Human Reproduction Open, pp. 1–12, 2020
doi:10.1093/hropen/hoaa012

human reproduction open

ORIGINAL ARTICLE

Surgery for women with endometrioma prior to *in vitro* fertilisation: proposal for a feasible multicentre randomised clinical trial in the UK

Abha Maheshwari[1,*], Jemma Healey[2], Siladitya Bhattacharya[3], Kevin Cooper[4], Lucky Saraswat[4], Andrew W. Horne[5], Jane Daniels[6], Suzanne Breeman[7], Kate Brian[8], Gwenda Burns[9], Jemma Hudson[10], and Katie Gillies[10]

[1]Aberdeen Fertility Centre, NHS Grampian, Aberdeen AB25 2ZL, UK [2]Health Service Research Unit, University of Aberdeen, Aberdeen AB25 2ZD, UK [3]School of Medicine, Medical Sciences and Nutrition, University of Aberdeen, Aberdeen, UK [4]Gynecology, NHS Grampian, Aberdeen AB25 2ZN, UK [5]The Queen's Medical Research InstituteEdinburgh, EH16 4TJ, UK [6]Faculty of Medical & Health Sciences, Nottingham, NG7 2UH, UK [7]Clinical Trials Unit, Health Services Research Unit, University of Aberdeen, Aberdeen AB25 2ZD, UK [8]Women's Voices, Royal College of Obstetricians and Gynaecologists, London, UK [9]Infertility Network, London, UK [10]Health Services Research Unit, University of Aberdeen, Aberdeen AB25 2ZD, UK

*Correspondence address. E-mail: abha.maheshwari@abdn.ac.uk https://orcid.org/0000-0002-3652-2447

Submitted on November 23, 2019; resubmitted on January 15, 2020; editorial decision on February 10, 2020

STUDY QUESTION: Is it feasible to undertake a randomised controlled trial to establish whether surgical removal of endometrioma or not, improves live birth rates from IVF?

SUMMARY ANSWER: A randomised controlled trial (RCT) comparing surgery versus no surgery to endometrioma prior to IVF is only feasible in UK if an adaptive rather than traditional study design is used; this would minimise resource wastage and complete the trial in an acceptable time frame.

WHAT IS KNOWN ALREADY: There is wide variation in the management of endometriomas prior to IVF, with decisions about treatment being influenced by personal preferences.

STUDY DESIGN, SIZE, AND DURATION: This was a mixed-methods study consisting of an online survey of clinicians, a focus group and individual interviews with potential trial participants.

PARTICIPANTS/MATERIALS, SETTING, METHODS: Endometriosis and fertility experts across the UK were invited to participate in an online anonymised questionnaire. Potential future trial participants were recruited from a tertiary care fertility centre and invited to participate in either individual interviews or focus groups.

MAIN RESULTS AND THE ROLE OF CHANCE: Clinicians and potential trial participants confirmed the need for an RCT to inform the management of an endometrioma prior to IVF. There were 126 clinicians who completed the survey, and the majority (70%) were willing to recruit to a trial. Half of those who responded indicated that they see at least 10 eligible women each year. The main barriers to recruitment were waiting lists for surgery and access to public funding for IVF. One focus group ($n = 7$) and five interviews were conducted with potential trial participants ($n = 3$) and their partners ($n = 2$). The findings from these discussions highlighted that recruitment and retention in a potential RCT could be improved by coordination between IVF and surgical services such that an operation does not delay IVF or affect access to public funding. Live birth was considered the most important outcome with an improvement of at least 10% considered the minimum acceptable by both patients and clinicians.

LIMITATIONS, REASONS FOR CAUTION: This feasibility study captured views of clinicians across the UK, but as patients were from a single Scottish centre, their views may not be representative of other areas with limited public funding for IVF.

WIDER IMPLICATIONS OF THE FINDINGS: There is a need for an appropriately powered RCT to establish whether or not surgical treatment of endometrioma prior to IVF improves live birth rates. There are logistical issues to be considered due to limited number of participants, funding of IVF and waiting times. These could be overcome in a RCT by using an adaptive design which would include a prospectively planned opportunity for modification of specified aspects of the study design based on interim analysis of the data, coordination of IVF treatments and endometriosis surgeries and international collaboration. Similar principles could be used for other questions in fertility where a traditional approach for randomised trials is not feasible.

STUDY FUNDING/COMPETING INTEREST(S): Funding was received from the NHS Grampian R&D pump priming fund (RG14437-12). S.B. is Editor-in-Chief of HROPEN, and A.W.H. is Deputy Editor of HROPEN. Neither was involved in the review of this manuscript. L.S. reports grants from CSO and NIHR to do endometriosis research, outside the submitted work. K.C. reports grants from NIHR/HTA and CSO during the conduct of the study. J.H.e., A.W.H., J.D., S.B.r., K.B., G.B., J.H.u. and K.G. report no conflict of interest.

