches. These were highly individual and often combined. Many women did not think of their menstrual bleeding in terms of blood volume, but were far more concerned about rate of blood flow, the presence of blood clots and controlling this flow in order to maintain concealment of their menstrual state. In addition, related pain, cramps and backache were the most significant issues they wanted to address. Many of these women felt this pain to be severe and debilitating and were completely unable to continue with their normal daily activities during menstruation, resulting in taking time off work.

It's a strange sensation because it's coming away from you and you have no control over it . . . I have blood clots sometimes the size of a fist and it's awful.

<div align="right">

15A
</div>

But it wasn't the actual bleeding part, it was more the pain … doubled over with the pain.

<div align="right">

5A
</div>

Then the last year they have got worse with pain and having to have time off work for a couple of days … It always felt like my insides were dropping out the pain on the first 2–3 days.

<div align="right">

5A
</div>

Women described complex behaviours and coping mechanisms to enable them to both manage their menstrual bleeding and to conceal their condition not only from work colleagues and friends, but also from family members. Preparatory coping behaviours included carrying extra sanitary protection at all times, frequent changes and checking of blood loss, and having a spare change of clothes (sometimes two).

It was every handbag had some sanitary wear in it just in case . . . I always knew it was best to stock up at that time ready for next time.

<div align="right">

29B
</div>

Avoidance behaviours included rescheduling activities at home and at work, cancelling meetings both social and professional, taking time off sick from work and remaining at home during menstruation. Other mechanisms included having someone else take care of children, or take their place or shift at work. These coping mechanisms for HMB allowed the woman to rest and ensured that their menstrual difficulties remained concealed. If these failed, women reported feeling embarrassment, guilt and shame, particularly if other people needed to help them. They often became distressed and upset, affecting self-esteem and mood.

It was on our staff party [date] it was and it was mid-afternoon and . . . I just thought I need to go to the ladies but . . . when I got to the ladies toilets . . . I had wet my, soiled my trousers, my underwear was saturated through and the pain and the contractions were really bad and it was just pouring out of me . . . I was too embarrassed to come out . . .

<div align="right">

39B
</div>

Heavy menstrual bleeding could have a negative and wide-ranging impact on women's QoL, which was most often related to their ability to function effectively in their daily roles, to be able to cope well with their responsibilities at home, at work and within their relationship with their partner during menstruation. The negative impact of HMB on their roles and responsibilities was influential in women's decisions to seek treatment.

I would say it's not trivial. You know to them it may be something that you know if they are a gynaecologist and they are dealing with it may be the least of someone's problems but to that woman it is the butt of her problems you know and it has an impact on her whole being her whole holistic being revolves around that one thing.

27B

Beyond peers and family, few women had sought information about treatment from other sources such as the Internet and all had a treatment preference prior to their medical consultation. The most requested treatment was the LNG-IUS. Women who expressed a preference for other standard treatments did not want the LNG-IUS because of its effect on fertility and also the invasive nature of the coil. Women who rejected the LNG-IUS because of its contraceptive effect were all in their forties and most strongly wanted to conceive a child.

Initial experiences of the levonorgestrel-releasing intrauterine system

Many women who were treated with the LNG-IUS found the coil fitting to be more distressing than anticipated and some felt unprepared psychologically for the procedure. While many doctors did not recommend the LNG-IUS for nulliparous women, multiparous women still had problems. They recounted feeling pain and discomfort in excess of other medical procedures within their previous experience and used emotive language.

I suppose I was a bit naïve, I suppose a picture might have helped. I was thinking of like a smear test but they actually have to open you up a bit more and obviously that's not very pleasant.

29B

I had it done [LNG-IUS fitting] and it really hurt. I think you would really have to make that clear to somebody. Even if they have not even had a baby or something. It hurts for a short time I suppose. But you wouldn't like, if you were doing that to somebody's neck it would be classed as torture or something you know it really is nasty. You know a bit sort of medieval.

29B

Some women had been told to take oral analgesia before the procedure, but many did not. Most women did not specifically mention or recall local anaesthesia in discussing their experience of the LNG-IUS insertion but those that did were dissatisfied with it.

I don't think she put any [expletive] numbing lotion there to be fair, there is no way on earth that she put anything to numb me because there was no numbness.

Women felt that health professionals taking consent for the procedure had downplayed the possible intensity of the pain and also the complexity of the procedure itself. Women had a strong sense that while this was a simple procedure for the health professional, it was not simple for them experiencing it.

She said it's going to be feeling like this, this, and this, you are going to feel a little bit of pain and it should be over and done within like 2 minutes, OK, right, not a problem. Oh how very wrong she was. The little [nationality] nurse I was holding her hand and I had to let go of her hand because I would have broke her fingers because in the end it was literally gripping on to the sides of the bed, tears coming down my eyes, excruciating pain . . . I felt sick, I felt giddy I just wanted to be sick and then she done it and the pain was just excruciating, she said you are going to have to wait like half an hour in here. I couldn't walk when I got off.

23A

Women then found the LNG-IUS troublesome within the first 3 months of treatment because irregular bleeding interfered with women's long-established behavioural coping mechanisms.

> *I was quite anxious because I kept thinking that at some point they could come back or they could start or I could spot and that did actually happen. Initially I was having spotting or a mild period pain and things like that but they just got less and less.*

Some considered removal of the LNG-IUS during this time but did not seek to do this partly because they were advised or anticipated it would improve over time, or they thought removal would be difficult, or they might lack an alternative treatment. Some women did not ask for the LNG-IUS to be removed because they did not want to go through the process of revisiting their GP and having a vaginal examination within such a short time.

> *The other day it [intermittent bleeding following treatment with LNG-IUS] drove me so mad somebody actually offered me pliers and I nearly took them [pause] you know.*
>
> *27B*

> *Tell you the truth I did actually think about going back to the doctor because I thought . . . perhaps there was something wrong, perhaps the coil perhaps wasn't working properly or perhaps something had happened and it was sort of making it go a bit weird but then I never sort of got round to making an appointment and then the next month would come and it would be a bit better and then by the third or fourth month it seemed to have settled down anyway.*

Most women attending for their routine check, around 6 weeks after IUS fitting, were asked by health professionals about whether or not their menstrual blood loss had reduced. However, women found that explaining blood loss in terms of volume was challenging and, moreover, felt that questioning changes in menstrual blood flow did not reflect a more desired holistic change for them.

> *I think sometimes [doctors] are dealing with the problem as opposed to the woman and that is the mistake; that is the mistake.*
>
> *27B*

Women expected a period of transition following the initial fitting of the LNG-IUS, perhaps reflecting information given by most health professionals about the initial side effects at the time the LNG-IUS was fitted. However, women were surprised at the amount of disruption to their cycle or amount of irregular bleeding as this was greater than either their peers or health professionals had indicated, and indeed some had thought their menstrual bleeding would cease completely within 1 or 2 months

> *It took a while, it didn't do much at the beginning, it took a while for it to do anything. I didn't think it was going to do anything at first but yeah, it took a while and I noticed it slowing down. That was the big thing for me. It didn't matter if it didn't stop or anything, just slowing it down. It has done it in the same year, it has done it bit by bit, it slowed down and slowed down and now for the last 2 or 3 months I have not had anything, not a thing.*
>
> *15B*

Women who had a positive experience of treatment were more inclined to talk about it, while women who had had a negative experience of treatment were more likely to remain silent. Women did not anticipate major changes in their menstrual cycle, such as irregular bleeding and spotting, or that this might last much longer than expected before they saw an improvement in their condition.

Experiences of the levonorgestrel-releasing intrauterine system after 12 months of treatment

Most women persevered with the LNG-IUS and at 12 months were satisfied with this treatment, experiencing a positive and often major impact on their QoL. The treatment took effect gradually over the first 3–6 months.

> *It started off really, really excellent then we tended to have a bit of a dip where it was a bit hit and miss some months but now it seems to have settled back down again so I don't know whether that was just a fluke or perhaps whether I was a bit more stressed at work at the time and you know busy in my life that made an impact on that but now on the whole is the best decision I could have made.*
>
> *26B*

Most women had, as expected, experienced amenorrhoea (cessation of periods) with this treatment after 12 months, but even those still having menstrual periods felt significantly better.

> *I was concerned because obviously they advised that if this [LNG-IUS] didn't control it, it would be a hysterectomy, which was obviously not what I would like. I was only 40, 39 or 40, at the time and so I had that fitted now for a while and everything is now wonderful so I have had no more problems any more.*
>
> *42B*

> *It is such a relief, it really is, after the pain that I used to have before and the tiredness and the flooding and the whole, it was almost like your whole life revolved around periods and being comfortable or uncomfortable and you know it was, it was awful and I was feeling constantly tired and thinking this is just, you know, that just awful and now that has all gone and it is just such a relief, it really is.*
>
> *16B*

> *It just sort of gave us I suppose my bit of life back, my bit of independence back that I could do, you know carry on without thinking you know have I leaked.*
>
> *26B*

Most women with a LNG-IUS were happy to continue with this treatment long term. Several women who had considered hysterectomy prior to commencing their treatment with the LNG-IUS now no longer considered surgery as necessary or appropriate use of health-service resource.

> *What I have now is not worth going through an operation for . . . If it hadn't worked and I was still having heavy periods then I would think about it but no, not with the coil, no it's not worth having that done, taking up space in the hospital just to have that.*
>
> *5B*

Experiences of standard treatments

Women who were prescribed standard treatments (most commonly oral tranexamic acid, but also norethisterone, medroxyprogesterone acetate injection and the oral contraceptive pill) had widely varying perceptions and experiences. They were regularly less satisfied with their treatment, with most having ceased prescribed medication at 12 months.

Women were generally dissatisfied with tranexamic acid, but did acknowledge some reduction in menstrual blood flow. Some were concerned about the physical size of tablets and the number that they had to take each day (two tablets, three times per day is commonly recommended). Most did not perceive significant improvements in the short term and so stopped taking their medication after 2 months.

If I take them it does keep the bleed, you know, under control. If I leave it a little bit too long before I take the next lot I am just like flooding again so it's, I am not really one to take tablets though, it's very hard.

32B

It's like remembering to take X, Y, Z and the amount of tablets all the time, that can be a problem . . . remembering to take tablets and making sure there's the right spaces in the day between these tablet taking that can be a problem in itself 'cos you might not be where your tablets are or you might be, you know, hung on at work doing extra hours and that and it's all that and it's all that added stress that you don't really need.

7B

I was taking, I had got good at remembering to take them, take them just before and take the maximum dose and I thought they are not making much difference at all, I still can't get out of bed without hopping to the bathroom cross-legged.

29B

The experience of the few women prescribed oral and injectable contraception for HMB was also generally not satisfactory.

I hardly have any periods, I think I have 2 days of bleeding, don't get any pain but 2 weeks before my period is due I get terrible depression where I am always crying. I just feel dead depressed and I just let small things become big things and it's just when I am coming up to my period.

43B (oral contraceptive)

So I had the injection in [month] I haven't stopped bleeding since, now it's not working and I am getting chronic migraines all the time. And I have ballooned in weight in the last 3 months so suddenly it is not working, not suiting me. I am changing my treatment as of next week.

18B (Depo-Provera)

Oral contraception was continued for other reasons such as familiarity, reversible contraceptive effect and lack of interference with other medical treatments.

I have been taking the iron tablets, which help with the dizzy spells and things like that. I did try the other medication [tranexamic acid] but I tried it and I felt, no, I didn't feel anything changes much and therefore I thought is there any point taking another medicine if it doesn't make any difference? I still have the pain and I still felt the same and the blood flow seems to be no better so I haven't taken it again.

28B

Women perceived standard treatments to be a short-term measure. Many who discontinued treatment elected not to try any other treatment, preferring to manage their condition themselves by changing coping behaviours and strategies. This was particularly true of the women who had a long history of HMB and those who anticipated the onset of menopause within the next few years.

Some women were concerned about the physical amount of tablets that they would have to take over the course of their remaining years of menstruation. Most disliked the thought of taking medication, with concerns about the long-term effects of the medications prescribed. For some, this was related to potential unknown side effects.

They [tranexamic acid] did work but I didn't like the quantity of tablets that I was taking.

18A

I will probably go back to the doctor and get some more tablets . . . unless there was something else that they could try instead rather than stay on them long term. But in the short term, yes . . . I know they research them but they never know until years later the effects. It might not have an effect now but in years to come and I'd rather not take the risk of causing anything long term. If I had a choice I would rather not take the tablets.

30B

Women who were initially treated with standard treatments commonly changed their treatment over time, with six women in this sample eventually opting not to use any treatment at all and manage without: two resuming standard treatment after intervals of using no treatment and four women switching to the LNG-IUS. One woman had a hysterectomy.

Discussion

This study has found women's expectations, and experiences of use and effectiveness, of medical treatments for HMB to vary considerably and change over time. These findings highlight a range of implications for clinical practice, which are discussed in *Implications for Practice* below, and may be used to enhance more informed shared decision-making and improve care for HMB when women seek help.

Strengths and limitations

The work provides qualitative insights in to how women experience current medical treatments for HMB, with the advantage of a longitudinal design, over a timeframe of around 1 year. Methods were used to enhance rigour, including purposeful sampling, involving researchers from differing disciplines in analysis and respondent validation. The sample is described in some detail, which may aid assessment of relevance of findings beyond the study and trial context. However, it is recognised that this was inevitably a selected sample of women willing to articulate their experiences. Although participants were engaged in the context of clinical trial recruitment, we actively included those unwilling to be randomised and those with preferences for a particular treatment for HMB.

Comparison with existing work

Previous work has found that health professionals may be preoccupied with volume of blood loss, within a biomedical model, rather than addressing the impact on women's health.[95,96] Although highlighted in more recent research,[89] assessing reduction in blood volume for treatment efficacy,[97] or what reduction in menstrual blood loss is meaningful to women,[98] remains a focus of research. Higher volume of daily menstrual blood loss has been correlated with limitations in functioning and social activities.[99] Women in this study were as concerned with menstrual pain and psychological distress affecting their health and QoL as with HMB itself. This finding is consistent with other work.[95,96] Going beyond reduction in volume of blood loss alone, measures of QoL are now becoming more widely used to determine treatment efficacy.[69,93]

More than 1 year following initial treatment with the LNG-IUS most women were satisfied with their treatment and had experienced an improvement in their perceived QoL. This finding concurs with the quantitative results of the ECLIPSE trial in *Chapters 3* and *4*.[93] Narrative accounts were overwhelmingly positive and a number of women used emotive language to describe the improvement in their health and well-being. Most women had experienced amenorrhoea, but even women who still experienced regular menstrual bleeding felt better. The change in perceptions of health and well-being was always associated with being able to fulfil roles and responsibilities. Some women described an improvement in physical well-being, but the emphasis for all was on the psychological benefits of being able to maintain the societal norms of concealment of menstruation. The later acceptability of the LNG-IUS reported here concurs with Lete *et al.*,[100] who found that 87.9% of pre-menopausal women in their study would use a second LNG-IUS device.

Women who were prescribed tranexamic acid initially took the medication as prescribed and experienced some benefit, but most did not feel better quickly enough to continue with their treatment. However, others have found long-term use of tranexamic acid to be well tolerated over a 15-month timeframe, although emergent adverse effects such as headaches, menstrual and back pain were commonly reported.[86]

Implications for practice

These qualitative findings suggest that discussion with women about medical treatments for HMB should not be confined to orthodox clinical considerations, such as volume of heavy bleeding or a woman's wish to avoid having an IUS or contraception. Rather, health professionals could also helpfully consider the wider impacts of HMB on women's lives and respond to these appropriately in sharing decision-making with women about choosing and using treatments. Women's high expectations of a prompt effect from treatments may be understood in terms of their prolonged toleration of the impacts of HMB and its concealment, which had typically preceded seeking medical help. Potentially high expectations of treatment (and the potential influence of peers) should be appreciated and explored appropriately when discussing treatment.

In assessing treatments with women, practitioners should consider that issues such as effect on rate of blood flow, pain and well-being, as well as wider functioning socially or at work, may be as or more important to individuals than amount of menstrual blood loss. Women's accounts also underline other points for helping women make more informed decisions. These include avoiding 'underplaying' initial problems that may be unanticipated by those considering the LNG-IUS, by ensuring that women are appropriately prepared (physically and psychologically) for fitting the LNG-IUS in practice, and in particular for troublesome changes in pattern of bleeding that may commonly occur and can be prolonged.

Unpredictable irregular bleeding with the LNG-IUS was more problematic than the volume of blood loss for women in this study owing to its impact on their established behavioural coping mechanisms for HMB. Irregular bleeding is recognised as a common reason for discontinuing the LNG-IUS.[48,81] Nevertheless, the current qualitative findings, in line with the strong pragmatic clinical trial evidence reported in *Chapters 3* and *4*, underline that women can be advised that, if they persevere with leaving their IUS in place, they are likely to experience significant improvement in menstrual symptoms and related QoL.

Women's perceptions of other treatment options, as a more short-term measure they could control themselves (stopping and starting as they wished), may provide helpful information to offer women considering what may suit their individual circumstances. Information provided may highlight that improvements in the impact of HMB may be expected, but may be variable and may not be as substantial as for the LNG-IUS, in line with the trial evidence. Women may also be given advice regarding practical considerations, such as the inconvenience of taking several large tablets in the case of the most commonly prescribed oral medication, tranexamic acid.

The findings also highlight an important wider context for discussing HMB treatments in relation to effect on fertility and desire to conceive among women in their forties presenting to services with change in menstrual experience, often at a time of other life transitions.

Chapter 6 Economic evaluation alongside the ECLIPSE trial

Introduction

In this section, the economic evaluation that was carried out alongside the ECLIPSE RCT is reported. The objective of the ECLIPSE trial was to compare the clinical effectiveness and cost-effectiveness of the LNG-IUS and usual medical treatment. Following the recommendations of NICE, that surgical intervention should be the last course of treatment for patients with HMB, the clinical effectiveness and cost-effectiveness of non-surgical interventions has become the primary focus of clinicians and decision-makers alike.

In 2007, NICE introduced guidelines[1] for the LNG-IUS to be used for first-line treatment of HMB, but this was based on limited evidence of cost-effectiveness from one trial.[50] An updated review identified seven further randomised trials[38–44] which compared the LNG-IUS with non-hormonal treatments, but none was considered cost-effective.

The aim of the analysis reported here is to conduct an economic evaluation alongside the ECLIPSE trial to provide robust evidence on the cost-effectiveness of the LNG-IUS compared with usual medical treatment. Furthermore, as impact on QoL is the primary outcome measure in HMB, the outcome was measured using a number of different instruments, including the EQ-5D, Short Form questionnaire-6 Dimensions (SF-6D) and MMAS. Although MMAS is disease specific, the EQ-5D and SF-6D are both general health-related instruments that represent the extra-welfarist approach to measure outcomes. However, at the inception of this trial in 2002, the use of SF-6D was in its infancy and still being developed.[101] It was included in the current study as a methodological component. In the intervening years much has been written on the different results produced by the EQ-5D and SF-6D instruments, especially regarding floor and ceiling effects. Further exploration of the methodological differences is beyond the scope of this study. Therefore, for clarity, the base case will report results based on the EQ-5D, although results using the other instrument will also be reported. The influence on the cost-effectiveness results of using these different extra-welfarist measures is explored briefly. As the ECLIPSE trial has been described in *Chapters 3* and *4* of this report, only a brief overview of the trial is provided here as part of the methods. The remaining section of the methods describes the model, the derivation of transition probabilities and utility values using the trial data and resource use.

Methods

A model-based economic evaluation in the form of a cost–utility analysis (CUA) based on an outcome of cost per QALY was carried out alongside the ECLIPSE trial.[93] The analysis took a UK NHS perspective in a primary care setting and provides an assessment of the difference in costs and QALYs. The QALY encapsulates quantity of life and QoL into a single metric. Two analyses based on two different time horizons are considered. The first is analysis based on the outcome at 2 years and the second is based on the outcome at 5 years. As the economic evaluation was carried out alongside a trial, an intention-to-treat analysis was adopted. A societal perspective to include private costs to women was considered, but deemed not to be feasible given the resource constraints for data collection.

Participants and trial design

The key information required to understand the economic evaluation is summarised again here. Briefly, 571 women with HMB from 63 UK centres were randomised between February 2005 and July 2009. Women between 25 and 50 years of age presenting to their GP with HMB, occurring over at least three consecutive cycles, provided written informed consent to participate. Women were randomised to having a LNG-IUS fitted, or usual medical treatment, chosen by the GP and the woman, based on contraceptive needs or desire to avoid hormonal treatment.

Usual medical treatment options included mefenamic acid, tranexamic acid, norethisterone, a combined oestrogen–progestogen or progestogen-only oral contraceptive pill (any formulation), or medroxyprogesterone acetate injection.[1,18] The particular medical treatment was specified prior to randomisation. Treatment reviews by GPs were carried out at 6 weeks and 3 months. Subsequently, treatments could be changed or discontinued owing to perceived lack of benefit, side effects, change in contraception requirements, referral for endometrial ablation or hysterectomy as per usual practice.[1,18] Treatment changes reported by patients were confirmed with the GP.

Model

A model-based analysis was used as a vehicle for the within-trial economic analysis to comprehensively account for the changes in QoL that occurred while the women were taking these treatments, but all the QoL data used in the model are based on primary data collected in the trial.

As outlined in *Quality-of-life measures*, data on utilities (or QoL) were collected at baseline and at 6 months, 1, 2 and 5 years. The trial data showed that, because of its non-curative nature, women were changing their treatment more frequently within this time to identify the best method for managing HMB, and this process had an influence on their QoL. The analysis does not lend itself to a regression framework because patients change between different health states on a monthly basis and QoL was not measured at that frequency. It would therefore be inappropriate to infer QALYs in a typical trial-based analysis from the QoL scores at the time point they happen to be taken in the trial. The most suitable method to capture changes in QoL occurring throughout the trial and provide a robust CUA is to represent these experiences as health states in a decision model, which follows the process of management of HMB used in the ECLIPSE trial. Therefore, QoL values were attached to the health states. A measure of change in utility from baseline to the end point of the trial, as in a typical trial-based analysis that does not use a decision model, would not accurately capture the health states that women had experienced throughout the time span of the trial. Furthermore, a trial-based analysis, without a decision model, would not comprehensively capture the time spent in health states or the associated repetitive costs and resource use. The decision model, based on trial data, provides a more realistic explanation of the utility pathway, providing information that can be synthesised with other data and projected forward.