Key words: endometrioma / endometriosis / IVF / live birth / surgery

WHAT DOES THIS MEAN FOR PATIENTS?

Endometriosis is a condition where tissue similar to the womb lining is found in other parts of the body. It can cause fertility problems, and women may need treatments such as *in vitro* fertilisation (IVF). About a third of women with endometriosis have cysts called endometriomas which are often removed before starting IVF to try to improve the chances of success. It is possible, however, that removing endometriomas could make IVF less likely to work. Researchers want to carry out a trial to see whether removing endometriomas before IVF is a good idea. They consulted doctors and patients and realised it would be difficult to get enough people to take part in a traditional trial as IVF funding is often limited and it would mean delays in treatment. Instead of a traditional trial, they would rather use a more flexible type of study where you can start with small numbers of participants and adapt the study from what you learn as you go. It means you can change what you expect to find as you get more evidence. This paper proposes using this kind of trial to work out whether surgery for endometriomas is beneficial.

Introduction

Endometriosis affects 1 in 10 women of reproductive age, 30–40% of whom experience infertility (Redwine, 1999) and 17–44% have an endometrioma. “”Many of these women require *in vitro* fertilisation (IVF) for treatment of their infertility.

Surgical removal or ablation of an endometrioma prior to IVF has traditionally been believed to prevent exacerbation of disease, reduce the dose of gonadotrophins required for ovarian stimulation, reduce the risk of infection during oocyte retrieval and bring down treatment costs. Removing an endometrioma can also provide a histological diagnosis which helps to rule out rare cases of malignancy. It is also possible that the procedure may also improve the chances of spontaneous conception.

Data from more recent studies suggest that surgery (and the need for postsurgical recovery) for endometrioma can delay IVF treatment. Additionally, some reports suggest that surgical damage to ovarian tissue can compromise ovarian reserve (Raffi *et al.*, 2012a, Somigliana *et al.*, 2012) leading to fewer oocytes and lower success rates. This is disputed by other studies which have failed to show a similar effect (Muzii *et al.*, 2014).

Consequently, there is considerable uncertainty at the present time (Dunselman *et al.*, 2014; Hamdan *et al.*, 2015) about whether operating on an endometrioma prior to IVF is beneficial or not. A single randomised controlled trial ($n = 99$) has failed to show improved pregnancy rates after surgery (Demirol *et al.*, 2006), but surgical treatment of endometriosis and endometrioma(s) prior to IVF continues to be widely practiced (Raffi *et al.*, 2012b), despite its attendant risks and costs.

Improving fertility in women with endometriomas is one of the top 10 priorities identified by the James Lind Alliance priority setting partnership for endometriosis (Horne *et al.*, 2017). Following repeated calls from clinicians, patients and policymakers to address this uncertainty, a well-designed adequately powered, randomised controlled trial (RCT) evaluating the effectiveness of surgery on endometrioma prior to IVF is overdue. We searched various trial registries (ISRCTN; ANZCTR; ChiCTR) for ongoing/proposed trials; only one trial was found registered in 2008 (ISRCTN35880386) which was stopped after a year (personal communication (T Gelbaya). This was conducted unfunded in a single centre. Another protocol was published recently (Tomassetti *et al.*, 2018), but the trial was never conducted. The major challenges to date are the complexity of designing and conducting a successful trial and limited numbers of eligible participants.

These illustrate that it is a difficult trial to recruit and traditional trial designs may not be appropriate. We conducted a study to determine a feasible design of a future multicentre RCT of surgery for endometrioma prior to IVF in UK.

Materials and Methods

In order to assess the feasibility of an RCT, we undertook a survey of clinicians' views and conducted a focus group, followed by individual interviews with potential trial participants. It was anticipated that the combined views from these two groups would provide information on key components of a future trial including the choice of participants, intervention, comparison and outcomes, sample size, likely accrual rate and duration of trial and optimum trial design to enhance recruitment and retention.

Clinicians' survey

SurveyMonkey™ software was used to create an online survey for completion by fertility and endometriosis specialists. Appropriate steps were taken to construct this questionnaire such as relevant wording,

sequencing and consideration of time taken to complete it. A simple self-administered online survey was constructed to assess: minimal non-identifying baseline demographic data; prevalence and current practices of managing endometrioma in women undergoing IVF; treatment preferences using four clinical scenarios each describing two alternative approaches to endometrioma prior to IVF (drainage of endometrioma and ablation of base versus excision, complete excision of endometrioma versus partial excision, aspiration of endometrioma under ultrasound guidance versus aspiration at laparoscopy or use of medical treatment using a GnRH agonist before and after surgery versus no GnRH agonist); willingness to recruit to a trial of pre-stated/any clinical scenarios; barriers to recruitment and retention in a trial; and preferred primary outcome.

There was space for free comments at the end to capture any further thoughts not covered in the above sections. The questionnaire was peer-reviewed by two experts to assess content validity and readability. It was piloted amongst five experts who were advised to 'think aloud' as a cognitive probe in the presence of a researcher to assess the face validity of the questions.

Clinicians were asked to quantify current levels of clinical certainty on a 5-point linear scale (definitely prefer A, may prefer A, undecided, may prefer B, definitely prefer B) about alternative treatment approaches (with a variety of available techniques and approaches to treat endometrioma) in controversial clinical scenarios. This was to assess and quantify the strength of their preferences for treatment approaches and assess equipoise at the level of the clinician (individual equipoise) and within the clinical community (community equipoise) (Young *et al.*, 2006).

Administering the questionnaire

An email containing a link to the questionnaire with a covering letter was sent to two groups of clinicians (fertility specialists and endometriosis experts). The lead clinician of each centre in the UK was sent a personalised email with a covering letter and a link to the survey with request to forward the survey to all clinicians in their unit. The contact details of each IVF centre are publicly available on the Human Fertilisation and Embryology Authority (HFEA) website. A reinforcement email was also sent through the Reproductive Medicine Clinical Studies Group (RMCSG), North of England Reproductive medicine Group, Scottish Fertility Group and the Principal Investigators of two ongoing clinical trials (E-Freeze: ISRCTN-61225414 and PreEMPT: ISRCTN-97865475). The questionnaire was also sent to all British Society of Gynaecological Endoscopy (BSGE) members via the British Society of Gynaecology Endoscopy (BSGE) secretary.

Focus group and individual interviews

A focus group and individual interviews were conducted with women (and where appropriate their partners) to explore trial acceptability amongst potential participants, investigate views on trial recruitment and retention and identify outcomes most relevant to them. Women who were potentially eligible to participate in the proposed trial were identified from the Aberdeen Fertility Centre (a tertiary centre for the North of Scotland). Women and partners were invited to take part in the feasibility study by the treating clinicians and were sent an invitation pack. A topic guide for both the interviews and the focus group were developed in consultation with patient and partners. It explored the following areas: background knowledge; willingness to take part in an RCT; barriers to research participation (specifically to an RCT); retention in proposed trial; preferred outcomes; and presentation of information about the RCT. Interviews and the focus group were audio-recorded and transcribed verbatim.

Ethical approval was obtained by North of Scotland Research Ethics committee (Ref number—IRAS: 245259; R&D Reference—RG14437-12).

Data analysis

Clinicians' survey

Data were summarised descriptively using numbers and percentages. Responses between the types of specialists were compared using Fisher's exact test as the expected frequencies were less than five. Analyses were performed in Stata 15 software (StataCorp 2017. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC).

Focus group and individual interviews

Data were coded and analysed using a Template Analysis variation of thematic analysis to incorporate the inductive and deductive nature of the interviews (Brooks *et al.*, 2015). The initial thematic framework focused on practical aspects for trial design and delivery. This was developed by J.H. then reviewed by K.G. for consistency and further refinement. This framework was then applied consistently across all transcripts.

As this was a feasibility study a formal sample size calculation was not done.

Results

Clinicians' survey

There were 126 responses to the questionnaire, of which 59 (47%) were from endometriosis specialists and 67 (55%) were from fertility specialists. Most offered both publicly (NHS) funded as well as privately funded treatment.

The annual number of women with pre IVF endometriomas cared for by each clinician varied: 22 (17%) reported seeing 8–10 women, 66 (52%) saw more than 10 women, and the remainder managed fewer than 8 women per year. The responses from fertility and endometriosis experts were similar.

The diameter of an endometrioma below which the specialists would withhold surgery also varied, with the majority choosing 3 cm (29%) or 4 cm (25%) as cut-off values. There were some differences between the two groups of gynaecologists, as fewer endometriosis specialists (22% versus 13%) had a size threshold. In the free comments section, several endometriosis specialists mentioned that they acted as technicians at request of their IVF colleagues who actually made the decision to operate. Most specialists stated that they would prefer to proceed straight to IVF rather than surgery in women over the age of 40.