A state transition[102] model was developed using Microsoft Excel® (Microsoft Corporation, Redmond, WA, USA). A Markov model appropriately combines data for each of the pathways and takes account of the cyclical and repetitive nature of events and facilitates a simulation approach. All parameters used in the model were based on the trial data.

Model structure

The model structure was informed by clinical input and the pathways followed by women in the trial. Several iterations of the structure were made as the model was continuously modified to ensure that it most closely reflected the trial data. Although the structure was not informed by literature, the health states for the model were developed using trial data and by adapting those used previously.[1,103] The inclusion of menopause as a health state was considered but the 2-year time horizon of the analysis and the age of the patients, between 25 and 50 years, meant very few women, if any, would be reaching

menopause within 2 years. At the 5-year time point the inclusion of menopause was considered but deemed inappropriate as this specific information was not captured (see *Additional model assumptions for the analysis based on the outcome at 5 years*). For similar reasons, death was not included as a health state: the interventions do not cause death and the age of the women meant that the likelihood of death from natural causes was minimal. As models are a simplification of reality, it is necessary to include only transitions between health states that commonly occur for the majority of patients, and this was not the case for death and menopause.

Figure 7 presents the clinical pathways and the progress of the two cohorts of women in the ECLIPSE trial who were randomised to either the LNG-IUS or usual medical treatment. The pathway for each treatment can be represented in the same structure.

Definition of health states for levonorgestrel-releasing intrauterine system pathway

- HMB: every woman in the ECLIPSE trial cohort suffers from HMB.
- LNG-IUS: all women have the LNG-IUS inserted.
- Well with the LNG-IUS: following the insertion of the LNG-IUS women are satisfied with treatment and remain in this health state.
- Symptomatic with the LNG-IUS: following the insertion of the LNG-IUS, adverse effects may arise and women may feel unsettled with their current treatment. From this state women may remain in the state or change to an alternative treatment, (i.e. usual treatment, surgery or no treatment).
- Surgery: women may choose to have surgery, which includes endometrial ablation or hysterectomy, determined from the data. Women can move to surgery from 'symptomatic with the LNG-IUS', 'change to usual medical treatment' or 'no treatment'.
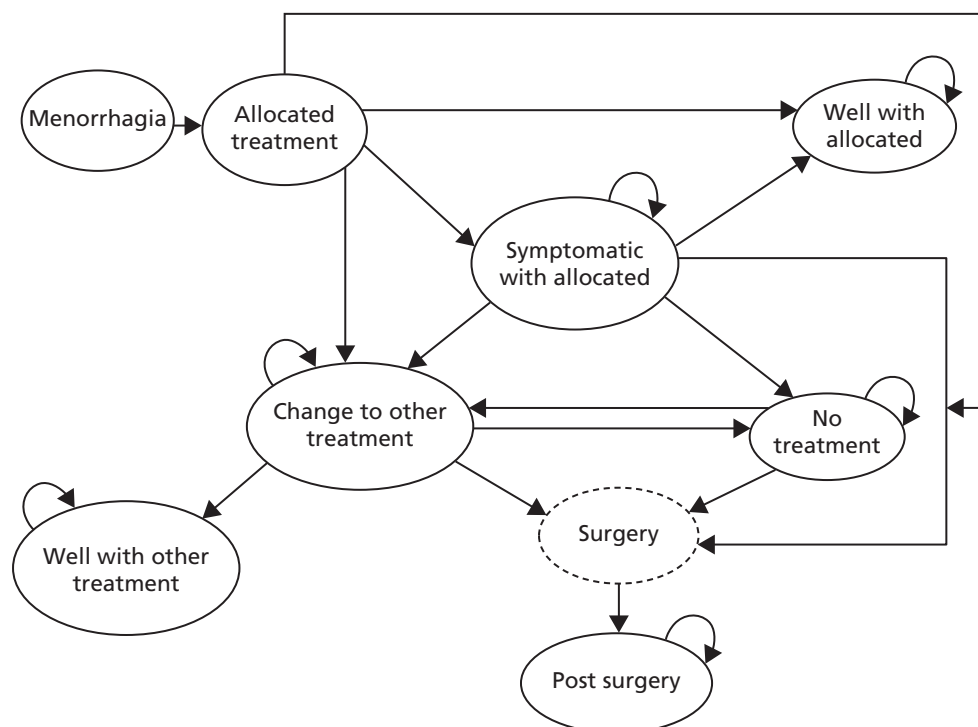
FIGURE 7 Clinical pathway for the LNG-IUS and usual medical treatment.

- Post surgery: following 'surgery', women will then remain in the 'post-surgery' state.
- Change to usual medical treatment: if women have chosen 'no treatment' or are 'symptomatic with the LNG-IUS' they may choose to begin usual medical treatment. Women may immediately become 'well with usual medical treatment' or may experience adverse effects and feel unsettled with usual medical treatment and then eventually become 'well with usual medical treatment'. Alternatively, women may choose to have 'surgery' or 'no treatment'. If women are in this state at 3 months, the cost of a GP review will be incurred. Women can change to usual medical treatment from the LNG-IUS in the first cycle. (see *Model assumptions* for further details).
- Well with usual medical treatment: following 'change to usual medical treatment' women are satisfied with this treatment and remain in this health state.
- No treatment: following the LNG-IUS, 'symptomatic with the LNG-IUS' and 'change to usual medical treatment', women may choose to discontinue treatment altogether. Following this, women may decide to 'change to usual medical treatment' or have 'surgery'.

Definition of health states for usual medical treatment pathway

- HMB: every woman in the ECLIPSE trial cohort suffers from HMB.
- Usual medical treatment: all women have usual medical treatment.
- Well with usual medical treatment: following 'usual medical treatment' women are satisfied and remain in this health state.
- Symptomatic: following usual medical treatment, adverse effects may arise and women may feel unsettled. From this state, women may have an alternative treatment (i.e. 'change to the LNG-IUS', 'surgery' or 'no treatment').
- Surgery: women may choose to have surgery, which includes endometrial ablation or hysterectomy, determined from the data. Women can move to surgery from 'symptomatic with usual medical treatment', 'change to the LNG-IUS' or 'no treatment'.
- Post surgery: following 'surgery', women will then remain in the 'post-surgery' state.
- Change to the LNG-IUS: if women have chosen 'no treatment' or are 'symptomatic with usual medical treatment' they may choose to begin treatment with the LNG-IUS. Women may immediately become 'well with the LNG-IUS' or may experience adverse effects and feel unsettled with the LNG-IUS and then eventually become 'well with the LNG-IUS'. Alternatively, women may choose to have 'surgery' or 'no treatment'. Women can change to the LNG-IUS from usual medical treatment in the first cycle. (see *Model assumptions* for further details).
- Well with the LNG-IUS: following 'change to the LNG-IUS' women are satisfied with treatment and remain in that health state.
- No treatment: following 'usual medical treatment', 'symptomatic with usual medical treatment' and 'change to the LNG-IUS' women may choose to discontinue treatment altogether. Following this, women may decide to 'change to the LNG-IUS' or have 'surgery'.

Time cycle

A monthly time cycle was used in the analysis as this represents the clinically meaningful changes observed in treatment and resource use. The ECLIPSE trial data showed that, in many cases, treatments had been altered and women had visited their GP within a month of treatment administration. A monthly time cycle was therefore required to reflect the data and incorporate the associated costs and treatment changes into the CUA. A half-cycle correction is not applicable in this model because the monthly time cycles used in the model are genuine discrete time events and this transition did not occur in the middle of a cycle. On this basis, a correction would not have made a substantial difference to the costs and outcomes.[104]

The following assumptions were made and developed based on clinical expertise from practising primary care and gynaecology clinicians, in addition to standard treatment protocol. The first set of analyses to be carried out was based on the outcome at 2 years post intervention. Some additional assumptions were required to extend the analysis to include a 5-year time horizon and these are reported separately. For both sets of analyses, the assumptions were agreed prior to conducting the analysis. Unless otherwise stated, the assumptions used for the analysis at 2 years were maintained for the 5-year analysis.

Model assumptions

Model assumptions for the analysis based on the outcome at 2 years

- A woman is 'well' with the allocated treatment if she does not change or stop treatment. Some of these women may not be 'well', but are coping with treatment, and the utility values for the 'well' with allocated treatment state reflect this.
- A woman who is 'well with the LNG-IUS' and 'well with usual medical treatment' may report different QoL scores. This is because the LNG-IUS tends to stop bleeding, whereas oral medication helps to manage the bleeding.
- A woman who is 'well with the LNG-IUS' or 'well with usual medical treatment' cannot spontaneously become 'symptomatic'.
- Based on the data, if in the first cycle women move from the allocated treatment to an alternative state other than 'well', it is assumed that they move to either the 'change to alternative treatment' or 'no treatment' state. It is assumed that they do not move to the 'symptomatic' state in the first cycle because insufficient time has elapsed for this to be established and so it is assumed they changed for other reasons.
- From the second cycle onwards, if women change from their allocated treatment they do not go to 'well' but to the 'symptomatic' state and move on from there.
- For the transition to 'surgery', data were collected on whether or not a woman had ablation or hysterectomy, but not the precise technique (e.g. thermal balloon endometrial ablation or microwave endometrial ablation). Data on the weighted likelihood of surgery undertaken were taken from a previous study.[103] It is assumed that if a woman in the trial has endometrial ablation then it will be for her first ablation and we apply the cost for first-line endometrial ablation techniques.
- Once a woman has changed from the allocated treatment, it is not possible for the woman to move back to the allocated treatment.
- It is assumed that if a woman 'changes to the other treatment', she must spend at least one cycle in 'change to other treatment' before she can move to 'well with other treatment'. This is assumed as it will take at least one menstrual cycle for any effect to become apparent.
- Adjustment for baseline differences in QoL were not been incorporated because of the risk of imposing a difference at every point in time over the time period of analysis.

Additional model assumptions for the analysis based on the outcome at 5 years

- Only data up to 5 years will be considered in the analysis.
- The utilities and transition probabilities for the 2-year analysis have now been updated to reflect the changes in 5-year data. For example, where a woman did not change treatment during the first 2 years she would have been classed as 'well'. However, if the same was later observed to change treatment at 4 years, at this time point in the 5-year analysis she would be moved to 'symptomatic' and so the new utility values and transition probabilities reflect the updated data.
- Explicit data were not available on whether or not women were experiencing menopause.
- In total across both arms of the trial, in eight cases (five in usual treatment and three in the LNG-IUS), women who were in the 'well' state at the 2-year analysis stated that they have stopped treatment because it was no longer needed. Where a woman stops treatment and the notes state that treatment is 'no longer needed' the woman will remain in the 'well' state.
- As few women have stopped treatment in the 'well' state because treatment is no longer needed, the cost of stopping treatment is negligible and there is little value in adding complexity to the analysis for this group of women.

- Where a woman stops treatment and provides no reason or states 'lack of efficacy', these women will move to 'symptomatic' first and then reach the 'stop treatment' health state. The utility value for the 'stop treatment' health state therefore reflects both those who may be experiencing menopause and those who no longer want to try any other treatment for the condition.
- A paper reporting the results for the 2-year cost-effectiveness analysis for the ECLIPSE trial has been published,[69] so to maintain consistency with the previous analysis, costs reported here are not inflated to 2013 and all are reported in 2011 prices.

Calculating transition probabilities

To use the trial data in the model, it was necessary to identify the distribution of women in the states at any time and the transition probabilities between health states.

For every patient, the time in months from randomisation to every event (i.e. 'stop treatment') was calculated. For each arm of the trial, the movements of all patients between states for every time cycle were identified to determine the transitions made in each cycle. Then from each state (e.g. the LNG-IUS) the total number of transitions from this state were summed to calculate the transition probabilities to each health state, the number of people moving to each state (e.g. well) was divided by the total number of movements from the LNG-IUS. This process was repeated for every other transition from the state and then repeated for all remaining states.

In cases where women made moves that were not possible in the model, for example returning to allocated treatment after changing to another state, these women were censored at the last point a move was possible. This change only occurred in four cases and is advised against clinically. Therefore, this transition was not considered to be a typical pathway and does not represent the population. As the point in the cycle when the women moved back to 'usual medical treatment' is unknown, the women were censored at one cycle prior to the return to randomised treatment.

To be consistent and prevent biasing any data, the two patients who changed to the other treatment arm within the first month of treatment allocation were moved directly to this state without spending any time in symptomatic, as reflected in the data. The treatment was never taken and so considering these patients as symptomatic would not be accurate.

Where the woman 'changes to the other treatment arm' to 'well with the other treatment arm', one of the model assumptions set out above should again be highlighted. That is, women must spend at least one cycle in 'change to other treatment' to then enter the 'well with other treatment' state as it will take this long for the treatment effects to be seen.

As women cannot stay in surgery, it is constructed as a transition state and therefore all of the women in 'surgery' move to 'post surgery' after one cycle. 'Post surgery' and 'well' are both absorbing states. Therefore, as women cannot leave the state once they have entered, the probability of remaining in these states is fixed to one.

The calculation of the utility values assigned to each health state will be described in *Quality-of-life measures* following a description of the outcome measures collected in the trial.

Quality-of-life measures

Outcome measures were collected using both the EQ-5D and the SF-36 at baseline prior to randomisation, then by post at 6 months, and 1 and 2 years post randomisation. The booklet questionnaire given to women in the trial contained the generic EQ-5D 3-level questionnaire, which measures the impact of treatment on broader aspects of health-related QoL.[105] The SF-36 was converted into SF-6D using the algorithm.[106] Utility values for each state in the model were calculated by averaging the EQ-5D or SF-6D values for each woman in the given state at any given time.

Utility values for the individual states were calculated by averaging the EQ-5D and SF-6D values obtained by each woman in the given state at any given time. For example, if a woman is randomised to the LNG-IUS and then does not change treatment she is considered to be in the 'well' state for the remainder of the analysis, as outlined in the model assumptions previously. Therefore, all of the woman's utility values collected at 6 months and 1, 2 and 5 years will be assigned to the 'well' health state in the model. Similarly, if a woman is initially 'symptomatic with the LNG-IUS' and then moves to 'no treatment' at 2 or 5 years, the utility values for the intermediate months will be assigned to 'symptomatic' and the utilities for 2 and 5 years assigned to the 'no treatment' health state. This method was used to derive the utility values because the utility for the state is important, not the values associated with the individual woman's journey, as decision models are a reflection of the typical population.

Costs and resource use

Costs were collected from a UK NHS perspective. Data on health-care resource use, including GP or gynaecologist consultations, were collected from women alongside other outcome measures. Similar to the utility values, data on costs and resource use from the trial were collated and analysed according to the randomised treatment arm. The general health-care costs for both groups included health-care staff costs and the cost of the interventions. LNG-IUS fitting is estimated to take 20 minutes (informed by clinical experts within trial team), requires both a GP and nurse to be present and also requires disposable consumables. Treatment review by the GP was assumed to last 10 minutes (informed by clinical experts within trial team). Staff costs were calculated using nationally recognised reference costs.[107] The costs of standard medical treatment and the LNG-IUS were estimated from the *British National Formulary*.[23] Cost data on surgical interventions were taken from a previously published study and inflated to 2011 prices.[103,107] As recommended by NICE, a discount rate of 3.5% was applied to both costs and utilities as the model time horizon is beyond 1 year.[108] All costs are reported in 2011 prices in UK sterling (£) using the UK hospital and community health services index.[96,107] *Tables 10–12* present the data used in the analysis.

In cases where women were prescribed a combination of usual medical treatments, a weighted average of the cost was taken. Similarly, repeat prescription costs were calculated based on the average weighted cost of repeat prescriptions in the 'change to usual medical treatment' state of the LNG-IUS arm. As the most commonly prescribed usual medical treatments involved GP review for effectiveness at 3 months, it was assumed that GP review occurs at 3 months.

Where one-off costs were incurred, such as the cost of LNG-IUS removal, these costs were assigned to the transition arrow.

Analysis

An incremental CUA provides information on the difference in costs and QALYs between the LNG-IUS and usual medical treatment and is reported in terms of incremental cost-effectiveness ratios (ICERs), as cost per QALY gained. If a treatment is less costly and generates a greater number of QALYs, dominance is said to occur. Analysis was by intention to treat to provide a pragmatic estimate of ICERs. The base-case analysis and three deterministic sensitivity analyses were carried out using the EQ-5D. An additional sensitivity analysis, repeating the base case and its three deterministic sensitivity analyses, was carried out using the SF-6D.

TABLE 10 Cost data used in the analysis at 2 and 5 years

Cost component	Unit cost (£)	Source
LNG-IUS		
Consultation (GP, 10 minutes)	26.67	PSSRU 2011[107]/expert opinion
Insertion		
GP (20 minutes)	53.33	PSSRU 2011[107]/expert opinion
Practice nurse (20 minutes)	17.00	PSSRU 2011[107]/expert opinion
Device cost	88.00	BNF 62[23]
Sterile pack (insertion)	21.63	NICE (inflated to 2011)
Discontinuation		
GP (10 minutes)	26.67	PSSRU 2011[107]/expert opinion
Practice nurse (10 minutes)	8.50	PSSRU 2011[107]/expert opinion
Sterile pack (removal)	3.77	NICE (Inflated to 2011)
Follow-up		
6-week review (GP, 10 minutes)	26.67	PSSRU 2011[107]/expert opinion
3-month review (GP, 10 minutes)	26.67	PSSRU 2011[107]/expert opinion
Usual medical treatment		
Progestogen (Cerazette®, Merck Sharp & Dohme Ltd)	8.68	BNF 62[23]
Tranexamic acid (Cyclokapron®, Meda)	14.30	BNF 62[23]
Mefenamic acid (Ponstan®, Chemidex)	15.72	BNF 62[23]
Norethisterone	2.18	BNF 62[23]
COC (Microgynon®, Bayer)	2.82	BNF 62[23]
Medroxyprogesterone acetate injections	6.01	BNF 62[23]
Consultation (GP, 10 minutes)	26.67	PSSRU 2011[107]/expert opinion
3- or 6-month review (GP, 10 minutes)	26.67	PSSRU 2011[107]/expert opinion
Discontinuation (GP, 10 minutes)	26.67	PSSRU 2011[107]/expert opinion
Surgery		
Cost of surgery	1720.18	Weighted cost from trial

BNF, *British National Formulary*.

TABLE 11 Health state utility data used in the model

Health state	EQ-5D at 2 years	EQ-5D at 5 years	PSA distribution (EQ-5D) at 2 years	PSA distribution (EQ-5D) at 5 years	SF-6D value[a] at 2 years	SF-6D at 5 years	PSA distribution (SF-6D) at 2 years	PSA distribution (SF-6D) at 5 years
LNG-IUS								
LNG-IUS	0.756	0.756	Beta (653, 211)	Beta (653, 211)	0.597	0.597	Beta (10,204, 6883)	Beta (10,204, 6883)
Well with LNG-IUS	0.980	0.808	Beta (1169, 297)	Beta (1139, 271)	0.598	0.603	Beta (17,912, 12061)	Beta (21,541, 14,192)
Symptomatic with LNG-IUS	0.744	0.759	Beta (130, 45)	Beta (394, 125)	0.589	0.587	Beta (3464, 2418)	Beta (6523, 4591)
Change to usual medical treatment	0.817	0.710	Beta (20, 5)	Beta (30, 12)	0.596	0.586	Beta (1066, 723)	Beta (892, 630)
Well with usual medical treatment	0.714	0.702	Beta (66, 26)	Beta (86, 37)	0.594	0.578	Beta (2032, 1390)	Beta (1487, 1084)
No treatment	0.785	0.785	Beta (70, 19)	Beta (134, 37)	0.604	0.613	Beta (2108, 1380)	Beta (5056, 3188)
Surgery	0.620	0.666	Linked to post surgery	Linked to post surgery	0.430	0.442	Linked to post surgery	Linked to post surgery
Post surgery	0.827	0.855	Beta (59, 12)	Beta (145, 25)	0.574	0.590	Beta (330, 245)	Beta (1067, 742)
Usual medical treatment								
Usual medical treatment	0.714	0.714	Beta (514, 206)	Beta (514, 206)	0.603	0.603	Beta (9892, 6519)	Beta (9892, 6519)
Well with usual medical treatment	0.728	0.737	Beta (528, 197)	Beta (329, 118)	0.592	0.595	Beta (9664, 6647)	Beta (8004, 5450)
Symptomatic with usual medical treatment	0.756	0.745	Beta (311, 100)	Beta (673, 230)	0.606	0.602	Beta (5168, 3359)	Beta (10,247, 6779)
Change to LNG-IUS	0.694	0.706	Beta (49, 21)	Beta (85, 35)	0.627	0.603	Beta (2494, 1484)	Beta (1624, 1067)
Well with LNG-IUS	0.801	0.822	Beta (282, 70)	Beta (316, 68)	0.595	0.603	Beta (4069, 2766)	Beta (5782, 3805)
No treatment	0.766	0.784	Beta (223, 68)	Beta (354, 98)	0.586	0.586	Beta (3548, 2509)	Beta (3548, 2509)
Surgery	0.619	0.642	Linked to post surgery	Linked to post surgery	0.454	0.461	Linked to post surgery	Linked to post surgery
Post surgery	0.825	0.855	Beta (64, 14)	Beta (145, 25)	0.606	0.615	Beta (2136, 1391)	Beta (4129, 2586)

PSA, probabilistic sensitivity analysis.
a Values used in sensitivity analysis 4.
Utility values are rounded to three decimal places. α and β values for the PSA distribution are rounded to the nearest whole number. The source of the data is the ECLIPSE trial.