Current preferences and recruitment to a trial

To explore current practices (to determine equipoise) and identify the interventions to compare in a trial, specific questions were asked (Table I) for the following interventions. There was no difference between the two groups of specialists in terms of current practices or willingness to recruit to an RCT comparing the two interventions.

Table I Current preferences and recruitment to proposed trial on surgery to endometrioma prior to IVF.

	Current preferences						Recruitment to trial				
	Definitely prefer A	May prefer A	Undecided	May prefer B	Definitely prefer B	No response	Definitely yes	Maybe yes	Undecided	Maybe no	Definitely no	No response
Laparoscopic treatment (A) or transvaginal aspiration (B)	101(80)	13(10)	3(2)	2(2)	1(1)	6(5)	34(27)	14(11)	21(17)	13(10)	26(21)	18(14)
Laparoscopic excision (A) or laparoscopic drainage and ablation of base (B)	59(47)	29(23)	9(7)	12(10)	11(9)	6(5)	40(32)	29(23)	18(14)	8(6)	12(10)	19(15)
Laparoscopic complete excision (A) or Laparoscopic partial excision (B)	60(48)	34(27)	7(6)	11(9)	7(6)	7(6)	31(25)	20(16)	27(21)	12(10)	16(13)	20(16)
Pre-operative GnRHagonist (A) or no pre-operative GnHagonist (B)	13(10)	14(11)	29(23)	24(19)	39(31)	7(6)	35(28)	24(19)	24(19)	10(8)	14(11)	19(15)
Post-operative GnRHagonist (A) or no post-operative GnRHagonist (B)	26(21)	25(20)	23(18)	19(15)	26(21)	7(6)	39(31)	27(21)	22(17)	6(5)	13(10)	19(15)

Values are *n* (%)

Laparoscopic treatment or transvaginal aspiration. Most respondents ($n = 114$, 90%) preferred laparoscopic treatment when compared to transvaginal aspiration with no difference between the two group of specialists (88 vs 92%, *P* value 0.529). For randomising participants to an RCT comparing these two interventions, 38% said yes and 31% said no with 17% being undecided.

Laparoscopic excision or laparoscopic drainage and ablation of base. There were 88 (70%) who preferred laparoscopic excision. For recruiting to an RCT, 70% said yes with no difference between the specialists (*P* value 0.510).

Laparoscopic complete excision or laparoscopic partial excision. Three quarters of respondents ($n = 94$, 75%) preferred laparoscopic complete excision. There was a mixed response for randomising participants to an RCT, with 41% saying yes, 23% saying no and 21% were undecided.

Pre-operative GnRH agonist or no pre-operative GnRH agonist. Half of the respondents preferred no pre-operative GnRH agonist. For recruiting to a trial on this, 47% would participate, 19% would not participate and 24% were undecided.

Post-operative GnRH agonist or no post-operative GnRH agonist. The proportions of respondents who preferred to use post-operative GnRH agonist (41%) treatment were similar to those who did not (36%). Over half of the respondents (52%) were willing to recruit participants to an RCT comparing post-operative versus no post-operative GnRH agonist.

Willing to randomise to any trial comparing treatments to endometrioma. When asked if they would be willing to randomise to any trial comparing treatments to endometrioma, 62.4% indicated that they would, while 20% were undecided.

Primary outcome

Clinicians in both groups were asked to choose a preferred primary outcome out of clinical pregnancy, live birth or healthy baby rate. Most agreed that live birth rate was the preferred primary outcome (61%). Just over half the respondents (55%) agreed that >10% improvement in the primary outcome would be acceptable in the context of an RCT, 7% indicated that it should be >15% whereas 25% were content with a 5–7% improvement. According to the clinicians, the proportion of eligible women who might be willing to participate in an RCT varied from less than 25% to 50–75%. A third thought that women would be willing to be followed up.

Identified barriers for recruitment and retention

When asked about barriers to recruitment and retention in the trial, the concerns raised by both group of clinicians in free text were: NHS waiting lists for surgery leading to delay in accessing IVF; different specialists carrying out surgery and IVF; and availability of public funding for IVF. Endometriosis surgeons specifically advised that decision is usually made by fertility clinicians.

Focus group and individual interviews

A total of 12 participants (nine women and three men) were recruited to participate in individual interviews or focus groups between October and December 2018. It was important to involve couples as any infertility treatment involves both partners. Three individual interviews were with women, one was a joint interview with both partners

and one focus group (consisting of five women and two men) was conducted. Interviews lasted between 60 and 120 min.

Six main themes emerged from the exploratory work with potential trial participants. Each of these key areas is discussed in detail below.

Acceptability of the research question

All participants felt that the research question 'surgery or no surgery for endometrioma prior to IVF' was important.

P2 'Because if the doctors don't know and you don't know then there's no real . . . you can't really make a decision because you don't know what's going on really do you. So the more information you have the more evidence there is, you're able to make better choices.'

Participants acknowledged that the requirement for this research was to improve treatment options and improve decision-making in this context. It was accepted that participating in a clinical trial was 'a gamble' but noted that high-quality data was needed to enable patients to make well-informed choices in the future.

P7 'For future people going through this because I think once you've been through anything first hand you have empathy for anybody who is going to go through it in the future. So if, I suppose if we, or if I can be involved in something that is going to bring somebody up another step towards that goal then yes, I would be involved.'

P11 'But like I said, if I could help in any kind of study then I would, just because . . . I just would. Yeah, I just would.'

Timing of the invitation to participate in trial

There was consensus amongst all patients that introduction and recruitment to a trial should be as close to diagnosis of endometrioma (and any recommendation for IVF) as possible. The timeliness of this approach was considered important as the process of fertility investigations and treatments can become stressful and overwhelming very quickly, leaving little reserve for consideration of trial participation. Inviting women at diagnosis would foster feelings of 'doing something to support their IVF' as well as build on the momentum of developing a treatment plan, being active participants rather than passive recipients of care.

P1 'Yeah, as soon as you know that you're going to get IVF, strike while the iron is hot. Let couples think about it from the beginning, "We could be a part of a trial, we don't know what side we're going to be on but it's going to be a trial, we know all the options.'

P7 'Yeah, I agree, probably at that start, at the first scan . . . because you're still full of zip and you know, you think, "Okay, great, yeah, okay, we're going to do that, yeah, I'll be part of that, that's fine.'

Women identified that they expect decisions (about treatment and trial participation) to be made in conjunction with their partners. Partners expressed the view that that even in the unlikely event they disagreed about trial participation they would support the woman's preferences.

Waiting times for surgery

Participants had mixed feelings about waiting times about surgery, especially in areas where they had to wait for IVF as well. Participants asserted that 'knowing what would happen and when' was of greater importance than delays from waiting lists. It is noteworthy that patients suggested an upper threshold of 3 months for any additional wait time on the trial. Their reservations would likely be resolved through clear and open communication about why this delay existed/was necessary.

P3 'I think what I was saying, if you're telling me today, "Hello, you can be part of . . . if you're not part of this trial you're waiting a year to have IVF because there's the waiting list", I would say, "Okay, this is the waiting list in NHS', 'If you take part in this trial", if you're not going to make my year longer I will take part you know, or you make it longer by a few months, you know what I mean.'

P1 [standardisation of wait times] 'I think you would get way more people to come on board, not that they know what one they're going to get, way more people would know that within six months "I've had my operation, recovered and I'm going through IVF" and the other people is "I'm going straight into IVF, I'll hopefully have a baby at the end of nine months and then I'll have my operation" Yeah, I think you'd get more people.'

P5 'The brackets for IVF are sort of under 35, 35 to 38, 38 and above or thereabouts. So anything that goes within six months to a year that already moves you from one bracket, potentially to one bracket to the next, so that would be too much I think. Anything that goes over you know, around half a year or so.'

P10 'So yeah, I don't know. I think it [standardising wait times across all trial centres] would help you to maybe recruit more people.'

P11 'Once you get to the stage that we're at you don't care! The E-Freeze thing as well, when we signed up for it we were thinking we've waited so long anyway, what's another couple of months or whatever going to . . . you know, what's an extra couple of months? It's nothing really, compared to what we've been through already, it would be silly not to do it.'

P12 'I don't know what it's like for other people but we've had a fairly long journey to this point and so time, I think, is already going to be something that's on people's minds no matter what age you are, just in terms of how long you're going through this stuff. Not necessarily in terms of running out of time for having a baby and so to know that that wouldn't be dragged out any longer than necessary would be a really positive thing, definitely.'

Confidence in the clinical team

All participants identified confidence in their clinical team as a key influencer to participating in a RCT. Most participants felt that a member of the clinical team rather than academic team should introduce the trial. However, being knowledgeable about the trial and being able to answer questions competently would be key for their decision-making about participating. Participants asserted that honest and comprehensive communication from outset about the risks and benefits of treatments including waiting times and deviations from standard pathways, supports them in feeling in control and in making informed decisions.

P10 'I suppose probably the consultant in a way because then you can ask them more medical questions about how it might impact your treatment, and then maybe the researcher because then you could ask practical questions about the research, and then maybe the nurse. But I don't have a . . . I don't think I would have strong, what's the word, preferences. I think as long as I knew who I could speak to if I wanted, if I needed more information.'