TABLE 12 Probability parameters used in the analysis

Probability parameters	Monthly probability (2-year analysis)	PSA distribution	Monthly probability (5-year analysis)	PSA distribution
LNG-IUS				
LNG-IUS to well with the LNG-IUS	0.639	(182, 103)[a]	0.502	(143, 142)[a]
LNG-IUS to symptomatic with the LNG-IUS	0.253	(72, 213)[a]	0.389	(111, 174)[a]
LNG-IUS to change to usual medical treatment	0.067	(19, 266)[a]	0.067	(19, 266)[a]
LNG-IUS to no treatment	0.042	(12, 73)[a]	0.042	(12, 273)[a]
Remain well with the LNG-IUS	1	Fixed	1	Fixed
Symptomatic with the LNG-IUS to well with the LNG-IUS	0	Fixed	0	Fixed
Remain symptomatic with the LNG-IUS	0.907	(700, 72)[a]	0.953	(2236, 111)[a]
Symptomatic with the LNG-IUS to change to usual medical treatment	0.035	(27, 745)[a]	0.015	(36, 2311)[a]
Symptomatic with the LNG-IUS to no treatment	0.041	(32, 740)[a]	0.022	(52, 2295)[a]
Symptomatic with the LNG-IUS to surgery	0.017	(13, 759)[a]	0.010	(23, 2324)[a]
Remain change to usual medical treatment	0.708	(109, 45)[a]	0.864	(311, 49)[a]
Change to usual medical treatment to well with usual medical treatment	0.208	(32, 122)[a]	0.092	(33, 327)[a]
Change to usual medical treatment to no treatment	0.045	(7, 147)[a]	0.028	(10, 350)[a]
Change to usual medical treatment to surgery	0.039	(6, 148)[a]	0.017	(6, 354)[a]
Remain well with usual medical treatment	1	Fixed	1	Fixed
No treatment to change to usual medical treatment	0	(1, 547)[a]	0	(1, 2393)[a]
Remain no treatment	0.984	(540, 8)[a]	0.995	(2383, 11)[a]
No treatment to surgery	0.016	(10, 538)[a]	0.004	(10, 2384)[a]
Surgery to post surgery	1	Fixed	1	Fixed
Remain post surgery	1	Fixed	1	Fixed
Usual medical treatment				
Usual medical treatment to well with usual medical treatment	0.402	(115, 171)[a]	0.225	(64, 221)[a]
Usual medical treatment to symptomatic with usual medical treatment	0.566	(162, 124)[a]	0.744	(212, 73)[a]
Usual medical treatment to change to the LNG-IUS	0.007	(2, 284)[a]	0.007	(2, 283)[a]
Usual medical treatment to no treatment	0.024	(7, 279)[a]	0.025	(7, 278)[a]
Remain well with usual medical treatment	1	Fixed	1	Fixed
Symptomatic with usual medical treatment to well with usual medical treatment	0	Fixed	0	Fixed
Remain symptomatic with usual medical treatment	0.901	(1474, 162)[a]	0.943	(3450, 210)[a]
Symptomatic with usual medical treatment to change to the LNG-IUS	0.049	(80, 1556)[a]	0.025	(92, 3568)[a]

TABLE 12 Probability parameters used in the analysis (*continued*)

Probability parameters	Monthly probability (2-year analysis)	PSA distribution	Monthly probability (5-year analysis)	PSA distribution
Symptomatic to no treatment	0.040	(65, 1571)[a]	0.026	(95, 3565)[a]
Symptomatic to surgery	0.010	(17, 1619)[a]	0.006	(23, 3637)[a]
Remain change to the LNG-IUS	0.603	(120, 79)[a]	0.845	(496, 91)[a]
Change to the LNG-IUS to well with the LNG-IUS	0.312	(62, 137)[a]	0.107	(63, 524)[a]
Change to the LNG-IUS to no treatment	0.045	(9, 190)[a]	0.027	(16, 571)[a]
Change to the LNG-IUS to surgery	0.040	(8, 191)[a]	0.020	(12, 575)[a]
Remain well with the LNG-IUS	1	Fixed	1	Fixed
No treatment to change to the LNG-IUS	0.001	(1, 852)[a]	0	(2, 4207)[a]
Remain no treatment	0.992	(846, 7)[a]	0.998	(4200, 9)[a]
No treatment to surgery	0.007	(6, 847)[a]	0.002	(7, 4202)[a]
Surgery to post surgery	1	Fixed	1	Fixed
Remain post surgery	1	Fixed	1	Fixed

PSA, probabilistic sensitivity analysis.
a Dirichlet distribution.
α and β values for the PSA distribution are rounded to the nearest whole number.

Uncertainty in the model was explored by conducting both deterministic and probabilistic sensitivity analysis (PSA). Population heterogeneity was not considered by assessing the cost-effectiveness according to population subgroups because there was no a priori reason for expecting different treatment effects in different subgroups.

Deterministic sensitivity analysis at 2 years:

1. Mean utility values for each state used in the base case were replaced by the median utility value. Previously, some published studies[109,110] used the median and not the mean value, which greatly impacts the cost-effectiveness results and is argued to be inappropriate.[103] The current analysis assesses the impact of using such values when primary data are collected.
2. The assumptions used in the UK national guidelines costing template replaced the expert opinion and trial data that were used in the base case. This change applied to the clinical staff member present for the initial consultation, which used a practice nurse and was assumed to be 10 minutes for the initial consultation and insertion (GP in the base case for the initial consultation and 20 minutes for the insertion with both GP and practice nurse), treatment review by a nurse at 6 weeks (GP was used in the base case) only for those with a LNG-IUS fitted and annual follow-up for both treatment groups thereafter (no annual follow-up in the base case).[1]
3. In the base case, it was assumed that when an EQ-5D completion date and notification of change of treatment coincide, the EQ-5D value will belong to the subsequent state. In sensitivity analysis 3, we assigned the EQ-5D value to the state prior to the change.
4. The base-case analyses and the three deterministic analyses (described in 1 to 3 above) were repeated using the SF-6D instead of the EQ-5D to generate QALYs.

The base-case analysis was also carried out for the 5-year time horizon, but it was not deemed necessary to carry out a complete set of deterministic sensitivity analyses if the base-case results at 5 years were unchanged from the results at 2 years. However, the analysis at the 5-year end point was carried based on QALYs generated by both the EQ-5D and SF-6D. The PSA simultaneously changes all the utilities and probability parameters as presented in *Tables 11* and *12*. However, some parameters are fixed, for

example surgery to post surgery, as that is an essential feature of the model. For each parameter, a distribution is assigned and a value for each parameter is randomly drawn from the assigned distribution. This is repeated 1000 times and the range of incremental cost and QALY results for the LNG-IUS and usual medical treatment are presented on the cost-effectiveness plane.[111] These 1000 values were used to construct a cost-effectiveness acceptability curve (CEAC) to illustrate the probability of the LNG-IUS being more cost-effective than usual medical treatment across a range of monetary values that decision-makers may be willing to pay for an additional QALY. This was carried out using both the EQ-5D and SF-6D.

Distributions for the probabilistic sensitivity analysis

A Dirichlet distribution was used for the transition probabilities because this distribution allows for a possible movement to more than two states. Where a transition between states is possible, but did not occur in the trial data, a value of 1 is added to all such transitions to enable the Dirichlet distribution to be assigned and to compensate for not observing the transition in the data.[112] In *Table 12* it can be seen that some distributions are fixed at the given value. This is because these transitions are not dependent on the data and logically it would not be appropriate for the transition to be made probabilistic. Utility parameters are assigned a beta distribution.

Costs are not assigned distributions because they are primarily unit costs, which are variable, but not uncertain. Where costs are uncertain, the PSA on utilities and transitions will be sufficient to account for this uncertainty. Furthermore, as distributions for these costs and transitions would be particularly complex, the small number of uncertain costs and very small cost values are unlikely to affect the results in a manner sufficient to warrant using a separate distribution.

Results

The results of the base-case analysis and sensitivity analyses are presented in two parts. Part 1 depicts the results using the EQ-5D for the outcome at 2 years and 5 years. Part 2 shows the results using the SF-6D for 2 years and 5 years. The results of the model are presented here as aggregate outcomes. The intermediate outcomes of the model are presented in *Appendix 12* for the usual medical treatment group and *Appendix 13* for the LNG-IUS group.

Part 1a: European Quality of Life-5 Dimensions at 2 years

Table 13 presents a summary of the base-case and the deterministic sensitivity analysis results using the EQ-5D.

The base-case results show that the LNG-IUS costs £100 more than usual medical treatment, as the LNG-IUS costs £430 and usual medical treatment costs £330. However, the LNG-IUS also generated 0.067 more QALYs than usual medical treatment, as the LNG-IUS generated 1.580 QALYs and usual medical treatment 1.513 QALYs. The ICER for the base-case analysis shows that the LNG-IUS generates £1600 per additional QALY when compared with usual medical treatment.

In all three of the deterministic sensitivity analyses, the findings supported the base-case results. However, the ICER in each analysis did differ. Sensitivity analyses 1 and 2 had a slightly less favourable effect, increasing the ICER to £2030 and £1640 per QALY gained respectively, while sensitivity analysis 3 resulted in a more favourable effect on the ICER, with a reduction to £1560 per additional QALY.

The results of the PSA in *Figure 8* illustrate the distribution of the incremental costs and effects from 1000 Monte Carlo simulations. It is shown that the majority of the incremental costs and effects lie in the upper right-hand quadrant of the cost-effectiveness plane. This indicates that the LNG-IUS is both more costly and more effective than usual medical treatment. The uncertainty is then summarised in relation to the changes in the decision-makers' threshold for considering an intervention cost-effective in *Figure 9*. This figure depicts the CEAC, which shows that, from £2000 per QALY, the LNG-IUS has a greater probability of being the more cost-effective intervention. This probability increases to over 90% at approximately £4000 per QALY and this figure is maintained at the subsequent thresholds.

TABLE 13 Base-case and deterministic sensitivity analysis results using the EQ-5D at 2 years

Analysis	Total mean costs per intervention (£)	Total mean QALYs per intervention	ICER (£) (vs. usual medical treatment)
Summary of base-case deterministic results			
Usual medical treatment	330	1.513	1600
LNG-IUS	430	1.580	–
Mean difference	**100**	**0.067**	–
Deterministic sensitivity analysis 1ᵃ			
Usual medical treatment	330	1.590	2030
LNG-IUS	430	1.643	–
Mean difference	**100**	**0.053**	–
Deterministic sensitivity analysis 2ᵇ			
Usual medical treatment	340	1.513	1640
LNG-IUS	450	1.580	–
Mean difference	**110**	**0.067**	–
Deterministic sensitivity analysis 3ᶜ			
Usual medical treatment	330	1.514	1560
LNG-IUS	430	1.582	–
Mean difference	**100**	**0.068**	–

a Deterministic sensitivity analysis 1 used median utility values.
b Deterministic sensitivity analysis 2 used NICE assumptions.
c Deterministic sensitivity analysis 3 assigning the EQ-5D completion date utility for change treatment if change treatment date is the same as the EQ-5D completion date.
Costs are rounded to nearest £10. QALYs are rounded to three decimal places.
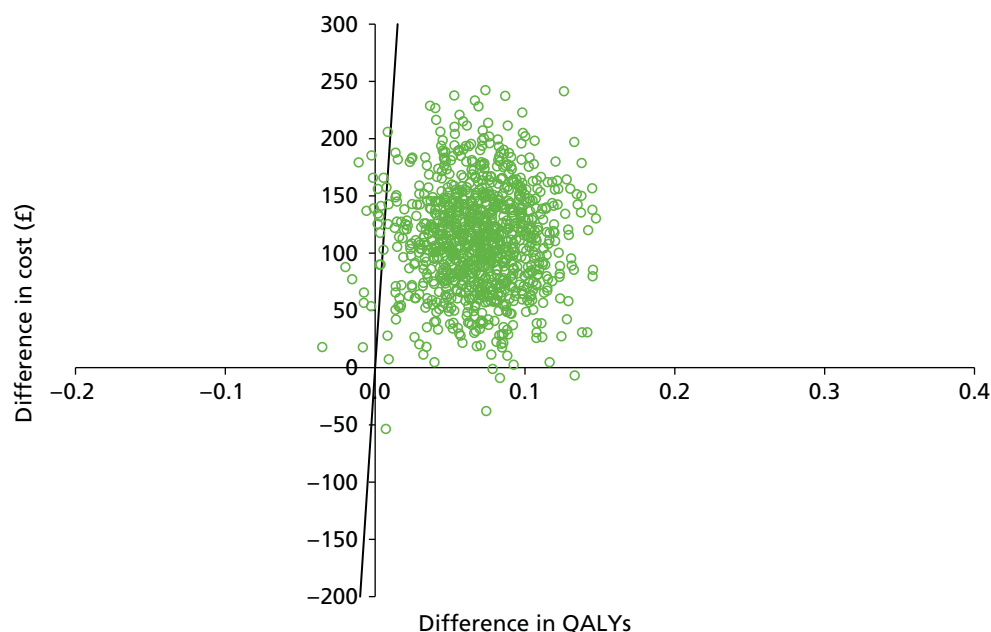
FIGURE 8 Results of the PSA (EQ-5D) at 2 years. The sloping line indicates the £20,000/QALY threshold.
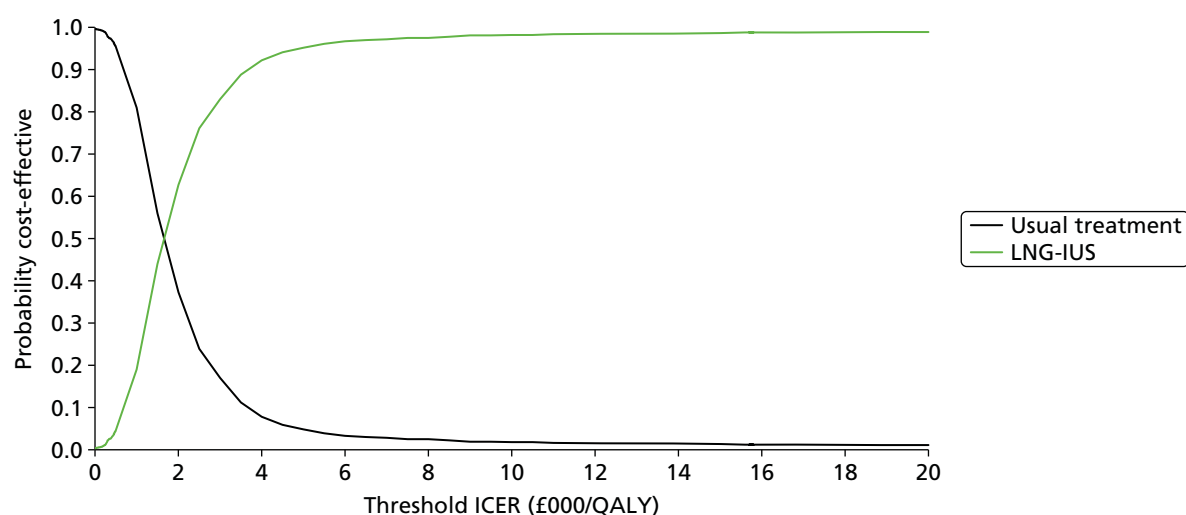
FIGURE 9 Cost-effectiveness acceptability curves for usual medical treatment and the LNG-IUS using the EQ-5D at 2 years. Note that the CEACs have stabilised well before the £20,000 per QALY value. Extending these curves could mean loss of clarity in the early part.

Part 1b: European Quality of Life-5 Dimensions at 5 years

The base-case results at 5 years for the EQ-5D (summarised in *Table 14*) show that the LNG-IUS costs £10 more than usual medical treatment (the LNG-IUS costs £517; usual treatment costs £507). The LNG-IUS also generates 0.090 more QALYs than usual treatment (the LNG-IUS generated 3.698 QALYs and usual medical treatment 3.608 QALYs). Therefore, the additional costs of the LNG-IUS compared with usual medical treatment was £114 per additional QALY gained.

The results of the PSA at 5 years are presented in *Figures 10* and *11* in the CEAC. The CEAC shows that the LNG-IUS has a greater probability of being the more cost-effective intervention from approximately £500 per QALY onwards.

TABLE 14 Base-case analysis results using the EQ-5D at 5 years

Analysis	Total costs per intervention (£)	Total QALYs per intervention	ICER (£) (vs. usual medical treatment)
Summary of base-case deterministic results			
Usual medical treatment	507	3.608	114
LNG-IUS	517	3.698	–
Mean difference	**10**	**0.090**	**–**
Costs are rounded to nearest £10. QALYs are rounded to three decimal places.			
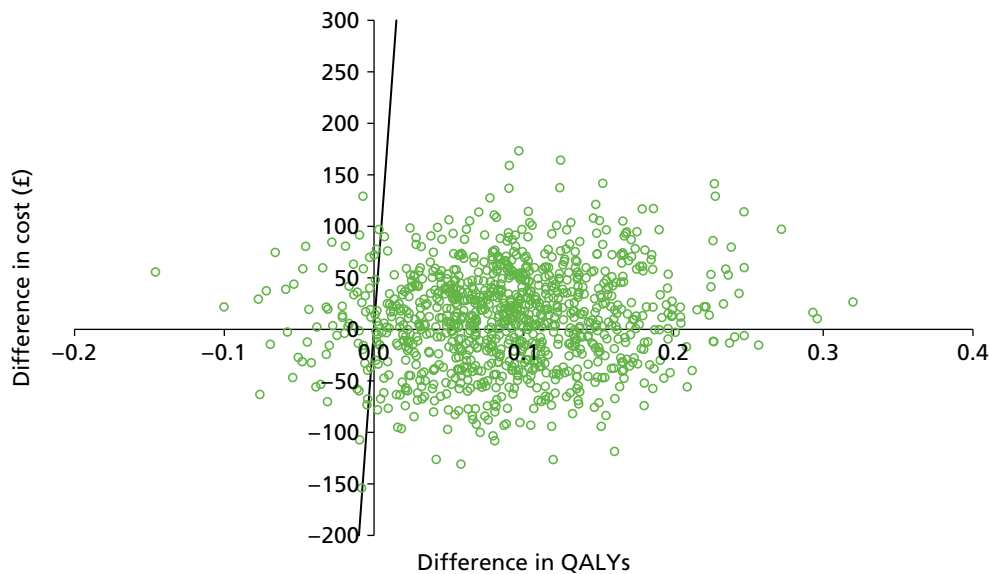
FIGURE 10 Results of the PSA (EQ-5D) at 5 years. The sloping line indicates the £20,000/QALY threshold.
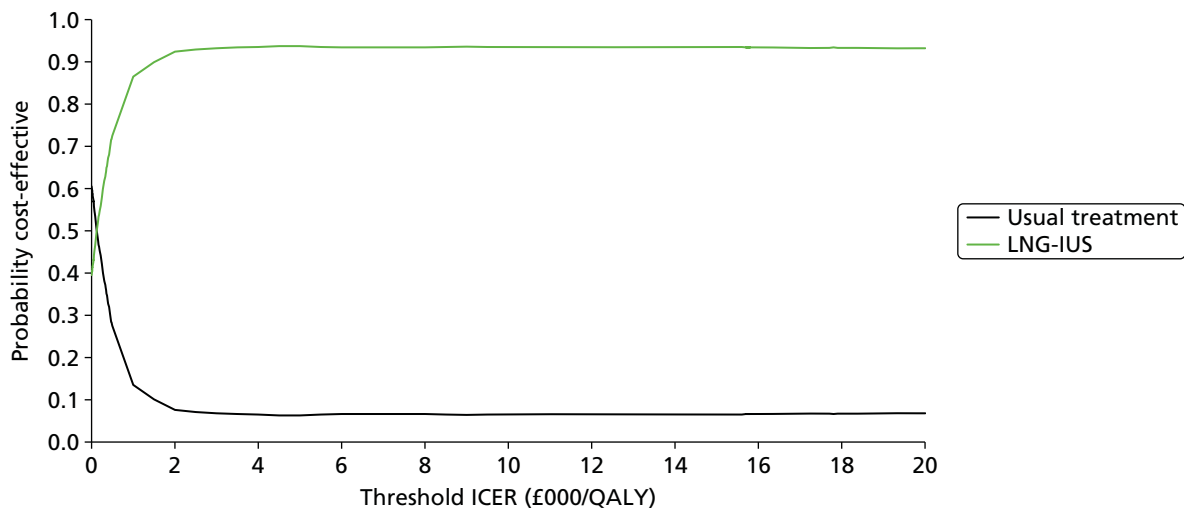
FIGURE 11 Cost-effectiveness acceptability curves for usual medical treatment and the LNG-IUS using the EQ-5D at 5 years. Note that the CEACs have stabilised well before the £20,000 per QALY value. Extending these curves could mean loss of clarity in the early part.

Part 2a: Short Form questionnaire-6 Dimensions at 2 years

Table 15 presents the base-case and deterministic sensitivity analysis results for the SF-6D.

The base-case results show that the LNG-IUS costs £100 more than usual medical treatment as the LNG-IUS costs £430 and usual medical treatment costs £330. However, when the SF-6D is used, usual medical treatment is shown to generate 0.002 more QALYs than the LNG-IUS, as usual medical treatment generated 1.200 QALYs and the LNG-IUS 1.198 QALYs. Therefore, it is shown that usual medical treatment dominates the LNG-IUS.

In two of the three deterministic sensitivity analyses, the findings supported the base-case results. In sensitivity analyses 1 and 2, usual medical treatment dominated the LNG-IUS, whereas in sensitivity analysis 3 usual medical treatment does not dominate the LNG-IUS. The LNG-IUS was shown to be more effective than usual medical treatment and more expensive, generating an ICER of approximately £110,000.

**TABLE 15** Base-case and deterministic sensitivity analysis results using the SF-6D at 2 years

Analysis	Total mean costs per intervention (£)	Total mean QALYs per intervention	ICER (vs. usual medical treatment)
Summary of base-case deterministic results			
Usual medical treatment	330	1.200	Dominates
LNG-IUS	430	1.198	–
Mean difference	**100**	**−0.002**	**–**
Deterministic sensitivity analysis 1[a]			
Usual medical treatment	330	1.215	Dominates
LNG-IUS	430	1.215	–
Mean difference	**100**	**0**	**–**
Deterministic sensitivity analysis 2[b]			
Usual medical treatment	340	1.200	Dominates
LNG-IUS	450	1.198	–
Mean difference	**110**	**−0.002**	**–**
Deterministic sensitivity analysis 3[c]			
Usual medical treatment	330	1.198	£112,340
LNG-IUS	430	1.199	–
Mean difference	**100**	**0.001**	**–**

a Deterministic sensitivity analysis 1 used median utility values.
b Deterministic sensitivity analysis 2 used NICE assumptions.
c Deterministic sensitivity analysis 3 assigning the SF-6D completion date utility for change treatment if change treatment date is the same as the SF-6D completion date.
Costs are rounded to nearest £10. QALYs are rounded to three decimal places.

The results of the PSA in *Figure 12* are presented on the cost-effectiveness plane and show the distribution of the incremental costs and effects from the 1000 Monte Carlo simulations. In this case, when the SF-6D is used to generate the effectiveness outcome, it can be seen that the incremental costs and effects are spread between the upper right-hand and upper left-hand quadrant of the cost-effectiveness plane, indicating that the LNG-IUS is either more costly and more effective or more costly and less effective. This is then summarised in the CEAC in *Figure 13* to show that, for any threshold willingness to pay per QALY, usual medical treatment has the greater probability of being the more cost-effective intervention. This probability is 100% at £0 per QALY and decreases to 90% at approximately £20,000 per QALY.

Part 2b: Short Form questionnaire-6 Dimensions at 5 years

Table 16 presents the results based on using the SF-6D at 5 years. When the EQ-5D values used to generate QALYs were replaced with the SF-6D values, usual medical treatment generates 0.002 more QALYs than the LNG-IUS (usual treatment generated 2.797 QALYs and the LNG-IUS generated 2.795). The costs were unchanged from the previous analysis. Therefore, usual medical treatment is shown to be less costly, but generates more QALYs than the LNG-IUS and, thus, usual medical treatment is said to dominate the LNG-IUS.

The results of the PSA presented in *Figure 14* and the CEAC (*Figure 15*) show that, for any threshold willingness-to-pay per QALY, usual medical treatment has the greater probability of being the more cost-effective intervention.
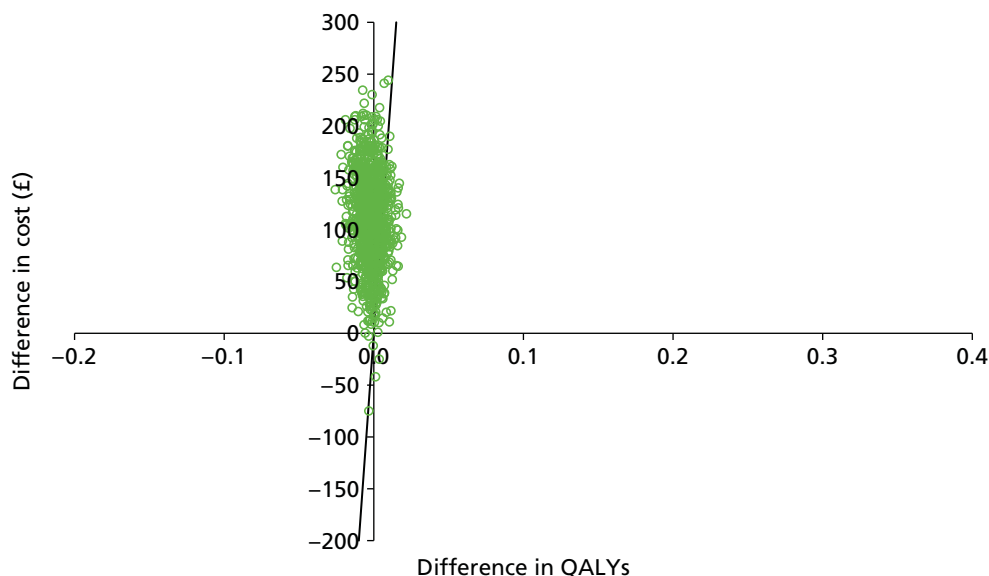
FIGURE 12 Results of the PSA (SF-6D) at 2 years. The sloping line indicates the £20,000/QALY threshold.
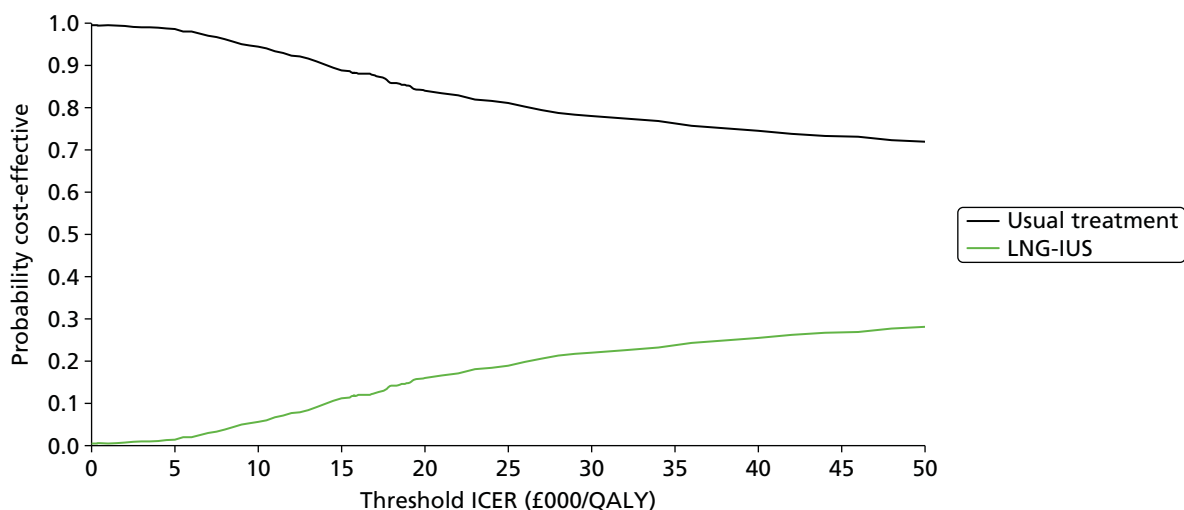
FIGURE 13 Cost-effectiveness acceptability curves for usual medical treatment and the LNG-IUS using the SF-6D at 2 years.

TABLE 16 Base-case analysis results using the SF-6D at 5 years

Analysis	Total costs per intervention (£)	Total QALYs per intervention	ICER (vs. usual medical treatment)
Summary of base-case deterministic results			
Usual medical treatment	507	2.797	Dominates
LNG-IUS	517	2.795	–
Mean difference	**10**	**–0.002**	**–**
Costs are rounded to nearest £10. QALYs are rounded to three decimal places.			
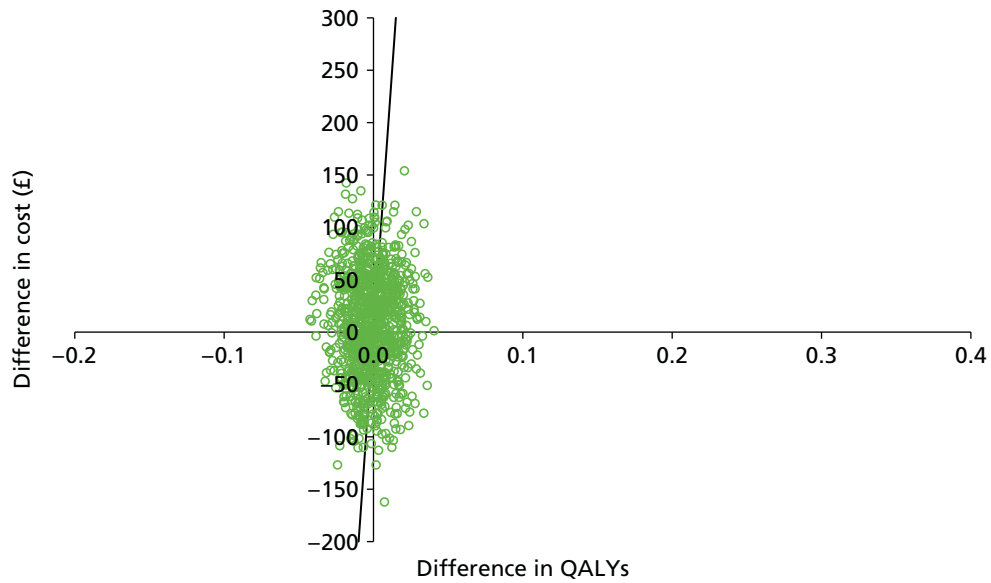
**FIGURE 14** Results of the PSA (SF-6D) at 5 years. The sloping line indicates the £20,000/QALY threshold.
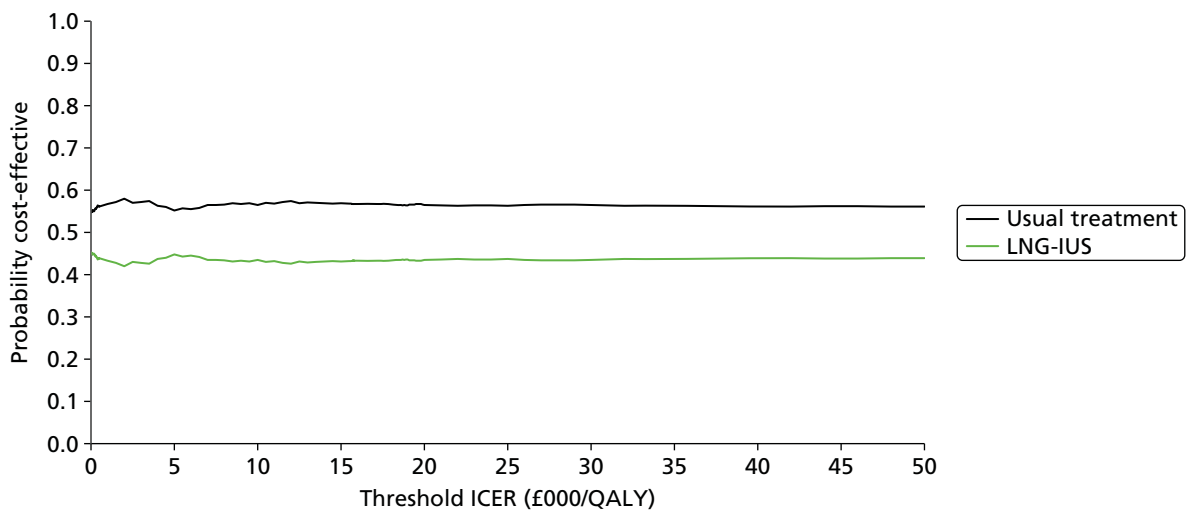
FIGURE 15 Cost-effectiveness acceptability curves for usual medical treatment and the LNG-IUS using the SF-6D at 5 years.

Discussion

Main findings at 2 years

In the primary care setting, treating HMB using the LNG-IUS costs more, but is also more effective than usual medical treatment. The relative cost-effectiveness of the LNG-IUS compared with usual medical treatment is £1600 per QALY. This means every additional QALY costs an extra £1600. The deterministic sensitivity analyses showed the uncertainty in this ICER to be in the range of £1560 to £2030 for an additional QALY gained. As the NICE guidelines recommend new interventions into practice if the ICER is below £20,000 per QALY, the LNG-IUS would be considered cost-effective and recommended as the primary choice for women who require treatment, have no preference against contraception or intrauterine device insertion and no contraindication to LNG-IUS insertion.

However, the importance of selecting the most appropriate QoL instrument is highlighted when the measure used to assess QoL is changed in sensitivity analysis 4 from the EQ-5D to the SF-6D. In sensitivity analysis 4, where utility values from the SF-6D are used rather than the EQ-5D, the cost-effectiveness results differ. In contrast to the findings using the EQ-5D, usual medical treatment is the more cost-effective intervention. Usual medical treatment was found to dominate the LNG-IUS in the base-case and the two deterministic sensitivity analyses. In the third deterministic analysis, the ICER for the LNG-IUS was over £100,000 per QALY, which is much higher than the £20,000/QALY threshold currently set by NICE; therefore, usual medical treatment would be recommended for implementation in clinical practice.

The difference in the cost-effectiveness results derived by using the alternative measures to value QoL has a considerable impact on the cost-effectiveness decision. The different measures resulted in a change, not just in the strength of cost-effectiveness of the same treatment, but also in what was the most cost-effective treatment. Therefore, the recommendation to decision-makers would differ depending on the QoL instrument used.

Main findings at 5 years

The overall findings based on the 5-year time horizon support the findings of the 2-year analysis for both the EQ-5D and the SF-6D. Using the EQ-5D, the LNG-IUS is more costly and more effective than usual medical treatment. The relative cost-effectiveness of the LNG-IUS compared with usual medical treatment is now £114 per QALY. The mean difference in QALYs generated using the EQ-5D between the 2-year and 5-year analysis has increased by 0.023. The costs have reduced from a difference of £100 between the two treatments at 2 years to a difference of £17 at 5 years. The ICER reduced from £1600 per QALY in the 2-year analysis to £114 per QALY in the 5-year analysis. These findings show that at 5 years the LNG-IUS becomes more cost-effective using the EQ-5D, relative to the 2-year analysis. Using the NICE £20,000 per QALY threshold, the LNG-IUS is cost-effective and should be recommended as first-line treatment for HMB.

However, in contrast to the findings, when the EQ-5D is used to estimate the QALY, the results based on the SF-6D show that usual medical treatment is again less costly and more effective. At 5 years, using the SF-6D, usual medical treatment again dominates. While the mean difference in QALYs generated between the treatments is the same for the SF-6D in the 2-year and 5-year analysis (–0.002), the reduction in cost difference between the treatments from 2 years and 5 years (£100 difference between treatments at 2 years and £17 difference at 5 years) has meant that the extent of the dominance has 'increased'. These findings show that at 5 years, using the SF-6D, usual medical treatment would be recommended as first-line treatment for HMB.

The reasons for the costs converging at the 5-year analysis could relate to the LNG-IUS having an initial one-off 'upfront' cost only and usual treatment a small, but continuous cost over time. Alternatively, owing to the non-curative nature of these treatments and the impact on QoL, it could be caused by the numbers of women who changed treatments and treatment arms in both arms of the trial, in addition to the number of women who had surgery, a cure, in both arms.

Strengths and limitations

The strength of this CUA is that it is based on data from the largest multicentre randomised trial undertaken for HMB. As the treatment is aimed at managing the condition, the changes in both QoL and costs throughout the woman's treatment journey are critical to the analyses and these are most appropriately captured by using the trial data to populate a model. The model structure was developed based on the women's pathway, data from the trial and supported by the advice of expert clinicians. All assumptions were agreed by the team in the model development stage prior to analysis. A further strength is that all data on resource use and outcomes were collected prospectively alongside the trial and the economic evaluation was an integral component of the trial design from the outset.

Some limitations exist as a result of some of the assumptions required for the model; for example, it was assumed that women are 'well' if they do not change treatment. Nonetheless, this assumption is mitigated by the fact that, if women enter the 'well' state, but are not well, this will be reflected in the overall utility value for 'well'. Furthermore, utility data were collected at 6 months and at 1, 2 and 5 years and were averaged, so there were insufficient data available to test the time dependency of health states. A further potential limitation is that baseline differences in QoL data at the outset of the trial were not adjusted to be the same. If we did adjust fully for the difference in baseline we would find the LNG-IUS to be less effective than usual medical treatment, but, in so doing, the assumption would be imposed that the difference would be maintained over the time horizon. This would not be an appropriate assumption since the probability of regression to the mean over time would be ignored. It is therefore acknowledged that the base-case results might be slightly overoptimistic, but it is not clear how much of an adjustment, if any, would be appropriate. The initial EQ-5D score in the LNG-IUS group was 0.042 higher at the outset of the trial than in usual medical treatment and this difference is significant ($p < 0.05$). These data are based on individuals who have been randomised, so it is assumed that this difference occurred by chance and it does not follow that this initial difference between groups would be sustained over the 2-year or 5-year time horizon in the absence of treatment. Adjustment for baseline, therefore, risks imposing a difference at every point in time over the time period of analysis.

Comparison with other studies

To our knowledge, this is the first study to conduct a CUA using prospectively collected primary data from a trial to directly compare the LNG-IUS with usual medical treatment for women with HMB. Other economic evaluations of the LNG-IUS have been carried out; however, none has drawn a comparison of the LNG-IUS directly against usual treatment using primary data. Another primary study has shown the LNG-IUS to be cost-effective, but the comparator was hysterectomy and the study is considered to have methodological flaws.[58,103] Other studies have compared alternative treatments using model-based analyses and secondary data from reviews,[1,113,114] although the studies typically compared the LNG-IUS, various oral treatments and various surgical techniques against one another. Two of these showed the LNG-IUS to be the most cost-effective intervention.[1,113] The remaining study suggested that hysterectomy was the optimal intervention, but the authors acknowledged that insufficient published data on the effectiveness of the LNG-IUS were available at the time.[114] Although the LNG-IUS is shown to be cost-effective against surgical interventions in these studies, it does not provide evidence for the LNG-IUS being cost-effective against other pharmaceutical treatments.

Implications and further research

The results, based on the decision-maker-recommended EQ-5D, provide clear evidence in support of the NICE guidelines that recommend that the LNG-IUS be considered the primary treatment for HMB. The main objective of this study was to provide evidence to decision-makers on the cost-effectiveness of two treatments for HMB; therefore, the primary results are based on the EQ-5D. However, as SF-6D data were also collected, the use of this measure was deemed worthy of exploration in the sensitivity analysis.

When the SF-6D is used to generate QALYs, the main base-case results are reversed and usual medical treatment is shown to be the more cost-effective intervention, and this is the case in analyses at both the 2-year and 5-year end points, hence the recommendation to decision-makers differs depending on the outcome measure used. The conflicting findings suggest that these measures may be capturing different aspects of QoL, which clearly has an impact on the results. Reasons why these instruments produce different results have been further explored by other authors,[115,116] but that research is methodological and so far inconclusive, and at the current time does not help to guide which instrument is more appropriate within the context of HMB, particularly when the results are so sensitive to that choice. It may well be that neither instrument is appropriate, with their focus exclusively on health-related QoL and, as evidenced in a recent review of measures for HMB,[69] consideration needs to be given to alternative measures, such as willingness to pay, which provides a broader assessment of well-being.

Additional work is ongoing by the authors of this report to explore the impact on the results of the use of WTP in a cost–benefit analysis. This is in contrast to the more commonly used measures, the EQ-5D and SF-6D, which may be limited by the recall periods used in these measures and their focus on health-related QoL.

Chapter 7 Discussion

Introduction

We completed the following distinct studies as part of this health technology assessment project:

- assessed the clinical effectiveness of the LNG-IUS compared with usual medical treatment for women seeking treatment for HMB in primary care in the short term (2 years following randomisation) through the ECLIPSE pragmatic, multicentre RCT
- assessed the clinical effectiveness in the medium term by follow-up of the ECLIPSE trial cohort of women for 5 years following randomisation
- determined the cost-effectiveness of the LNG-IUS compared with usual medical treatment in the short and medium term through a parallel economic evaluation
- explored the perspectives of trial participants, and women with HMB who declined to be randomised, in a longitudinal qualitative study
- demonstrated the reliability and validity properties of the MMAS.

Each of these studies has been described in detail, with the main findings reported and the conclusions derived in the light of any limitations identified at the end of each *Chapters 2–6*. This chapter attempts to focus on the key findings and limitations emerging from the ECLIPSE trial and associated studies. It is not a comprehensive summary of all of the issues raised, for which the reader is encouraged to consult *Chapters 2–6*.

Principal findings

Evidence for clinical effectiveness

Between February 2005 and July 2009, a total of 571 women with HMB from 63 UK centres were randomly assigned to either the LNG-IUS (285 women) or usual medical treatment (286 women). In the usual medical treatment group, 75% were initially treated with mefenamic acid or tranexamic acid or a combination of these two drugs.

At baseline, the women were substantially affected by HMB, as assessed with the condition-specific MMAS and general (SF-36) health-related scales. There were significant improvements over a period of 2 years in both the LNG-IUS and usual medical treatment groups, which were maintained to 5 years. However, improvements in average MMAS scores were greater within the LNG-IUS group than in the usual medical treatment group over the first 2 years, by an average of 13.4 points in an intention-to-treat analysis. Although MMAS scores and the majority of the SF-36 QoL domains at 5 years still favoured the LNG-IUS, these were no longer statistically significant, and contrast with the relatively large differences we noted at 2 years of follow-up.

The greater improvement with the LNG-IUS over the first 2 years was seen despite 36% of women having had the system removed, generally owing to the lack of effectiveness or to irregular prolonged bleeding, which are well-recognised reasons for discontinuing the LNG-IUS.[61,66] There were higher rates of discontinuation in those allocated usual medical treatments (64% at 2 years' follow-up). The diminution of a difference between trial treatment groups may be unsurprising given the high proportions of women who, by 5 years, had changed to a treatment that worked for them or had ceased bleeding, with only 15% still taking usual medical therapy and 47% having a LNG-IUS in place.

We expected fewer surgical interventions in the LNG-IUS group, but the rates were similarly low in the both groups, with approximate 10% at 2 years and 20% at 5 years. Compared with previous trials,[38,49,51,56,81] this low rate may reflect the eligibility criteria for the trial, since women who had fibroids or other disorders were excluded. However, this finding also suggests that women who present in primary care with an uncomplicated presentation of HMB can be both successfully treated at the initial point of contact, with persisting improvement in impact on QoL and avoid referral to secondary care.