P8 'So I think ultimately you have to be extremely clear what you're asking people to do, I think that's probably the key thing at the start.'

P3 'From my point of view it doesn't really matter who gives me the information, if I have questions as long as that person knows what they're talking about. I guess if you're there and you have the scan,

which will be done by a doctor, a consultant, whatever, they'll probably just mention it, "Oh, by the way we have this trial, and would you like to know more information?"'

Additionally, if trial staff were unable to be sensitive to the patients' emotional state and acknowledge the challenging context of balancing a diagnosis with an invitation to participate in a trial, this would also reduce their willingness to participate.

P1 'I think as long as you explain it, you make it all very clear and it's not complicated because when things start getting complicated then people are like, they're not really bothered, they don't want to do it. They just want to get to that end goal'.

P12 'Yeah, I guess it's a sensitivity and an understanding to people's individual situations. And I guess understanding that sometimes phone calls might come at the wrong time, might have to wait, that kind of thing. Conversations . . . maybe kind of a choice that they say, "No, I don't want to have that conversation now but next week might be better," that kind of thing. It's a particularly emotional time.'

Patient characteristics as barriers to participation

The age of the patient may act as a potential delay to treatment and may preclude them from accessing public funding for IVF due to eligibility thresholds. Participants noted that the severity of their condition may influence their decision to participate. They also indicated that additional appointments would be unacceptable, especially when considering associated travel distances and costs.

P1 'If you've got people coming as far as two hours away, I would try to do it, have all of it on one day so they don't have to keep toing and froing.'

P2 'As long as it's easy and its as seamless as possible and doesn't intrude too much . . . when things get a bit difficult is when people tend to pull out... It's when it becomes more complicated if you know what I mean. People like to be easy.'

Trial retention

Potential participants indicated that if they signed up to a trial, they would expect to remain committed to the trial until the end unless there was a medical need to withdraw. Participants reported that any trial in this area must run in parallel to typical clinic sessions for them to consider participation. This continuity in their treatment pathway would support them in feeling that the trial was not intruding on what they already had to do in their 'fertility journey'. Additionally, feeling that participating in the trial improved their access to experts and peer support would be a great incentive to initial decision to participate and for continuation.

P2 'So people might see it as being a benefit I think is what you're trying to say . . . . . people think it's more of a benefit because they think they're going to get more of a specialist interaction maybe.'

P2 'Because once you're invested, once you've signed up to something you've invested in you're quite happy to answer questions because . . . . . if I've joined up to a trial then when that trial finishes is when I stop, not when I want, not when I would want to stop so to speak.'

P10 'I think if you agree to take part in the trial then you've already . . . like I can't see why you would object to being contacted again to get the benefit out of you having taken part. I suppose the only . . . I guess if your pregnancy wasn't successful then maybe you wouldn't want to be speaking to people.'

P11 'No, I just know that if I've ever got a question they're just at the end of the phone.'

Outcomes of importance

Participants confirmed that live birth rate is the main outcome they expect to measure success. Participants found it difficult to gauge the level of improvement required to determine the level of success of either surgery or no surgery before IVF. Estimates ranged between 3 and 30% with 10% being the modal value.

P3 'I guess is . . . obviously will want live birth as the outcome because that's what you want as an individual. Probably if you want to find out whether this works or not you'd be looking more at maybe implantation rate, because you know, if you lose the baby later on, which we don't want, then that might not have nothing to do with your actual . . . actually your surgery, because you did the surgery to try to get more follicles, more eggs, more embryos and you want to implant them, that's probably what the surgery will do.'

P10 'For me I don't think it matters, a small increase, because I think when you're talking . . . like even if we say 15% and 20%, that's massive, I think that's a massive increase in the chances of it being successful.'

P2 'If you said to me "it's only 5%" I'd be like "Well that's not really anything really is it to be honest, it's so small".'

P3 'Oh, yeah, that's definitely significant for me if it's just like 10% increase of . . . '

Answers to specific questions posed by the feasibility study

Based on the clinicians' survey and qualitative work with potential participants, it was clear that a randomised trial is needed.

The factors that would enhance recruitment and retention, identified from this feasibility study could be divided in two groups: generic and specific for this trial. The generic factors were good communication; ownership of decision-making by participants; impact on clinical treatment; and extra travel. Specific factors related to surgery to endometrioma to improve the outcome of IVF were: information about trial early on in the infertility journey rather than at start of IVF; very close liaison between the fertility and endometriosis team so that every single eligible participant is approached for the trial; and a joint pathway to bypass public (NHS) waiting lists for surgery, so that patient care and eligibility for publicly funded IVF is not compromised for those who agree to take part in trial. Bypassing waiting lists could be justified as the number of participants would be small and it is important that surgery is compared to no surgery in ideal circumstances (with no delay) to determine if there is any benefit.