Evidence for cost-effectiveness

Alongside the ECLIPSE trial, an economic evaluation provided robust evidence on the cost-effectiveness of the LNG-IUS compared with usual medical treatment. The general health-related instruments, the EQ-5D and SF-6D, represent the extra-welfarist approach to measure cost-effectiveness. A model-based CUA, based on an outcome of cost per QALY, was carried out.

Using the EQ-5D, at 2 years the relative cost-effectiveness of the LNG-IUS compared with usual medical treatment was £1600 per QALY, which by 5 years was reduced to £114 per QALY. This increase in cost-effectiveness is caused both by the mean difference in QALYs generated increasing by 0.023 and by the costs reducing from a difference of £100 between the two treatments at 2 years, to a difference of £17 at 5 years. However, using the SF-6D to produce utility values, usual medical treatment was assessed as less costly and more effective at both 2 and 5 years.

It can be seen that the difference in the cost-effectiveness results derived by using the alternative measures to value QoL will have a considerable impact on cost-effectiveness decisions. The different measures did not just change the strength of cost-effectiveness of the same treatment, but the most cost-effective treatment itself changed. Therefore, any recommendation to decision-makers will differ depending on the QoL instrument used.

Qualitative insights

There was a need to better understand women's experiences of medical treatments for HMB from a qualitative perspective, alongside quantitative measures. The qualitative longitudinal design of our study over a time frame of around 1 year involved a purposeful sample of women willing to articulate their experiences. This included women participating in the trial and, in addition, those not willing to be randomised because they had a specific preference for a particular treatment for HMB.

The results show that women's experiences and expectations of medical treatments for HMB vary considerably and change over time. In addition to a range of practical considerations for treatments, they emphasise that practitioners should consider that rate of menstrual blood flow, pain and well-being, alongside wider functioning socially or at work, may be as or more important to women as menstrual blood loss experienced.

Women had high expectations of a prompt effect from treatments. Unpredictable irregular bleeding with the LNG-IUS was more problematic than the volume of blood loss for women in this study, owing to its impact on their established behavioural coping mechanisms for HMB.

Measurement properties of Menorrhagia Multi-Attribute Scale

We investigated further the psychometric properties of the MMAS. We demonstrated that, in addition to having high face validity, the MMAS has good convergent and discriminant validity and test–retest reliability. This supported our decision to use the MMAS to measure the impact of HMB and its treatment on women in the ECLIPSE trial.

Strengths and limitations

The strengths of this health technology assessment project is that it is based around a large, well-designed randomised trial. The ECLIPSE study is the world's largest trial comparing medical treatments for HMB over the short to medium term, rather than 6 or 12 months, as in previous studies.[24,33] It had a multicentre design and included patients ethnically representative of the UK population, with relatively low rates of loss to follow-up and the assessment of outcomes over 2 and 5 years (84% and 74% of outcomes collected, respectively). Treatment options in the usual medical treatment group accommodated women's preference for hormonal or non-hormonal treatments and requirements for contraception. Our study also had a primary outcome measure, a patient-reported psychometrically valid condition-specific instrument, MMAS, that better reflects women's personal experience of the burden of HMB. The qualitative findings that describe patient experience were afforded rigour by purposeful sampling, respondent validation and longitudinal design, and support our decision to use the MMAS.

The data provide valuable information for women and clinicians considering what to expect from the medical treatments for HMB over the medium term in a real-world situation. Changes in both QoL and cost throughout women's treatment journey are critical to the economic analysis and these were appropriately captured prospectively in the ECLIPSE trial, with the economic evaluation designed as an integral component of the trial from the outset. The model structure was developed based on the women's pathway data from the trial and supported by the advice of expert clinicians, with all assumptions agreed prior to analysis.

Some limitations of our study should be noted. The range of options available for medical treatment complicates any efforts to compare the LNG-IUS with individual agents. However, the choice among the various treatments is representative of current clinical practice. Although the interventions studied in this trial represent options available in primary care settings in the UK, insertion of intrauterine devices is not part of primary care in all health-care settings, and in some circumstances, it requires consultation with a gynaecologist.

In addition, substantial numbers of patients switched treatments over the course of the study; however, these crossovers would be expected to result in an underestimation of the benefits that might be achieved with perfect compliance. A range of sensitivity analyses did not change the clinical effectiveness conclusions.

Implications for practice

The results provide unique and valuable practical information for women and primary care practitioners when considering choice of, and what to expect from, medical treatments for HMB. Our clinical results support the NICE guidelines,[1] which recommend that medical treatment options should be considered when women initially present with HMB in primary care and that the LNG-IUS should be considered first. Women with HMB, a relatively normal-sized uterus on examination and no other risk factors (i.e. intermenstrual bleeding, post-coital bleeding or irregularity of the menstrual cycle which has been assessed by an endometrial biopsy) can be successfully treated and expect a significant reduction in the range of negative impacts of HMB on QoL. Although the LNG-IUS is more effective in this regard than other usual medical options within the first 2 years of treatment, both the LNG-IUS and usual medical treatments are shown to be helpful initial choices 5 years later. The low overall rates of surgery underline the importance and feasibility of initial medical management of women with HMB in primary care and the avoidance of referral to secondary care.

Using the EQ-5D, there is clear cost-effectiveness evidence to support NICE guidelines that recommend the LNG-IUS should be considered the first-line medical treatment for HMB. However, when the SF-6D is used to generate QALYs, the main base-case results are reversed and usual medical treatment is shown to be more cost-effective at both the 2- and 5-year time-points. Therefore, the recommendation to decision-makers differs depending on the outcome measure used.

Qualitative data indicate that clinicians in primary care should also focus on the wider context for those presenting with HMB rather than solely on the amount of blood loss experienced by women. A general acknowledgement of their HMB, which results in a medical treatment option, should be instigated at the point of contact. Wider public awareness is needed among women so that, should they choose to seek help for HMB, they can benefit from such treatment. This will result in significant impact on the QoL of many thousands of women.

Recommendations for research

It will be important to continue assessment of the ECLIPSE trial participants to 10 years, as we expect that over half of our cohort will have reached menopause at that stage, as our initial recruitment age was a median of 42 years. This would examine the clinical and health service trajectories that women initially treated for HMB in primary care may follow in the longer term. Relevant data are lacking in these contexts. Such further research may helpfully identify the nature of continuing use of treatments, particularly the LNG-IUS, and any differential impact on frequency of surgical interventions.

The difference in the results of the economic outcomes arising from the different instruments used to derive utilities may be explained by the limited recall periods used in these measures and their focus only on health-related QoL. These conflicting findings suggest that these measures may be capturing different aspects of QoL, which clearly has an impact on the results. Further consideration of alternative measures to assess cost-effectiveness, such as WTP, which provides a broader assessment of well-being,[70] is also needed.

Acknowledgements

Contributions of authors

Janesh K Gupta (Cochief Investigator, Professor of Obstetrics and Gynaecology) contributed to the conception and design of all aspects of the project, the conduct of the trial, the interpretation of the results, and the writing and editing of the report.

Jane P Daniels (Triallist, Birmingham Clinical Trials Unit, Deputy Director) contributed to the design of the trial, oversight of the conduct of the trial, the interpretation of the results, and the writing and editing of the report.

Lee J Middleton (Senior Statistician) contributed to the design of the trial, conducted the statistical analyses, prepared the clinical effectiveness results and contributed to the writing of the report.

Helen M Pattison (Professor of Health Psychology) contributed to the conception and design of the trial, the design and analysis of the MMAS validation study, provided input into the qualitative study and the interpretation of the results, and contributed to the writing of the report.

Gail Prileszky (Research Associate) conducted and analysed the qualitative study.

Tracy E Roberts (Professor of Health Economics) and **Sabina Sanghera** (Health Economics Research Fellow) designed, conducted and prepared the results of the economic evaluation and contributed to the writing of the report.

Pelham Barton (Reader in Mathematical Modelling) provided advice on model structure and analysis, and the interpretation of model results.

Richard Gray (Triallist and Professor of Statistics) contributed to the conception and design of the trial, the conduct of the trial, the interpretation of the short-term results and the writing of *Chapter 3*.

Joe Kai (Cochief Investigator, Professor of Primary Care) contributed to the conception and design of all aspects of the project, the conduct of the trial, the interpretation of the results and the writing and editing of the report.

Other contributors

We thank the following people:

Trial Steering Committee

Jim Thornton (chairperson), Irwin Nazareth (independent primary care practitioner), Jayne Fountain (independent statistician), Bill MacKenzie (independent gynaecologist), Klim McPherson (independent epidemiologist) and Elaine Nicholls (patient representative).

Independent Data Monitoring and Ethics Committee

Mary-Ann Lumsden (chairperson), Nick Freemantle and Amanda Farrin (independent statisticians).

Publications

Gupta JK, Kai J, Middleton L, Pattison H, Grey R, Daniels JP. Levonorgestrel intrauterine system compared to usual medical treatment for heavy menstrual bleeding. *N Engl J Med* 2013;**368**:128–37.

Sanghera S, Roberts TE, Barton P, Frew E, Daniels J, Middleton L, *et al.* Levonorgestrel-releasing intrauterine system vs. usual medical treatment for menorrhagia: an economic evaluation alongside a randomised controlled trial. *PLOS ONE* 2014;**9**:e9189.

Pattison H, Daniels J, Kai J, Gupta J. The measurement properties of the menorrhagia multi-attribute quality-of-life scale: a psychometric analysis. *BJOG* 2011;**118**:1528–31.

Data sharing statement

Requests for data sharing for secondary research purposes can be addressed to the corresponding author.

References

1. National Institute for Health and Care Excellence. *Heavy Menstrual Bleeding: Guideline CG44.* London: NICE; 2007.

2. Warner PE, Critchley HO, Lumsden MA, Campbell-Brown M, Douglas A, Murray GD. Menorrhagia II: is the 80-mL blood loss criterion useful in management of complaint of menorrhagia? *Am J Obstet Gynecol* 2004;**190**:1224–9. http://dx.doi.org/10.1016/j.ajog.2003.11.016

3. Warner PE, Critchley HO, Lumsden MA, Campbell-Brown M, Douglas A, Murray GD. Menorrhagia I: measured blood loss, clinical features, and outcome in women with heavy periods: a survey with follow-up data. *Am J Obstet Gynecol* 2004;**190**:1216–23. http://dx.doi.org/10.1016/j.ajog.2003.11.015

4. Coulter A, Kelland J, Long A, Melvill A, O'Meara S. The management of menorrhagia. *Eff Health Care* 1995;**9**:218–26.

5. Vilos GA, Lefrebve G, Graves GR. Guidelines for the management of abnormal uterine bleeding. *J Obstet Gynaecol Can* 2001;**106**:1–6.

6. Albers JR, Hull SK, Wesley RM. Abnormal uterine bleeding. *Am Fam Physician* 2004;**69**:1915–26.

7. Critchley HO, Warner P, Lee AJ, Brechin S, Guihse J, Graham B. Evaluation of abnormal uterine bleeding: comparison of three outpatient procedures within cohorts defined by age and menopausal status. *Health Technol Assess* 2004;**8**(34). http://dx.doi.org/10.3310/hta8340

8. Harlow SD, Campbell BC. Host factors that influence the duration of menstrual bleeding. *Epidemiology* 1994;**5**:352–5. http://dx.doi.org/10.1097/00001648-199405000-00017

9. Shapley M, Jordan K, Croft PR. Abnormal bleeding patterns associated with menorrhagia in women in the community and in women presenting to primary care. *Fam Pract* 2007;**24**:532–7. http://dx.doi.org/10.1093/fampra/cmm068

10. Nicholson WK, Ellison SA, Grason H, Powe NR. Patterns of ambulatory care use for gynecologic conditions: a national study. *Am J Obstet Gynecol* 2001;**184**:523–30. http://dx.doi.org/10.1067/mob.2001.111795

11. Coulter A, Kelland J, Peto V, Rees MC. Treating menorrhagia in primary care. An overview of drug trials and a survey of prescribing practice. *Int J Technol Assess Health Care* 1995;**11**:456–71. http://dx.doi.org/10.1017/S0266462300008679

12. Whiteman MK, Kuklina E, Jamieson DJ, Hillis SD, Marchbanks PA. Inpatient hospitalization for gynecologic disorders in the United States. *Am J Obstet Gynecol* 2010;**202**:541–6. http://dx.doi.org/10.1016/j.ajog.2009.12.013

13. Gath D, Osborn M, Bungay G, Iles S, Day A, Bond A, *et al.* Psychiatric disorder and gynaecological symptoms in middle aged women: a community survey. *BMJ* 1987;**294**:213–18. http://dx.doi.org/10.1136/bmj.294.6566.213

14. Garside R, Britten N, Stein K. The experience of heavy menstrual bleeding: a systematic review and meta-ethnography of qualitative studies. *J Adv Nurs* 2008;**63**:550–62. http://dx.doi.org/10.1111/j.1365-2648.2008.04750.x

15. 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. http://dx.doi.org/10.1016/j.ejogrb.2006.08.008

16. Cromwell DA, Mahmood TA, Templeton A, van der Meulen JH. Surgery for menorrhagia within English regions: variation in rates of endometrial ablation and hysterectomy. *BJOG* 2009;**116**:1373–9. http://dx.doi.org/10.1111/j.1471-0528.2009.02284.x

17. Higham JM, Shaw RW. Clinical associations with objective menstrual blood volume. *Eur J Obstet Gynecol Reprod Biol* 1999;**82**:73–6. http://dx.doi.org/10.1016/S0301-2115(98)00224-3

18. Royal College of Obstetrics and Gynaecology (RCOG). *The Initial Management of Menorrhagia – Evidence-Based Clinical Guidelines No. 1*. London: RCOG Press; 1998.

19. Higham JM, Shaw RW. Risk-benefit assessment of drugs used for the treatment of menstrual disorders. *Drug Saf* 1991;**6**:183–91. http://dx.doi.org/10.2165/00002018-199106030-00004

20. Milsom I, Andersson K, Andersch B, Rybo G. A comparison of flurbiprofen, tranexamic acid and a levonorgesrel releasing intrauterine contraceptive device in the treatment of idiopathic menorrhagia. *Am J Obstet Gynecol* 1991;**164**:879–83. http://dx.doi.org/10.1016/S0002-9378(11)90533-X

21. Lethaby A, Duckitt K, Farquhar C. Non-steroidal anti-inflammatory drugs for heavy menstrual bleeding. *Cochrane Database Syst Rev* 2013;**1**:CD000400. http://dx.doi.org/10.1002/14651858.cd000400.pub3

22. Farquhar C, Brown J. Oral contraceptive pill for heavy menstrual bleeding. *Cochrane Database of Syst Rev* 2009;**4**:CD000154. http://dx.doi.org/10.1002/14651858.cd000154.pub2

23. Joint Formulary Committee. *British National Formulary*. 62nd edn. London: BMJ Group and Pharmaceutical Press; 2011.

24. Irvine GA, Campbell-Brown MB, Lumsden MA, Heikkila A, Walker JJ, Cameron IT. Randomised comparative trial of the levonorgestrel intrauterine system and norethisterone for treatment of idiopathic menorrhagia. *Br J Obstet Gynaecol* 1998;**105**:592–8. http://dx.doi.org/10.1111/j.1471-0528.1998.tb10172.x

25. Suvisaari J, Lähteenmäki P. Detailed analysis of menstrual bleeding patterns after postmenstrual and postabortal insertion of a copper IUD or a levonorgestrel-releasing intrauterine system. *Contraception* 1996;**54**:201–8. http://dx.doi.org/10.1016/S0010-7824(96)00189-8

26. Nilsson CG, Luukkainen T, Diaz J, Allonen H. Clinical performance of a new levonorgestrel-releasing intrauterine device. A randomized comparison with a Nova-T-copper device. *Contraception* 1982;**25**:345–56. http://dx.doi.org/10.1016/0010-7824(82)90092-0

27. Lethaby AE, Cooke I, Rees M. Progesterone or progestogen-releasing intrauterine systems for heavy menstrual bleeding. *Cochrane Database Syst Rev* 2005;**4**:CD002126. http://dx.doi.org/10.1002/14651858.cd002126.pub2

28. Stewart A, Cummins C, Gold L, Jordan R, Phillips W. The effectiveness of the levonorgestrel-releasing intrauterine system in menorrhagia: a systematic review. *BJOG* 2001;**108**:74–86. http://dx.doi.org/10.1111/j.1471-0528.2001.00020.x

29. 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. http://dx.doi.org/10.1016/S0029-7844(97)00226-3

30. 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 Endoscopy* 1998;**7**:61–5. http://dx.doi.org/10.1046/j.1365-2508.1998.00165.x

31. 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. http://dx.doi.org/10.1136/bmj.316.7138.1122

32. Cameron IT, Leask R, Kelly RW, Baird DT. The effects of danazol, mefenamic acid, norethisterone and a progesterone-impregnated coil on endometrial prostaglandin concentrations in women with menorrhagia. *Prostaglandins* 1987;**34**:99–110. http://dx.doi.org/10.1016/0090-6980(87)90267-X

33. Reid PC, Virtanen-Kari S. Randomised comparative trial of the levonorgestrel intrauterine system and mefenamic acid for the treatment of idiopathic menorrhagia: a multiple analysis using total menstrual fluid loss, menstrual blood loss and pictorial blood loss assessment charts. *BJOG* 2005;**112**:1121–5. http://dx.doi.org/10.1111/j.1471-0528.2005.00642.x

34. Fraser IS, McCarron G. Randomized trial of 2 hormonal and 2 prostaglandin-inhibiting agents in women with a complaint of menorrhagia. *Aust N Z J Obstet Gynaecol* 1991;**31**:66–70. http://dx.doi.org/10.1111/j.1479-828X.1991.tb02769.x

35. Wellington K, Wagstaff AJ. Tranexamic acid: a review of its use in the management of menorrhagia. *Drugs* 2003;**63**:1417–33. http://dx.doi.org/10.2165/00003495-200363130-00008

36. Coulter A, Long A, Kelland J, O'Meara S, Sculpher M, Song F, *et al.* Managing menorrhagia. *Qual Health Care* 1995;**4**:218–26. http://dx.doi.org/10.1136/qshc.4.3.218

37. Cameron IT, Haining R, Lumsden MA, Thomas VR, Smith SK. The effects of mefenamic acid and norethisterone on measured menstrual blood loss. *Obstet Gynecol* 1990;**76**:85–8.

38. Rauramo I, Elo I, Istre O. Long-term treatment of menorrhagia with levonorgestrel intrauterine system versus endometrial resection. *Obstet Gynecol* 2004;**104**:1314–21. http://dx.doi.org/10.1097/01.AOG.0000143824.16435.91

39. Kaunitz AM, Bissonnette F, Monteiro I, Lukkari-Lax E, Muysers C, Jensen JT. Levonorgestrel-releasing intrauterine system or medroxyprogesterone for heavy menstrual bleeding: a randomized controlled trial. *Obstet Gynecol.* 2010;**116**:625–32. http://dx.doi.org/10.1097/AOG.0b013e3181ec622b

40. Gheit SA. Bleeding patterns associated with progestin-only contraceptives: a prospective controlled trial comparing Mirena versus progestin only pill. *Middle East Fertil Soc J* 2009;**14**:216–19.

41. Shaaban MM, Zakherah MS, El-Nashar SA, Sayed GH. Levonorgestrel-releasing intrauterine system compared to low dose combined oral contraceptive pills for idiopathic menorrhagia: a randomized clinical trial. *Contraception* 2011;**83**:48–54. http://dx.doi.org/10.1016/j.contraception.2010.06.011

42. Küçük T, Ertan K. Continuous oral or intramuscular medroxyprogesterone acetate versus the levonorgestrel releasing intrauterine system in the treatment of perimenopausal menorrhagia: a randomized, prospective, controlled clinical trial in female smokers. *Clin Exp Obstet Gynecol* 2008;**35**:57–60.

43. Endrikat J, Shapiro H, Lukkari-Lax E, Kunz M, Schmidt W, Fortier M. A Canadian, multicentre study comparing the efficacy of a levonorgestrel-releasing intrauterine system to an oral contraceptive in women with idiopathic menorrhagia. *J Obstet Gynaecol Can* 2009;**31**:340–7.

44. Kari S. *Randomized, Open-Label, Multicenter Study Comparing the Efficacy and Safety of Mirena to a Three-Month Administration of Danazol Followed by a 3-Month Observation Period in Women Aged over 30 with Functional Menorrhagia*. Berlex, Canada, 2005. Report No.: A00696/DE00650-303003.

45. Endrikat J, Vilos G, Muysers C, Fortier M, Solomayer E, Lukkari-Lax E. The levonorgestrel-releasing intrauterine system provides a reliable, long-term treatment option for women with idiopathic menorrhagia. *Arch Gynecol Obstet* 2012;**285**:117–21. http://dx.doi.org/10.1007/s00404-011-1902-1

46. Muenzen R. A multicenter, randomized, open label, parallel group, active control study to evaluate the efficacy and safety of LNG IUS (Mirena) as compared to medroxyprogesterone acetate during 6 cycles of treatment in patients with idiopathic menorrhagia. *BSP Report* 2009. A38313.

47. 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. http://dx.doi.org/10.1016/S0301-2115(02)00408-6

48. 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. http://dx.doi.org/10.1055/s-2002-32434

49. 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. http://dx.doi.org/10.1111/j.1471-0528.2006.00863.x

50. Hurskainen R, Teperi J, Rissanen P, Aalto A-M, 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. http://dx.doi.org/10.1016/S0140-6736(00)03615-1

51. Marjoribanks J LA, Farquhar C. Surgery versus medical therapy for heavy menstrual bleeding. *Cochrane Database Syst Rev* 2006;**2**:CD003855. http://dx.doi.org/10.1002/14651858.cd003855.pub2

52. de Souza SS, Camargos AF, de Rezende CP, Pereira FAN, Araujo CAA, Silva Filho AL. A randomized prospective trial comparing the levonorgestrel-releasing intrauterine system with thermal balloon ablation for the treatment of heavy menstrual bleeding. *Contraception* 2010;**81**:226–31. http://dx.doi.org/10.1016/j.contraception.2009.09.012

53. Shaw R, Symonds IM, Tamizian O, Chaplain J, Mukhopadhyay S. Randomised comparison of thermal balloon ablation and levonorgestrel intrauterine system in patients with idiopathic menorrhagia. *Austral N Z J Obstetric Gynaecol* 2007;**47**:335–40. http://dx.doi.org/10.1111/j.1479-828X.2007.00747.x

54. Tam WH, Yuen PM, Ng DPS, Lung PL, Lok IH, Rogers MS. Health status function after treatment with thermal balloon endometrial ablation and levonorgestrel intrauterine system for idiopathic menorrhagia: a randomised study. *Gynecol Obstetr Investi* 2006;**62**:84–8. http://dx.doi.org/10.1159/000092660

55. Kupperman M, Varner RE, Summitt RL, Learman LA, Ireland C, Vittinghoff E, *et al.* Effect of hysterectomy vs medical treatment on health-related quality of life and sexual functioning. *JAMA* 2004;**291**:1447–55. http://dx.doi.org/10.1001/jama.291.12.1447

56. Malak KA, Shawki O. Management of menorrhagia with the levonorgestrel intrauterine system versus endometrial resection. *Gynecolog Surg* 2006;**3**:275–80. http://dx.doi.org/10.1007/s10397-006-0234-9

57. 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. http://dx.doi.org/10.1136/bmj.c3929

58. 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. http://dx.doi.org/10.1001/jama.291.12.1456

59. 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. http://dx.doi.org/10.1016/S0301-2115(02)00076-3

60. 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. *Br J Obstet Gynaecol* 1996;**103**:142–9. http://dx.doi.org/10.1111/j.1471-0528.1996.tb09666.x

61. 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. http://dx.doi.org/10.1111/j.1471-0528.1998.tb09968.x

62. 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. http://dx.doi.org/10.1111/j.1471-0528.1998.tb10209.x

63. Garratt AM, Ruta DA, Abdalla MI, Buckingham JK, Russell IT. The SF36 health survey questionnaire: an outcome measure suitable for routine use within the NHS? *BMJ* 1993;**306**:1440–4. http://dx.doi.org/10.1136/bmj.306.6890.1440

64. Boyle MH, Torrance GW. Developing multiattribute health indexes. *Med Care* 1984;**22**:1045–57. http://dx.doi.org/10.1097/00005650-198411000-00007

65. Habiba M, Julian S, Taub N, Clark M, Rashid A, Baker R, *et al.* Limited role of multi-attribute utility scale and SF-36 in predicting management outcome of heavy menstrual bleeding. *Eur J Obstet Gynecol Reprod Biol* 2010;**148**:81–5. http://dx.doi.org/10.1016/j.ejogrb.2009.09.021

66. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal*. London: NICE; 2013.

67. Stead ML, Crocombe WD, Fallowfield LJ, Selby P, Perren TJ, Garry R, *et al.* Sexual activity questionnaires in clinical trials: acceptability to patients with gynaecological disorders. *Br J Obstet Gynaecol* 1999;**106**:50–4. http://dx.doi.org/10.1111/j.1471-0528.1999.tb08084.x

68. DeVellis R. *Scale Development: Theory and Applications*. Thousand Oaks, CA: Sage; 2003.

69. Sanghera S, Frew E, Kai J, Gupta J, Elizabeth Roberts T. An assessment of economic measures used in menorrhagia: a systematic review. *Soc Sci Med* 2013;**98**:149–53. http://dx.doi.org/10.1016/j.socscimed.2013.09.016

70. Frew E, Wolstenholme JL, Whynes DK. Willingness-to-pay for colorectal cancer screening. *Eur J Cancer* 2001;**37**:1746–51. http://dx.doi.org/10.1016/S0959-8049(01)00200-3

71. Pattison H, Daniels JP, Kai J, Gupta JK. The measurement properties of the menorrhagia multi-attribute quality-of-life scale: a psychometric analysis. *BJOG* 2011;**118**:1528–31. http://dx.doi.org/10.1111/j.1471-0528.2011.03057.x

72. Protheroe J, Bower P, Chew-Graham C, Peters TJ, Fahey T. Effectiveness of a computerized decision aid in primary care on decision making and quality of life in menorrhagia: results of the MENTIP randomized controlled trial. *Med Decis Making* 2007;**27**:575–84. http://dx.doi.org/10.1177/0272989X07306785

73. Thirlaway K, Fallowfield L, Cuzick J. The Sexual Activity Questionnaire: a measure of women's sexual functioning. *Qual Life Res* 1996;**5**:81–90. http://dx.doi.org/10.1007/BF00435972

74. 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. I. Introduction and design. *Br J Cancer* 1976;**34**:585–612. http://dx.doi.org/10.1038/bjc.1976.220

75. Cohen J. *Statistical Power Analysis for the Behavioural Sciences*. New York, NY: Academic Press; 1977.

76. Verbeke G, Molenberghs G. *Linear Mixed Models for Longitudinal Data*. New York, NY: Springer; 2000.

77. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. *BMJ* 2011;**342**:d40. http://dx.doi.org/10.1136/bmj.d40

78. Rubin DB. *Multiple Imputation for Non-response in Surveys*. New York, NY: Wiley; 1987. http://dx.doi.org/10.1002/9780470316696

79. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003;**41**:582–92. http://dx.doi.org/10.1097/01.MLR.0000062554.74615.4C

80. Shapley M, Jordan K, Croft PR. Why women consult with increased vaginal bleeding: a case-control study. *Br J Gen Pract* 2002;**52**:108–13.

81. 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. http://dx.doi.org/10.1111/j.1479-828X.2007.00747.x

82. Ewies AA. Levonorgestrel-releasing intrauterine system – the discontinuing story. *Gynecol Endocrinol* 2009;**25**:668–73. http://dx.doi.org/10.1080/09513590903159656

83. Chapple A. Menorrhagia: women's perceptions of this condition and its treatment. *J Adv Nurs* 1999;**29**:1500–6. http://dx.doi.org/10.1046/j.1365-2648.1999.01038.x

84. O'Flynn N. Menstrual symptoms: the importance of social factors in women's experiences. *Br J Gen Pract* 2006;**56**:950–7.

85. Chapple A, Ling M, May C. General practitioners' perceptions of the illness behaviour and health needs of South Asian women with menorrhagia. *Ethn Health* 1998;**3**:81–93. http://dx.doi.org/10.1080/13557858.1998.9961851

86. Chapple A. Iron deficiency anaemia in women of South Asian descent: a qualitative study. *Ethn Health* 1998;**3**:199–212. http://dx.doi.org/10.1080/13557858.1998.9961862

87. Shapley M, Jordan K, Croft PR. Increased vaginal bleeding and psychological distress: a longitudinal study of their relationship in the community. *BJOG* 2003;**110**:548–54. http://dx.doi.org/10.1046/j.1471-0528.2003.02458.x

88. Shapley M, Jordan K, Croft PR. Increased vaginal bleeding: the reasons women give for consulting primary care. *J Obstet Gynaecol* 2003;**23**:48–50. http://dx.doi.org/10.1080/0144361021000043245

89. Santer M, Wyke S, Warner P. What aspects of periods are most bothersome for women reporting heavy menstrual bleeding? Community survey and qualitative study. *BMC Womens Health* 2007;**7**:8. http://dx.doi.org/10.1186/1472-6874-7-8

90. Vuorma S, Teperi J, Hurskainen R, Aalto AM, Rissanen P, Kujansuu E. Correlates of women's preferences for treatment of heavy menstrual bleeding. *Patient Educ Couns* 2003;**49**:125–32. http://dx.doi.org/10.1016/S0738-3991(02)00069-1

91. Coulter A, Peto V, Doll H. Patients' preferences and general practitioners' decisions in the treatment of menstrual disorders. *Fam Pract* 1994;**11**:67–74. http://dx.doi.org/10.1093/fampra/11.1.67

92. Santer M. Heavy menstrual bleeding: delivering patient-centred care. *Br J Gen Pract* 2008;**58**:151–2. http://dx.doi.org/10.3399/bjgp08X277258

93. Gupta J, Kai J, Middleton L, Pattison H, Gray R, Daniels J, *et al.* Levonorgestrel intrauterine system versus medical therapy for menorrhagia. *N Engl J Med* 2013;**368**:128–37. http://dx.doi.org/10.1056/NEJMoa1204724

94. Corbin J, Strauss A. Grounded theory research – procedures, canons and evaluative criteria. *Z Soziol* 1990;**19**:418–27. http://dx.doi.org/10.1007/bf00988593

95. O'Flynn N, Britten N. Diagnosing menstrual disorders: a qualitative study of the approach of primary care professionals. *Br J Gen Pract* 2004;**54**:353–8.

96. O'Flynn N, Britten N. Menorrhagia in general practice – disease or illness. *Soc Sci Med* 2000;**50**:651–61. http://dx.doi.org/10.1016/S0277-9536(99)00318-4

97. Magnay JL, Nevatte TM, Seitz C, O'Brien S. A new menstrual pictogram for use with feminine products that contain superabsorbent polymers. *Fertil Steril* 2013;**100**:1715–21. http://dx.doi.org/10.1016/j.fertnstert.2013.08.028

98. Lukes AS, Muse K, Richter HE, Moore KA, Patrick DL. Estimating a meaningful reduction in menstrual blood loss for women with heavy menstrual bleeding. *Curr Med Res Opin* 2010;**26**:2673–8. http://dx.doi.org/10.1185/03007995.2010.526098

99. Lukes AS, Baker J, Eder S, Adomako TL. Daily menstrual blood loss and quality of life in women with heavy menstrual bleeding. *Womens Health* 2012;**8**:503–11. http://dx.doi.org/10.2217/whe.12.36

100. Lete I, del Carme Cuesta MD, Marín JM, Martínez M, Bermejo A, Arina R. Acceptability of the levonorgestrel intrauterine system in the long-term treatment of heavy menstrual bleeding: how many women choose to use a second device? *Eur J Obstet Gynecol Reprod Biol* 2011;**154**:67–70. http://dx.doi.org/10.1016/j.ejogrb.2010.07.040

101. Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Econ* 2004;**13**:873–84. http://dx.doi.org/10.1002/hec.866

102. Belevantsev VL, Markov GA, Terleeva OP, Shulepko EK. Model of transition from anodization to microarc stage. *Izv Sib Otd an Khim* 1989;**6**:73–80.

103. Roberts TE, Tsourapas A, Middleton LJ, Champaneria R, Daniels JP, Cooper KG, *et al.* Hysterectomy, endometrial ablation, and levonorgestrel releasing intrauterine system (Mirena) for treatment of heavy menstrual bleeding: cost effectiveness analysis. *BMJ* 2011;**342**:d2202. http://dx.doi.org/10.1136/bmj.d2202

104. Briggs A, Sculpher M, Claxton K. *Decision Modelling for Health Economic Evaluation*. Oxford, UK: Oxford University Press; 2006.

105. Brooks R. Quality of life measures. *Crit Care Med* 1996;**24**:1769. http://dx.doi.org/10.1097/00003246-199610000-00027

106. Kharroubi SA, Brazier JE, Roberts J, O'Hagan A. Modelling SF-6D health state preference data using a nonparametric Bayesian method. *J Health Econ* 2007;**26**:597–612. http://dx.doi.org/10.1016/j.jhealeco.2006.09.002

107. Curtis L, Netten A. *Unit Costs of Health and Social Care 2011*. PSSRU, University of Kent; 2012.

108. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal*. London: NICE; 2008.

109. 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) http://dx.doi.org/10.3310/hta8030

110. 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. http://dx.doi.org/10.1017/S0266462300012277

111. Andronis L, Barton P, Bryan S. Sensitivity analysis in economic evaluation: an audit of NICE current practice and a review of its use and value in decision-making. *Health Technol Assess* 2009;**13**(29). http://dx.doi.org/10.3310/hta13290

112. Briggs A, Clark T, Wolstenholme J, Clarke P. Missing . . . presumed at random: cost-analysis of incomplete data. *Health Econ* 2003;**12**:377–92. http://dx.doi.org/10.1002/hec.766

113. Lui Z, Doan QV, Blumenthal P, DuBois RW. A systematic review evaluating health-related quality of life, work impairment, and health-care costs and utilization in abnormal uterine bleeding. *Value Health* 2007;**10**:183–94. http://dx.doi.org/10.1111/j.1524-4733.2007.00168.x

114. You JHS, Sahota DS, MoYuen P. A cost-utility analysis of hysterectomy, endometrial resection and ablation and medical therapy for menorrhagia. *Hum Reprod* 2006;**21**:1878–83. http://dx.doi.org/10.1093/humrep/del088

115. Brazier JE, Kolotkin RL, Crosby RD, Williams GR. Estimating a preference-based single index for the Impact of Weight on Quality of Life-Lite (IWQOL-lite) instrument from the SF-6D. *Value Health* 2004;**7**:490–8. http://dx.doi.org/10.1111/j.1524-4733.2004.74012.x

116. Whitehurst DGT, Bryan S, Lewis M. Systematic review and empirical comparison of contemporaneous EQ-5D and SF-6D group mean scores. *Med Decis Making* 2011;**31**:E34–44. http://dx.doi.org/10.1177/0272989X11421529

Appendix 1 Reported reasons for discontinuing treatment up to 2 years' follow-up: usual medical treatment group

Reason for discontinuing	Frequency
0–6 months	
Cannot swallow tablets	1
Depression, weight gain	1
Diarrhoea	1
Did not like tablets	1
Disliked taking tablets	2
Disliked treatment	1
Family history of blood clots	1
Hair loss	1
Irregular bleeding, eczema	1
Irregular bleeding, prolonged bleeding, depression	1
Irregular, prolonged bleeding, required operation	1
Lack of effectiveness, irregular bleeding, depression	1
Lack of efficacy	26
No longer needed	2
None given	8
PCOS	1
Pelvic pain	1
Pelvic pain, prolonged bleeding	1
Pregnant	1
Prolonged bleeding, skin allergy	1
Sister had pulmonary embolism	1
Skin allergy	1
Trying for baby	1
Total	**57**
6–12 months	
Disliked taking tablets	2
Disliked treatment	2
Irregular bleeding	1
Irregular bleeding, prolonged bleeding, pelvic pain, depression, skin allergy	1
Lack of effectiveness, pelvic pain, depression, weight gain	1

Reason for discontinuing	Frequency
Lack of efficacy	14
Lack of efficacy, irregular bleeding	2
Lack of efficacy, required operation	1
No longer needed	1
None given	11
Pelvic pain, dislikes taking tablets	1
Pregnant	3
Vomiting, irregular bleeding, prolonged bleeding, pelvic pain, depression	1
Weight gain	1
Total	**42**
12–24 months	
Chest pain	1
Concerned about long-term effects	1
Disliked taking tablets	1
Disliked treatment	2
Fibroids for operation	1
High blood pressure	1
Irregular bleeding	1
Irregular bleeding, prolonged bleeding	1
Irregular bleeding, prolonged bleeding, pelvic pain	1
Lack of effectiveness, irregular bleeding, prolonged bleeding	1
Lack of effectiveness, irregular bleeding, prolonged bleeding, pelvic pain	1
Lack of efficacy	24
Lack of efficacy, irregular bleeding	1
Lack of efficacy, pelvic pain	2
Lack of efficacy, upset tummy	1
Lack of efficacy, vomiting, diarrhoea	1
Lack of efficacy, required operation	1
Menopausal	1
No longer needed	2
None given	13
Pelvic pain	1
Pelvic pain, weight gain, headaches, leg cramps	1
Prolonged bleeding, depression	2
Sickness, dizziness	1
Wanted to see if periods had returned to normal without medication	1
Total	**64**

PCOS, polycystic ovary syndrome.

Appendix 2 Reported reasons for discontinuing treatment up to 2 years' follow-up: the levonorgestrel-releasing intrauterine system group

Reason for discontinuing	Frequency
0–6 months	
Abdominal pain	1
Coil expelled	3
Heavy bleeding	1
Irregular bleeding, depression, pelvic pain	1
Irregular, prolonged bleeding	1
Lack of effectiveness, irregular bleeding, pelvic pain, coil expelled	1
Lack of efficacy	1
Leg pains, depressed	1
None given	2
Pain	2
Pain in legs	1
Pelvic pain	2
Pelvic pain, infection	1
Prolonged bleeding	7
Prolonged bleeding, required operation	1
Total	**26**
6–12 months	
Coil expelled	1
Depression, weight gain, bloating	1
Disliked treatment	3
For operation	1
Irregular bleeding, acne	1
Lack of effectiveness, irregular bleeding, prolonged bleeding	1
Lack of effectiveness, pelvic pain, depression	1
Lack of effectiveness, thread problems, loss of libido	1
Lack of effectiveness, vomiting, prolonged bleeding, pelvic pain, depression, clotting, anaemia	1
Lack of efficacy	3
Lack of efficacy, prolonged bleeding	1
Mood swings	1
None given	1
Not happy with treatment	1

Reason for discontinuing	Frequency
Prolonged bleeding	1
Uterine prolapse	1
Vomiting, diarrhoea, depression, mood swings	1
Total	**21**
12–24 months	
Coil expelled	1
Coil moved, pelvic pain	1
Depression, weight gain	1
Disliked treatment	1
For operation	2
Irregular bleeding, prolonged bleeding, weight gain	1
Lack of effectiveness, irregular bleeding, prolonged bleeding, pelvic pain, weight gain	1
Lack of effectiveness, prolonged bleeding, pelvic pain, depression	1
Lack of efficacy	7
Lack of efficacy, prolonged bleeding	1
Lack of efficacy, for operation	3
None given	3
Pelvic pain	1
Pelvic pain, depression, weight gain, tender breasts	1
Prolonged bleeding	1
Prolonged bleeding, pelvic pain	1
Swelling hands, body	1
Thread problems	1
Weight gain	2
Weight gain, metallic taste	1
Total	**32**

Appendix 3 Results of primary outcome: Menorrhagia Multi-Attribute Scale sensitivity analysis (overall difference between groups over 2 years)

Assumption	Difference between groups[a] (95% CI)
Excluding women who switched treatment or ceased taking treatment altogether	18.7 (14.8 to 22.6; $p < 0.001$)
Women who indicated they were no longer bleeding and did not complete MMAS given best possible score of 100	12.8 (9.3 to 16.3; $p < 0.001$)
As above, but also carrying forward a pre-surgery score for those undergoing hysterectomy/ablation	12.2 (8.6 to 15.8; $p < 0.001$)
Excluding MMAS from late returned booklets (> 3 months past due date)	14.5 (10.9 to 18.1; $p < 0.001$)
Missing responses given worst possible score (0)	11.9 (7.2 to 16.7; $p < 0.001$)
Using multiple imputation for missing responses	12.3 (9.2 to 15.4; $p < 0.001$)

a   Based on a repeated measures model. See *Chapter 3*, *Statistical analysis* for details.
Point estimates (95% CI, *p*-value) shown for differences between groups. Estimates > 0 from differences between groups favour LNG-IUS treatment.
0 = worst affected; 100 = not affected.

Appendix 4 Menorrhagia Multi-Attribute Scale overall score stratified by body mass index and treatment group up to 2 years' follow-up
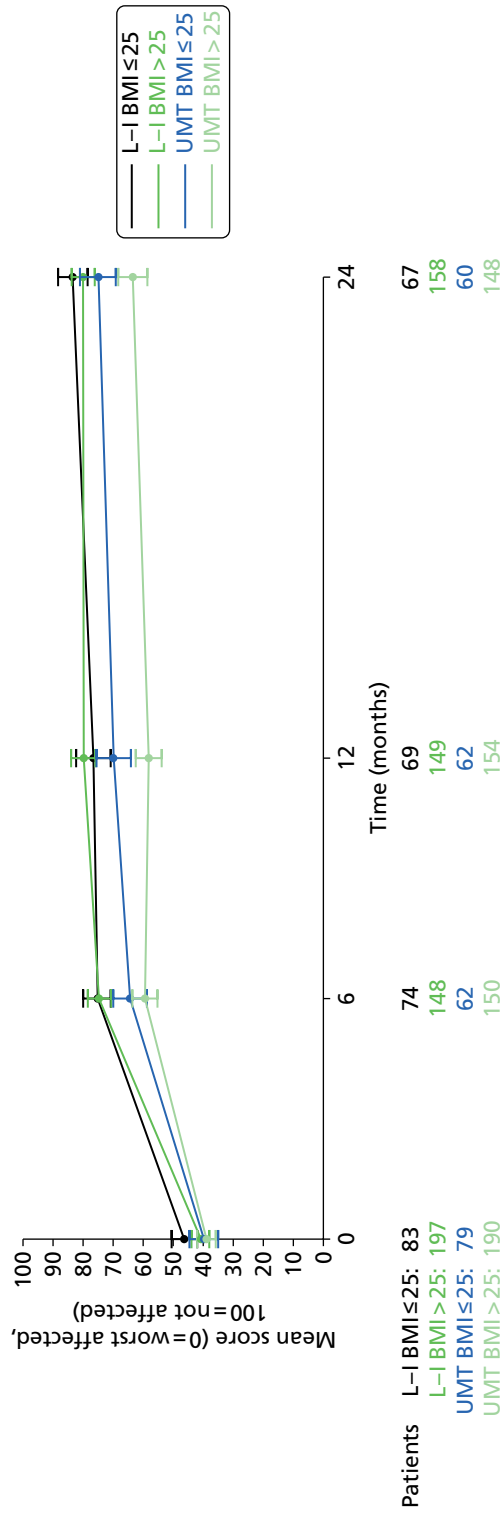
FIGURE 16 Menorrhagia Multi-Attribute Scale overall score stratified by BMI and treatment group up to 2 years' follow-up. L–I, LNG-IUS; UMT, usual medical treatment. Point estimates (95% CIs, *p*-value) shown for differences between groups. Estimates > 0 favours LNG-IUS treatment. 0 = worst affected; 100 = not affected. Based on a repeated measures model. See *Chapter 3, Statistical analysis*, for details.