The components of the trial deduced from the feasibility study were: population (women $<$40 years old undergoing IVF with endometrioma of $\geq$3 cm, as below 3 cm and above 40 years of age there was no equipoise); intervention (surgery to endometrioma followed by IVF); comparator (IVF without prior surgery for endometrioma); and outcomes: the primary outcome should be live birth.

Sample size calculation

The recommended improvement in live-birth rates after surgery was at least 10%. Using traditional approach we would need to randomise 954 women for 90% power (two-sided alpha 5%) to detect a 10% increase. For 80% power, we would need 712 women (Table II). This

Table II Sample size calculation for proposed trial on surgery to endometrioma prior to IVF.

Control event rate	Event rate in intervention group	Power	Enrolment ratio	Sample required (total)
30%	40%	80%	1	712
30%	40%	90%	1	952

Table III Number available for recruitment (based on questionnaire) for proposed trial on surgery to endometrioma prior to IVF from within UK.

No. of cases per year from respondents of the questionnaire	All respondents recruiting	70% of all respondents recruiting	Only Fertility specialist recruiting only	70% of fertility specialists recruiting
≥10	670	469	420	294
8–10	136	95.2	88	61.6
4–7	64	44.8	60	42
1–4	13	9.1	6	4.2
Total	883	618.1	574	401.8
50% recruitment	441.5	309.05	287	200.9
25% recruitment	220.75	154.52	143.5	100.45*

*Most pessimistic estimate

was based on assumption of 30% live birth rate (Human Fertilisation and Embryology Authority, 2019) after one embryo transfer episode in control group.

Likely accrual rate and duration of the trial

To assess the number of potential participants available in each UK centre, we undertook some projections based on the responses from the clinicians' survey (Table III). Using conservative estimates and the most pessimistic scenario, we should be able to recruit a maximum of 100 participants per year in the UK from the respondents' clinics (Table III). Hence, to recruit a minimum of 712 women, a minimum of 7 years will be needed. Given the minimum time needed for setting up the study (6 months) and follow-up to live birth (18 months, 3 months to have surgery, 3 months to have IVF treatment, 9 months for pregnancy and 3 months in case embryos are frozen and thawed and transferred in subsequent cycles), a minimum of 9 years would be needed to complete the RCT. This would increase to a minimum of 11 years, if the sample size was 952. This is clearly not expedient and therefore not feasible. An international collaboration with recruitment from multiple centres across the world would be ideal but is not practical, due to existing patient pathways of care for surgery and IVF in different countries as well as current systems of funding.

Suggested trial design

A limited number (i.e. 240) of eligible women can be recruited within the UK over a reasonable recruitment period (30-month) (Table III). A traditional trial design (Frequentist approach) with sample size calculation incorporating Type I and Type II error will not be feasible in this situation as there is little adaptation permitted and it is a long process. Hence, we need to think of non-traditional ways of doing this trial.

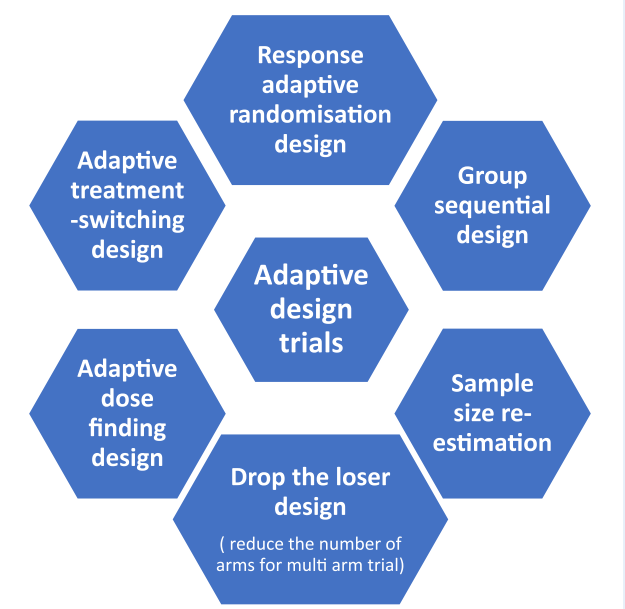

Figure 1 Scope of adaptation following interim analysis in adaptive clinical trial designs.