Appendix 5 Comprehensive results of Short Form-questionnaire-36 items up to 2 years' follow-up

Domain and comparison	Baseline	6 months	12 months	2 years	Overall[a]
Physical functioning					
Usual medical treatment, mean score (SD, *n*)	77.8 (24.7, 264)	79.3 (23.5, 206)	80.0 (23.8, 218)	83.4 (22.6, 230)	–
Change within group, point estimates (95% CI)	–	2.3 (–0.9 to 5.4; *p* = 0.15)	3.2 (0.4 to 5.9; *p* = 0.02)	6.9 (4.0 to 9.8; *p* < 0.001)	–
LNG-IUS, mean score (SD, *n*)	80.0 (20.4, 272)	85.5 (18.2, 216)	84.6 (22.5, 218)	86.3 (22.3, 244)	–
Change within group, point estimates (95% CI)	–	5.5 (2.7 to 8.3; *p* < 0.001)	4.3 (1.3 to 7.4; *p* = 0.006)	6.8 (3.8 to 9.9; *p* < 0.001)	–
Difference between groups, point estimates (95% CI)	**–**	**4.7 (1.2 to 8.1; *p* = 0.009)**	**2.5 (–1.1 to 6.1; *p* = 0.18)**	**1.5 (–2.1 to 5.0; *p* = 0.42)**	**2.7 (0.0 to 5.4; *p* = 0.05)**
Physical role					
Usual medical treatment, mean score (SD, *n*)	68.9 (26.2, 264)	74.1 (26.0, 205)	75.1 (26.9, 219)	80.1 (25.9, 232)	–
Change within group, point estimates (95% CI)	–	5.0 (1.1 to 8.8; *p* = 0.01)	6.2 (2.5 to 9.8; *p* = 0.001)	11.7 (7.9 to 15.4; *p* < 0.001)	–
LNG-IUS, mean score (SD, *n*)	72.1 (24.7, 276)	82.3 (22.0, 219)	83.3 (24.5, 219)	84.6 (24.5, 245)	–
Change within group, point estimates (95% CI)	–	9.9 (6.2 to 13.6; *p* < 0.001)	10.6 (6.9 to 14.3; *p* < 0.001)	13.2 (9.5 to 16.8; *p* < 0.001)	–
Difference between groups, point estimates (95% CI)	**–**	**7.0 (2.7 to 11.3; *p* = 0.001)**	**6.9 (2.5 to 11.4; *p* = 0.002)**	**3.4 (–0.9 to 7.6; *p* = 0.12)**	**5.9 (2.6 to 9.1; *p* < 0.001)**
Emotional role					
Usual medical treatment, mean score (SD, *n*)	69.8 (26.8, 264)	72.3 (28.2, 207)	75.7 (26.1, 220)	77.8 (27.2, 232)	–
Change within group, point estimates (95% CI)	–	1.8 (–2.2 to 5.9; *p* = 0.36)	5.3 (1.4 to 9.3; *p* = 0.009)	8.0 (3.9 to 12.2; *p* < 0.001)	–
LNG-IUS, mean score (SD, *n*)	71.9 (25.1, 276)	79.8 (24.1, 219)	79.7 (25.8, 219)	81.5 (25.3, 247)	–
Change within group, point estimates (95% CI)	–	8.1 (4.4 to 11.7; *p* < 0.001)	7.6 (3.9 to 11.4; *p* < 0.001)	10.2 (6.6 to 13.8; *p* < 0.001)	–
Difference between groups, point estimates (95% CI)	**–**	**7.1 (2.5 to 11.7; *p* = 0.002)**	**3.7 (–0.8 to 8.3; *p* = 0.11)**	**3.3 (–1.2 to 7.8; *p* = 0.15)**	**4.6 (1.3 to 8.0; *p* = 0.007)**

Domain and comparison	Baseline	6 months	12 months	2 years	Overall[a]
Social functioning					
Usual medical treatment, mean score (SD, n)	62.4 (25.9, 268)	68.3 (24.4, 211)	69.0 (26.5, 220)	75.4 (27.0, 233)	–
Change within group, point estimates (95% CI)	–	5.6 (1.9 to 9.3; p=0.003)	6.2 (2.6 to 9.8; p<0.001)	12.5 (8.8 to 16.2; p<0.001)	–
LNG-IUS, mean score (SD, n)	64.3 (24.5, 277)	74.6 (24.6, 221)	76.8 (24.2, 220)	78.9 (24.9, 248)	–
Change within group, point estimates (95% CI)	–	10.3 (6.8 to 13.8; p<0.001)	10.8 (7.4 to 14.2; p<0.001)	14.9 (11.4 to 18.4; p<0.001)	–
Difference between groups, point estimates (95% CI)	**–**	**5.6 (1.4 to 9.8; p=0.009)**	**6.2 (2.0 to 10.5; p=0.004)**	**2.9 (–1.4 to 7.2; p=0.18)**	**5.1 (2.0 to 8.1; p=0.001)**
Mental health					
Usual medical treatment, mean score (SD, n)	59.0 (19.8, 267)	62.5 (20.0, 210)	63.2 (20.5, 218)	66.3 (21.9, 232)	–
Change within group, point estimates (95% CI)	–	3.9 (1.4 to 6.5; p=0.003)	3.9 (1.4 to 6.5; p=0.003)	7.6 (5.0 to 10.3; p<0.001)	–
LNG-IUS, mean score (SD, n)	60.3 (19.3, 277)	63.8 (19.4, 220)	66.0 (21.7, 218)	68.8 (21.1, 246)	–
Change within group, point estimates (95% CI)	–	3.8 (1.3 to 6.4; p=0.003)	4.9 (2.3 to 7.6; p<0.001)	8.8 (6.1 to 11.4; p<0.001)	–
Difference between groups, point estimates (95% CI)	**–**	**0.6 (–2.5 to 3.8; p=0.70)**	**1.8 (–1.5 to 5.2; p=0.29)**	**1.7 (–1.7 to 5.1; p=0.33)**	**1.5 (–1.0 to 3.9; p=0.23)**
Energy/vitality					
Usual medical treatment, mean score (SD, n)	40.8 (21.7, 268)	44.7 (23.0, 211)	47.2 (21.4, 219)	51.8 (23.8, 232)	–
Change within group, point estimates (95% CI)	–	3.7 (0.6 to 6.8; p=0.02)	5.8 (2.9 to 8.7; p<0.001)	11.2 (7.8 to 14.7; p<0.001)	–
LNG-IUS, mean score (SD, n)	40.7 (20.9, 277)	50.5 (21.1, 220)	53.0 (22.0, 218)	55.7 (23.1, 248)	–
Change within group, point estimates (95% CI)	–	9.8 (6.6 to 13.1; p<0.001)	11.5 (8.3 to 14.7; p<0.001)	15.3 (12.3 to 18.3; p<0.001)	–
Difference between groups, mean score (95% CI)	**–**	**5.9 (2.0 to 9.7; p=0.003)**	**5.8 (2.1 to 9.5; p=0.002)**	**3.9 (0.0 to 7.8; p=0.05)**	**5.3 (2.5 to 8.2; p<0.001)**

Domain and comparison	Baseline	6 months	12 months	2 years	Overall[a]
Pain					
Usual medical treatment, mean score (SD, *n*)	49.5 (24.9, 268)	57.1 (25.4, 211)	59.9 (26.9, 220)	67.5 (28.3, 233)	–
Change within group, point estimates (95% CI)	–	5.9 (2.6 to 9.2; *p* < 0.001)	9.8 (6.2 to 13.3; *p* < 0.001)	17.7 (14.1 to 21.4; *p* < 0.001)	–
LNG-IUS, mean score (SD, *n*)	54.2 (24.9, 278)	68.6 (24.0, 221)	70.6 (26.7, 220)	72.9 (25.5, 248)	–
Change within group, point estimates (95% CI)	–	14.9 (11.3 to 18.5; *p* < 0.001)	16.9 (12.8 to 20.9; *p* < 0.001)	19.8 (16.0 to 23.6; *p* < 0.001)	–
Difference between groups, point estimates (95% CI)	**–**	**10.3 (6.1 to 14.4; *p* < 0.001)**	**9.0 (4.4 to 13.7; *p* < 0.001)**	**3.9 (–0.6 to 8.5; *p* = 0.09)**	**7.8 (4.5 to 11.0; *p* < 0.001)[b]**
General health perception					
Usual medical treatment, mean score (SD, *n*)	60.3 (21.9, 264)	61.3 (22.8, 207)	61.4 (22.8, 217)	64.0 (23.8, 233)	–
Change within group, point estimates (95% CI)	–	1.5 (–0.8 to 3.7; *p* = 0.20)	1.5 (–0.9 to 3.9; *p* = 0.22)	4.2 (1.4 to 7.1; *p* = 0.004)	–
LNG-IUS, mean score (SD, *n*)	61.8 (21.4, 274)	66.3 (21.2, 219)	67.0 (22.7, 218)	66.6 (23.0, 247)	–
Change within group, point estimates (95% CI)	–	3.4 (1.0 to 5.7; *p* = 0.005)	4.4 (1.8 to 6.9; *p* < 0.001)	4.8 (2.1 to 7.5; *p* < 0.001)	–
Difference between groups, point estimates (95% CI)	**–**	**2.7 (–0.3 to 5.8; *p* = 0.08)**	**3.7 (0.4 to 7.0; *p* = 0.03)**	**1.3 (–2.3 to 4.9; *p* = 0.46)**	**2.9 (0.3 to 5.4; *p* = 0.03)**

a Based on a repeated measures model. See *Chapter 3, Statistical analysis* for details.
b Some evidence here of time by treatment effect interaction (*p* ≤ 0.05).
Estimates > 0 from differences between groups favour the LNG-IUS. Results at each time point adjusted for baseline score. Change within group compared with baseline.
0 = worst score, 100 = unaffected.

Appendix 6 Comprehensive results of European Quality of Life-5 Dimensions up to 2 years' follow-up

Questionnaire and domain	Baseline	6 months	12 months	2 years	Overall[a]
EQ-5D					
Usual medical treatment, mean score (SD, n)	0.714 (0.276, 269)	0.747 (0.242, 209)	0.734 (0.268, 220)	0.782 (0.231, 231)	–
Change within group, point estimates (95% CI)		0.024 (–0.010 to 0.059; p=0.16)	0.018 (–0.023 to 0.059; p=0.38)	0.070 (0.037 to 0.103; p<0.001)	–
LNG-IUS, mean score (SD, n)	0.756 (0.243, 277)	0.780 (0.221, 218)	0.777 (0.268, 219)	0.807 (0.243, 247)	–
Change within group, point estimates (95% CI)	–	0.022 (–0.010 to 0.054; p=0.18)	0.018 (–0.016 to 0.051; p=0.30)	0.053 (0.021 to 0.085; p=0.001)	–
Difference between groups, point estimates (95% CI)	–	**0.018 (–0.020 to 0.056; p=0.36)**	**0.026 (–0.020 to 0.071; p=0.27)**	**0.007 (–0.031 to 0.045; p=0.72)**	**0.013 (–0.016 to 0.042; p=0.38)**
Health thermometer					
Usual medical treatment, mean score (SD, n)	69.7 (19.8, 246)	68.2 (19.4, 201)	67.9 (21.3, 208)	72.0 (18.7, 220)	–
Change within group, point estimates (95% CI)	–	–1.6 (–4.3 to 1.2; p=0.27)	–2.3 (–5.3 to 0.7; p=0.13)	2.8 (0.2 to 5.5; p=0.03)	–
LNG-IUS, mean score (SD, n)	70.3 (19.1, 250)	72.3 (18.4, 205)	73.0 (19.7, 211)	72.3 (20.8, 239)	–
Change within group, point estimates (95% CI)	–	1.0 (–2.1 to 4.0; p=0.54)	1.2 (–2.1 to 4.4; p=0.48)	1.8 (–1.2 to 4.8; p=0.24)	–
Difference between groups, point estimates (95% CI)	–	**3.1 (–0.4 to 6.5; p=0.08)**	**4.6 (0.8 to 8.4; p=0.02)**	**–0.7 (–4.2 to 2.7; p=0.68)**	**2.0 (–0.5 to 4.6; p=0.12)[b]**

a Based on a repeated measures model. See *Chapter 3, Statistical analysis* for details.
b Some evidence here of time by treatment effect interaction (p ≤ 0.05).
Estimates > 0 from differences between groups favour the LNG-IUS. Results at each time point adjusted for baseline score. Change within group compared with baseline.
EQ-5D: –0.59 = worst score and 1.0 = perfect score; health thermometer: 0 = worst score, 100 = perfect score.

Appendix 7 Comprehensive assessment of the Sexual Activity Questionnaire up to 2 years' follow-up

Domain and resopnse	Baseline	6 months	12 months	2 years	Overall[a]
Pleasure					
Usual medical treatment, mean score (SD, n)	10.9 (4.9, 199)	11.2 (4.8, 145)	11.1 (4.9, 144)	11.8 (4.4, 132)	–
Change within group, point estimates (95% CI)	–	–0.2 (–0.9 to 0.6; p=0.64)	–0.5 (–1.3 to 0.4; p=0.26)	0.6 (–0.3 to 1.5; p=0.21)	–
LNG-IUS, mean score (SD, n)	10.8 (4.9, 210)	11.4 (4.6, 158)	11.7 (4.3, 141)	11.5 (4.5, 136)	–
Change within group, point estimates (95% CI)	–	0.2 (–0.5 to 0.9; p=0.59)	0.5 (–0.4 to 1.4; p=0.26)	0.6 (–0.2 to 1.3; p=0.14)	–
Difference between groups, point estimates (95% CI)	**–**	**0.4 (–0.5 to 1.3; p=0.43)**	**0.9 (–0.2 to 1.9; p=0.10)**	**–0.1 (–1.1 to 0.9; p=0.83)**	**0.4 (–0.3 to 1.1; p=0.26)**
Discomfort					
Usual medical treatment, mean score (SD, n)	4.62 (1.69, 201)	4.67 (1.57, 148)	4.88 (1.49, 147)	4.78 (1.47, 137)	–
Change within group, point estimates (95% CI)	–	0.18 (–0.09 to 0.45; p=0.19)	0.40 (0.08 to 0.72; p=0.01)	0.16 (–0.14 to 0.46; p=0.29)	–
LNG-IUS, mean score (SD, n)	4.65 (1.48, 209)	4.79 (1.52, 159)	4.88 (1.41, 144)	4.85 (1.46, 136)	–
Change within group, point estimates (95% CI)	–	0.11 (–0.11 to 0.32; p=0.34)	0.13 (–0.15 to 0.40; p=0.36)	0.20 (–0.05 to 0.44; p=0.11)	–
Difference between groups, point estimates (95% CI)	**–**	**–0.03 (–0.32 to 0.27; p=0.87)**	**–0.20 (–0.54 to 0.14; p=0.24)**	**0.00 (–0.33 to 0.32; p=0.99)**	**–0.07 (–0.30 to 0.16; p=0.55)**

a Based on a repeated measures model. See *Statistical analysis* for details.
Estimates > 0 from differences between groups favour LNG-IUS treatment. Results at each time point adjusted for baseline score. Change within group compared with baseline.
Pleasure: 0 = low, 18 = high; discomfort: 0 = high, 6 = none.

Appendix 8 Serious adverse events reported within the ECLIPSE trial between February 2005 and August 2011

System (event or reason for hospitalisation)	Usual medical treatment, *n* (%) (*N* = 286)	LNG-IUS, *n* (%) (*N* = 285)
Total number of SAEs[a]	58 (20.3)	49 (17.2)
Number of women with SAEs[b]	51 (17.8)	46 (16.1)
Cancer	–	1 (0.4)
Colectomy	–	1 (0.4)
Cardiac	1 (0.4)	1 (0.4)
Chest pains/shortness of breath	1 (0.4)	1 (0.4)
Ear and labyrinth	–	1 (0.4)
Stapendectomy	–	1 (0.4)
Gastrointestinal	9 (3.1)	2 (0.7)
Cholecystectomy	4 (1.4)	–
Cholelithotomy	1 (0.4)	–
Recurrent abdominal pains	2 (0.7)	2 (0.7)
Hernia repair	1 (0.4)	–
Appendectomy	1 (0.4)	–
Immune system	2 (0.7)	2 (0.7)
Crohn's disease	1 (0.4)	–
Lupus	–	1 (0.4)
Allergic reaction to MRI contrast	–	1 (0.4)
Pulmonary sarcoidosis	1 (0.4)	–
Infections and infestations	2 (0.7)	1 (0.4)
Cellulitis	1 (0.4)	–
Food poisoning	–	1 (0.4)
Suspected viral meningitis	1 (0.4)	–
Metabolism and nutrition	–	1 (0.4)
Diabetes	–	1 (0.4)
Nervous system	2 (0.7)	2 (0.7)
Cluster headaches/migraine	1 (0.4)	–
Stroke	–	2 (0.7)
Non-specific headache	1 (0.4)	–

System (event or reason for hospitalisation)	Usual medical treatment, n (%) (N = 286)	LNG-IUS, n (%) (N = 285)
Pregnancy, puerperium and perinatal situations	7 (2.4)	1 (0.4)
Childbirth – post-partum recovery[c]	7 (2.4)	1 (0.4)
Psychiatric	3 (1.0)	–
Overdose	3 (1.0)	–
Renal and urinary	2 (0.7)	1 (0.4)
Incontinence surgery	2 (0.7)	–
Partial cystectomy	–	1 (0.4)
Reproductive system and breast	21 (7.3)	24 (8.4)
Endometrial ablation	2 (0.7)	1 (0.4)
Hysterectomy	17 (5.9)	19 (6.7)
Ovarian cyst removal	–	1 (0.4)
Uterine prolapse repair	–	1 (0.4)
Bartholin's cyst drainage	–	1 (0.4)
Dilatation and cutterage	1 (0.4)	–
Laparoscopy for pelvic pain	1 (0.4)	1 (0.4)
Other surgical and medical procedures	9 (3.1)	11 (3.9)
Blood transfusion	1 (0.4)	–
Breast reduction	1 (0.4)	2 (0.7)
Hip replacement	–	1 (0.4)
Tonsillectomy	–	1 (0.4)
Gastric band surgery	2 (0.7)	1 (0.4)
Cosmetic surgery	1 (0.4)	1 (0.4)
CT/MRI scan	2 (0.7)	–
Fracture repair	1 (0.4)	2 (0.7)
Back operation (not specified)	1 (0.4)	–
Abscess drainage following procedure	–	3 (1.0)
Death	–	1 (0.4)

CT, computerised tomography; MRI, magnetic resonance imaging; SAE, serious adverse event.
a All listed were classified as SAEs as they resulted in hospitalisation (apart from the recorded death). $p = 0.34$.
b $p = 0.59$ (seven women had two SAEs each in the usual medical treatment group; one woman had three SAEs in the LNG-IUS group, while another had two SAEs).
c All women had ceased treatment prior to becoming pregnant.
Eight women experienced two SAEs. Chest pains/shortness of breath and suspected viral meningitis ($n = 1$); cholecystectomy and cholelithotomy ($n = 1$); recurrent abdominal pains and non-specific headache ($n = 1$); childbirth – post-partum recovery and overdose ($n = 1$); incontinence and hysterectomy ($n = 1$); hysterectomy and laparoscopy for pelvic pain ($n = 1$); and hysterectomy and a blood transfusion ($n = 1$). One woman experienced three events: two strokes and a hip replacement.

Appendix 9 Reported reasons for discontinuing treatment over the whole period of follow-up: usual medical treatment group

Reason for discontinuing	Frequency
Cannot swallow tablets	1
Chest pain	1
Could not find the time to get tablets from chemist	1
Crohn's disease	1
Depression, weight gain	1
Diarrhoea	1
Did not like tablets	1
Did not start treatment, pregnant	1
Did not want to take medication	1
Did not want to take tablets	1
Did not want treatment	1
Disliked taking tablets	5
Disliked treatment	6
Doctor advised owing to age	1
Owing to cancer treatment	1
Family history of blood clots	1
Fibroids, required operation	1
Flooding	1
For operation	4
GP would not prescribe because of health	1
Hair loss	1
IVF treatment	1
Irregular bleeding	2
Irregular bleeding, eczema	1
Irregular bleeding, prolonged bleeding	1
Irregular bleeding, prolonged bleeding, depression	1
Irregular bleeding, prolonged bleeding, pelvic pain	1
Irregular bleeding, prolonged bleeding, pelvic pain, depression, skin allergy	1
Irregular, prolonged bleeding, required operation	1

Reason for discontinuing	Frequency
Lack of effectiveness, irregular bleeding, depression	1
Lack of effectiveness, irregular bleeding, pelvic pain, weight gain	1
Lack of effectiveness, irregular bleeding, prolonged bleeding	1
Lack of effectiveness, irregular bleeding, prolonged bleeding, pelvic pain	1
Lack of effectiveness, pelvic pain, depression, weight gain	1
Lack of effectiveness, prolonged bleeding, depression, weight gain	1
Lack of efficacy	72
Lack of efficacy, anaemia	1
Lack of efficacy, for operation	1
Lack of efficacy, irregular bleeding	4
Lack of efficacy, irregular bleeding, pelvic pain	1
Lack of efficacy, irregular bleeding	1
Lack of efficacy, operation	2
Lack of efficacy, pelvic pain	2
Lack of efficacy, upset tummy	1
Lack of efficacy, vomiting, diarrhoea	1
Lack of efficacy, required operation	2
Menopausal	1
Never took tablet	1
No longer needed	5
No periods	1
No reason given	1
None given	54
PCOS	1
Pelvic pain	2
Pelvic pain, dislikes taking tablets	1
Pelvic pain, prolonged bleeding	1
Pelvic pain, weight gain, headaches, leg cramps	1
Periods better	1
Periods stopped	3
Pregnant	4
Prolonged bleeding	1
Prolonged bleeding, depression	2
Prolonged bleeding, skin allergy	1
Sickness, dizziness	1
Sister had pulmonary embolism	1
Skin allergy	2

Reason for discontinuing	Frequency
Trying for baby	1
Vomiting, diarrhoea, depression	1
Vomiting, irregular bleeding, prolonged bleeding, pelvic pain, depression	1
Wanted a break from treatment	1
Wanted to see if periods had returned to normal without medication	1
Weight gain	2
Total	**228**

IVF, in vitro fertilisation; PCOS, polycystic ovary syndrome.

Appendix 10 Reported reasons for discontinuing treatment over the whole period of follow-up: levonorgestrel-releasing intrauterine system group

Reason for discontinuing	Frequency
Abdominal pain	1
Bacterial vaginosis	1
Coil expelled	7
Coil migrated, surgery required	1
Coil moved, pelvic pain	1
Constant bleeding	2
Depression	1
Depression, weight gain	1
Depression, weight gain, bloating	1
Did not have coil replaced	1
Did not want coil fitted	2
Did not want it fitted	7
Disliked treatment	5
Failed to fit coil	3
For operation	8
Heavy bleeding	1
Infection	1
Irregular bleeding, acne	1
Irregular bleeding, depression	1
Irregular bleeding, depression, pelvic pain	1
Irregular bleeding, pelvic pain	1
Irregular bleeding, prolonged bleeding, weight gain	1
Irregular, prolonged bleeding	1
Lack of effectiveness, irregular bleeding, pelvic pain, coil expelled	1
Lack of effectiveness, irregular bleeding, prolonged bleeding	1
Lack of effectiveness, irregular bleeding, prolonged bleeding, pelvic pain, weight gain	1
Lack of effectiveness, pelvic pain, depression	1
Lack of effectiveness, prolonged bleeding, pelvic pain, depression	1
Lack of effectiveness, thread problems, loss of libido	1
Lack of effectiveness, vomiting, prolonged bleeding, pelvic pain, depression, clotting, anaemia	1

Reason for discontinuing	Frequency
Lack of efficacy	19
Lack of efficacy, coil expelled	1
Lack of efficacy, operation	1
Lack of efficacy, prolonged bleeding	3
Lack of efficacy, for operation	5
Leg pains, depressed	1
Mood swings	1
Mood swings, depression	1
No longer needed	1
No reason given	1
None given	19
Not happy with treatment	1
Nulliparous, obese, uterine length 4.5 cm	1
Pain	2
Pain in legs	1
Pelvic pain	3
Pelvic pain, anaemia, for operation	1
Pelvic pain, depression, weight gain, tender breasts	1
Pelvic pain, infection	1
Periods improved	1
Pregnant	1
Prolonged bleeding	9
Prolonged bleeding, pelvic pain	1
Prolonged bleeding, for operation	1
Swab showed infection	1
Swelling hands, body	1
Technical difficulties	2
Thread problems	1
Thread problems, spotting, irritated cervix	1
Too painful	1
Uncomfortable	1
Uterine prolapse	1
Vomiting, diarrhoea, depression, mood swings	1
Weight gain	3
Weight gain, metallic taste	1
Total	**148**

Reason for discontinuing	Frequency

Appendix 11 Serious adverse events reported within the ECLIPSE trial between February 2005 and September 2014

System (event or reason for hospitalisation)	Usual medical treatment, n (%) (N = 286)	LNG-IUS, n (%) (N = 285)
Total number of SAEs[a,b]	68 (23.7)	58 (20.3)
Number of women with SAEs[c]	61 (21.3)	54 (18.9)
Cancer	–	1 (0.4)
Colectomy	–	1 (0.4)
Cardiac	1 (0.4)	1 (0.4)
Chest pains/shortness of breath	1 (0.4)	1 (0.4)
Heart attack	–	2 (0.7)
Ear and labyrinth	–	1 (0.4)
Stapendectomy	–	1 (0.4)
Gastrointestinal	9 (3.1)	2 (0.7)
Cholecystectomy	4 (1.4)	–
Cholelithotomy	1 (0.4)	–
Recurrent abdominal pains	2 (0.7)	2 (0.7)
Hernia repair	1 (0.4)	–
Appendectomy	1 (0.4)	–
Immune system	2 (0.7)	2 (0.7)
Crohn's disease	1 (0.4)	–
Lupus	–	1 (0.4)
Allergic reaction to MRI contrast	–	1 (0.4)
Pulmonary sarcoidosis	1 (0.4)	–
Infections and infestations	2 (0.7)	1 (0.4)
Cellulitis	1 (0.4)	–
Food poisoning	–	1 (0.4)
Suspected viral meningitis	1 (0.4)	–
Metabolism and nutrition	–	1 (0.4)
Diabetes	–	1 (0.4)
Nervous system	2 (0.7)	2 (0.7)
Cluster headaches/migraine	1 (0.4)	–
Stroke	–	2 (0.7)
Non-specific headache	1 (0.4)	–
Pregnancy, puerperium and perinatal situations	7 (2.4)	1 (0.4)
Childbirth – post-partum recovery[d]	8 (2.8)	2 (0.7)

System (event or reason for hospitalisation)	Usual medical treatment, n (%) (N = 286)	LNG-IUS, n (%) (N = 285)
Psychiatric	3 (1.0)	–
Overdose	3 (1.0)	–
Renal and urinary	2 (0.7)	1 (0.4)
Incontinence surgery	2 (0.7)	–
Partial cystectomy	–	1 (0.4)
Reproductive system and breast	21 (7.3)	24 (8.4)
Endometrial ablation	2 (0.7)	1 (0.4)
Hysterectomy	23 (8.0)	22 (7.7)
Ovarian cyst removal	–	1 (0.4)
Uterine prolapse repair	–	2 (0.7)
Bartholin's cyst drainage	–	1 (0.4)
Dilatation and cutterage	1 (0.4)	–
Laparoscopy for pelvic pain	1 (0.4)	1 (0.4)
Spontaneous miscarriage	–	1 (0.4)
Egg harvesting	1 (0.4)	
Other surgical and medical procedures	9 (3.1)	11 (3.9)
Blood transfusion	1 (0.4)	1 (0.4)
Breast reduction	1 (0.4)	2 (0.7)
Hip replacement	–	1 (0.4)
Tonsillectomy	–	1 (0.4)
Gastric band surgery	2 (0.7)	1 (0.4)
Cosmetic surgery	1 (0.4)	1 (0.4)
CT/MRI scan	2 (0.7)	–
Fracture repair	1 (0.4)	2 (0.7)
Back operation (not specified)	1 (0.4)	–
Abscess drainage following procedure	–	3 (1.0)
Tension of lower bowel	1 (0.4)	
Polyp removal	1 (0.4)	
Death	–	1 (0.4)

CT, computerised tomography; MRI, magnetic resonance imaging; SAE, serious adverse event.
a All listed were classified as SAEs as they resulted in hospitalisation (apart from the recorded death).
b $p = 0.32$.
c $p = 0.48$ (seven women had two SAEs each in the usual medical treatment group; one woman had three SAEs in the LNG-IUS group, while another had two SAEs).
d All women had ceased treatment prior to becoming pregnant.
Nine women experienced two SAEs. Chest pains/shortness of breath and suspected viral meningitis ($n = 1$); two heart attacks ($n = 1$); cholecystectomy and cholelithotomy ($n = 1$); recurrent abdominal pains and non-specific headache ($n = 1$); childbirth – post-partum recovery and overdose ($n = 1$); incontinence surgery and hysterectomy ($n = 1$); hysterectomy and laparoscopy for pelvic pain ($n = 1$); and hysterectomy and a blood transfusion ($n = 1$). One woman experienced three events: two strokes and a hip replacement.

Appendix 12 Intermediate outcomes for the usual medical treatment group

Month	Usual treatment, n	Well with usual, n	Symptomatic with usual treatment, n	Change to LNG-IUS, n	Well with LNG-IUS, n	No treatment, n	Surgery, n	Post surgery, n
Month 1	285	0	0	0	0	0	0	0
Month 2	0	64	211	2	0	7	0	0
Month 3	0	64	200	7	2	11	0	0
Month 4	0	64	190	10	6	13	1	0
Month 5	0	64	179	11	13	16	0	1
Month 6	0	64	174	8	17	17	1	1
Month 7	0	64	158	15	19	23	1	2
Month 8	0	64	147	14	23	30	1	3
Month 9	0	64	137	15	26	36	0	4
Month 10	0	64	128	17	28	40	1	4
Month 11	0	64	125	16	30	41	1	5
Month 12	0	64	121	14	31	44	2	6
Month 13	0	64	114	14	32	49	0	8
Month 14	0	64	107	9	35	55	3	8
Month 15	0	64	102	11	35	56	1	11
Month 16	0	63	93	14	36	59	2	12
Month 17	0	63	88	13	39	59	3	14
Month 18	0	63	84	12	41	62	0	17
Month 19	0	63	78	11	41	64	4	17
Month 20	0	63	70	14	41	65	4	21
Month 21	0	63	66	13	42	68	1	25
Month 22	0	63	62	16	42	68	1	26
Month 23	0	63	57	14	45	71	1	27
Month 24	0	63	54	16	46	71	0	28
Month 25	0	63	48	16	47	73	3	28
Month 26	0	63	45	15	49	75	0	31
Month 27	0	63	43	14	50	77	0	31
Month 28	0	63	42	15	50	77	0	31
Month 29	0	63	42	15	51	77	0	31
Month 30	0	63	40	15	52	78	0	31
Month 31	0	63	39	13	53	79	1	31
Month 32	0	63	37	13	53	81	0	32
Month 33	0	63	36	13	53	81	1	32
Month 34	0	63	36	12	53	81	1	33

Month	Usual treatment, *n*	Well with usual, *n*	Symptomatic with usual treatment, *n*	Change to LNG-IUS, *n*	Well with LNG-IUS, *n*	No treatment, *n*	Surgery, *n*	Post surgery, *n*
Month 35	0	63	36	11	53	82	0	34
Month 36	0	63	36	11	53	81	0	34
Month 37	0	63	36	11	53	80	0	34
Month 38	0	63	36	10	53	81	0	34
Month 39	0	63	36	9	53	82	0	34
Month 40	0	63	34	10	53	83	0	34
Month 41	0	63	33	10	53	84	0	34
Month 42	0	63	31	9	53	85	2	34
Month 43	0	63	29	7	53	87	0	36
Month 44	0	63	26	10	53	86	0	36
Month 45	0	63	22	7	56	90	0	36
Month 46	0	63	21	6	57	91	0	36
Month 47	0	63	18	6	57	93	1	36
Month 48	0	63	17	6	57	94	0	37
Month 49	0	63	15	5	57	97	0	37
Month 50	0	63	13	5	57	99	0	37
Month 51	0	63	12	6	57	99	0	36
Month 52	0	63	10	5	58	100	1	36
Month 53	0	63	9	5	58	100	1	37
Month 54	0	63	8	5	58	101	0	38
Month 55	0	63	6	3	59	102	2	38
Month 56	0	63	5	2	59	103	1	40
Month 57	0	63	5	2	59	103	0	41
Month 58	0	63	5	1	59	103	0	41
Month 59	0	63	4	1	59	102	0	41
Month 60	0	63	1	2	60	103	0	41

Appendix 13 Intermediate outcomes for the levonorgestrel-releasing intrauterine system group

Month	LNG-IUS, n	Well with the LNG-IUS, n	Symptomatic with the LNG-IUS, n	Change to usual, n	Well with usual, n	No treatment, n	Surgery, n	Post surgery, n
Month 1	285	0	0	0	0	0	0	0
Month 2	0	143	111	19	0	12	0	0
Month 3	0	143	104	10	13	14	1	0
Month 4	0	143	100	11	15	14	1	1
Month 5	0	143	98	9	18	15	0	2
Month 6	0	143	95	10	18	17	0	2
Month 7	0	143	92	10	19	18	1	2
Month 8	0	143	88	10	19	20	2	3
Month 9	0	143	83	10	19	24	1	5
Month 10	0	143	80	11	20	24	1	6
Month 11	0	143	78	11	21	24	0	7
Month 12	0	143	77	9	22	25	1	7
Month 13	0	143	70	10	22	27	3	8
Month 14	0	143	66	11	23	28	1	11
Month 15	0	143	62	9	24	32	1	12
Month 16	0	143	60	8	24	33	2	13
Month 17	0	143	60	7	24	32	2	15
Month 18	0	143	55	9	24	33	2	17
Month 19	0	143	54	8	25	32	2	19
Month 20	0	143	49	7	25	34	4	21
Month 21	0	143	47	9	25	34	0	25
Month 22	0	143	43	9	26	36	1	25
Month 23	0	143	41	8	27	37	1	26
Month 24	0	143	38	7	27	39	2	27
Month 25	0	143	38	6	27	40	0	29
Month 26	0	143	36	7	27	41	0	29
Month 27	0	143	35	7	27	42	0	29
Month 28	0	143	35	7	27	42	0	29
Month 29	0	143	35	6	27	42	1	29
Month 30	0	143	34	6	27	42	1	30
Month 31	0	143	34	6	27	42	0	31
Month 32	0	143	30	8	27	43	1	31
Month 33	0	143	30	7	28	43	0	32
Month 34	0	143	29	7	28	44	0	32

Month	LNG-IUS, *n*	Well with the LNG-IUS, *n*	Symptomatic with the LNG-IUS, *n*	Change to usual, *n*	Well with usual, *n*	No treatment, *n*	Surgery, *n*	Post surgery, *n*
Month 35	0	143	28	8	28	44	0	32
Month 36	0	143	28	7	29	44	0	32
Month 37	0	143	28	7	29	43	0	32
Month 38	0	143	27	8	29	43	0	32
Month 39	0	143	27	6	30	44	0	32
Month 40	0	143	25	6	30	46	0	32
Month 41	0	143	23	6	30	48	0	32
Month 42	0	143	21	5	30	49	1	32
Month 43	0	143	17	4	30	53	0	33
Month 44	0	143	15	3	31	55	0	33
Month 45	0	143	15	3	31	55	0	33
Month 46	0	143	15	3	31	54	0	33
Month 47	0	143	12	4	31	55	1	33
Month 48	0	143	12	3	32	54	0	34
Month 49	0	143	11	2	32	54	1	34
Month 50	0	143	10	1	32	56	0	35
Month 51	0	143	9	0	32	56	1	35
Month 52	0	143	8	1	32	56	0	36
Month 53	0	143	8	0	33	56	0	36
Month 54	0	143	6	0	33	56	2	36
Month 55	0	143	5	0	33	56	1	38
Month 56	0	143	3	0	33	58	0	39
Month 57	0	143	2	0	33	59	0	39
Month 58	0	143	2	0	33	59	0	39
Month 59	0	143	2	0	33	59	0	39
Month 60	0	143	1	0	33	60	0	39

Appendix 14 The ECLIPSE trial collaborators

Names in bold indicate the principal investigator for that centre.

Aberdeen Street Medical Centre, Birmingham

RP Kulshrestha and **S Kulshrestha**.

Albany House Medical Centre, Wellingborough

K O'Brien, W Coulson, **J Craig**, L Descals, A Kownacki and **E Montague**.

Ash Lodge Medical Centre, Chesterfield

B McKenzie and **CA Spooner**.

Ashby Health Centre, Ashby-de-la-Zouch

N Sanganee.

Ashfield Surgery, Sutton Coldfield

IP Collier, **SJ Cope**, A Goodwin and LJ Sheldrake.

Bellevue Medical Centre, Birmingham

C Hanly and **R Rajesh**.

Birmingham Heartlands Hospital, Birmingham

S Hutchon.

Birmingham Women's Hospital, Birmingham

T Bingham, **JK Gupta**, J Ingram, L Ingram and H Soneja.

Bluebell Medical Centre, Sheffield

R Pettinger.

Brailsford Medical Centre, Brailsford

C Sadler.

Bridge Street Surgery, Brigg

M Brown, J Chaffey and **V Cowan**.

Bulwell Health Centre, Nottingham

L Brain and **HS Pabla**.

Castle Donington Surgery, Derby

H Godridge, **MEA Horner**, PD Horton, MF McGhee, M McGrath, KR Sumner, GJ Ward-Campbell and S Williams.

Cavendish Surgery, Birmingham

P Madhavan.

Chapel Street Medical Centre, Derby

L Brewin, A Glenchcliffea and **DJ Young**.

Church Lane Surgery, Birmingham

K Heslington, D Hunt and **N Saikia-Varman**.

Church Road Surgery, Birmingham

M Jones, **L Kavi** and L Kidd.

Derby Lane Medical Centre, Derby

J Kai.

Dovecote Surgery, Warley

K Ladha and M Morris.

Dudley Park Medical Centre, Birmingham

K O'Neill, **JN Patel** and JE Waddell.

Eden Court Medical Practice, Birmingham

PG Beighton, FH Cole and RM Edwards.

Enki Medical Practice, Birmingham

A Thompson.

Fernley Medical Centre, Birmingham

J Ayres, J Davies, **TL Poltock**, M Smith and H Talbot.

Four Oaks Medical Centre, Birmingham

AJ Cuthbert and S Horwell.

Frankley Health Centre, Birmingham

C Jenkins, **G Sinha** and K Weaver.

Grange Hill Surgery, Birmingham

B Goodwin and **RP Patel**.

Granton Surgery, Birmingham

SE Debenham.

Hall Green Health, Birmingham

AD Cartmill, S Parker, E Pennington, SJ Pennington and GB Young.

Handsworth Wood Medical Centre, Birmingham

S Butler, PA Hamilton and L Warhad.

Hannage Brook Medical Centre, Wirksworth

J Elliott, **E Lawrence** and J North.

Harborne Medical Practice, Birmingham

CI Elliott.

Hawkesley Health Centre, Birmingham

S Jerome and **PA Shipman**.

High Street Surgery, Birmingham

R Bagchi and N Sheerin.

Hollybrook Medical Centre, Derby

C Blackwall, V Divers and J Siddall.

Jiggins Lane Medical Centre, Birmingham

HJ Parle and S Read.

Keynell Covert Surgery, Birmingham

SK Patodi.

Kings Mill Hospital, Nottingham

S Amis and S Moloney.

Lea Road Medical Practice, Wolverhampton

HH Leung, HU Obi and M Sidhu.

Leach Heath Medical Centre, Birmingham

N Ahmed, S R Carter, W Davies, H Harvey, **K Khanna**, PF Rice, **L Shapiro** and J Stone.

Limes Medical Centre, Alfreton

T Parkin.

Lister House, Derby

JM Craig, PJ Moss and V Steele.

The Goldenbrook Practice, Nottingham

JE Cartwright.

Lordswood House, Birmingham

MG Edward and **C Harrison**.

Meadowfields Practice, Derby

R Francis and P Hassall.

Northgate Practice, Aldridge

CE Lloyd.

Parkfield Medical Centre, Birmingham

J Haynes, **S Budh-Raja** and VP Budh-Raja.

Prestwood House Surgery, Derby

SP Patel and **JH Redferne**.

Queens Medical Centre, Nottingham

S Deb, A Lancaster, **MC Powell** and J Wragg.

Richmond Medical Practice, Solihull

DC O'Brien.

Ripley Medical Centre, Ripley

S Milner and DW Taylor.

Riverlyn Medical Centre, Nottingham

D Davies, **AK Tangri** and C Tangri.

Royal Derby Hospital, Derby

S Amer, G Prileszky, C Warlow and P Whatmough.

Saffron Group Practice, Leicester

CH Duncan.

Solihull Hospital, Solihull

W Bower and **S Irani**.

Springfield Medical Practice, Birmingham

S Rajput and VK Rajput.

St Mary's Surgery, Ely

K Young.

Staffa Health, Alfreton

RE Cooper.

Staffordshire General Hospital, Stafford

K Chin, D Scott and J Stacey.

The Castle Practice, Birmingham

C Dadd and **KG Joshi**.

The Colston Health Centre, Birmingham

K Sundar.

The Hawthorns Surgery, Sutton Coldfield

RM Broomhead, MP Clarke, **E Sangha** and J Wood.

The Karis Medical Centre, Birmingham

B Fitzgerald-Jones and **GR Harley-Mason**.

The Kings Norton Surgery, Birmingham

AD Coward, CM Coward and **J Pash**.

The New Sheepmarket Surgery, Birmingham

HM Little, AG Macdonald and JP Williams.

The Old Priory Surgery, Birmingham

J Marshall, H McBride and **KA Watson**.

The Poplars Surgery, Birmingham

SN Clay and **SA Walker**.

The Southgate Surgery, Birmingham

J Billington, H Griffiths and **R Mehta**.

The Spring Health Centre, Chesterfield

CJ Emslie.

Tower Hill Medical Centre, Birmingham

SS Bath, W Chiam, KR Gordon, M Latthe and DK Nandi.

University Medical Practice, Birmingham

VA Lilford and G Smith.

Victoria Health Centre, Nottingham

E Godfrey, S Griffiths and **C Porter**.

Wake Green Surgery, Birmingham

MS Baird, C Crombie, D O'Donnell and **V Horton**.

Wellside Medical Centre, Derby

RE Kelsey.

Wilson Street Surgery, Derby

JD Fletcher, F Hall and C Lenton.

Woodlands Road Surgery, Birmingham

MP Allen.

Wychall Lane Surgery, Birmingham

J Heanue and **JH McDonnell**.

Yardley Medical Centre, Birmingham

J Hancock and **BL Pattni**.

Yardley Wood Health Centre, Birmingham

ST Banerjee and S Pate.

**EME
HS&DR
HTA
PGfAR
PHR**

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library