The literature for rare diseases was explored; however, by definition a disease is called a rare disease when the prevalence is less than 1 in 2000. Endometrioma does not fulfil this criterion as at least 1 in 100 women undergoing IVF will have endometrioma. We also explored the cancer literature for potential trial designs because sample

Table IV Difference in Bayesian and Frequentist approach.

Frequentist	Bayesian
Data varies (calculated on Type I & II errors)	Data is fixed
Calculates *P* values & confidence interval	Calculates credible interval
No previous data ('prior')	Depends on prior and likelihood of observed data
Tends to be computationally less intensive	Computationally intensive due to integration over many parameters

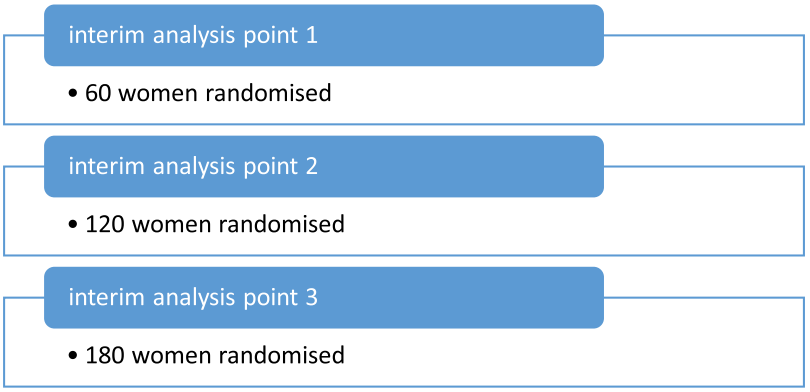

Figure 2 Suggested interim analysis points.

sizes are frequently restricted by the availability of participants. Two potential strategies are commonly used that could also be useful IVF trials in general, but specifically when sample size is restricted: adaptive designs (in particular the group sequential design below) and Bayesians approaches (Hampson *et al.*, 2014). In general, adaptive designs include prospectively planned opportunities for modification of one or more specified aspects of the trial design based on interim analysis of accumulating data. These modifications are made after the initiation of the trial, are usually pre-specified so that the validity and integrity of the study is not undermined. There are several types of adaptive trial designs (Fig. 1).

For our specific scenario the group sequential design has clear advantages. In group sequential trials, there are one or several pre-planned interim analyses of data accumulating when recruitment is still ongoing (Gsponer *et al.*, 2014). At each interim analysis, a decision is made about whether to continue with the trial or stop if early signals in the data suggest a treatment is ineffective, known as stopping for futility. However, a drawback of this approach in a more traditional, or Frequentist, framework is that we pay a penalty for taking multiple looks at the data. This inflates the Type I error, which requires changes to the analysis, but has the consequence of making a trial larger (in general).

A Bayesian, rather than Frequentist, approach is better suited to situations like this. Without mathematical details, a Bayesian approach allows us to incorporate available information into a prior probability distribution, which is called the 'prior'. As data from the trial accumulates, it is combined with the prior to produce a posterior probability distribution, which is known as the 'posterior' (Berry, 2006). The posterior allows us to make intuitive probabilistic statements, for example what is the probability of a specific size of treatment effect given the data in our trial (Gsponer *et al.*, 2014). This is contrast to a *P* value from a Frequentist analysis which gives us the probability of finding the observed treatment effect (or one more extreme) given that our null hypothesis (usually zero treatment effect) is true. Another advantage is that that we pay no penalty for multiple interim looks at the data, which makes Bayesian designs attractive when the number of eligible participants is low. Some of the main differences between the two approaches are given in Table IV.

A critique of Bayesian inference is that often a subjective prior is used. Constructing prior distributions is not trivial, and not unexpectedly different people may have different priors and therefore the same data can deliver different posteriors and conclusions depending on the choice of prior. However, this gets less likely as trial sample size increases. Rather than a weakness, we see this a strength of the Bayesian approach, by using different priors we can explore how sensitive results are to the choice of prior. Although Bayesian and adaptive designs are common in many other fields, we are not aware of their use the field of fertility.

Given our primary outcome of live birth takes time to report, we plan Interim futility analyses using an intermediate (or surrogate) outcome (Goldman *et al.*, 2008; Asakura *et al.*, 2017). Interim futility analyses will be based on clinical pregnancy (pregnancy at 7-week gestation) as a surrogate outcome because it has desirable properties: strongly correlated with and available 8 months earlier than the primary outcome of live birth. We suggest clinical pregnancy as the outcome for interim futility analysis because without achieving clinical pregnancy one cannot have live birth and most pregnancies continue once they reach this stage (Braakhekke *et al.*, 2014). A potential group sequential design is shown in (Fig. 2). In a design like this, we would do three interim

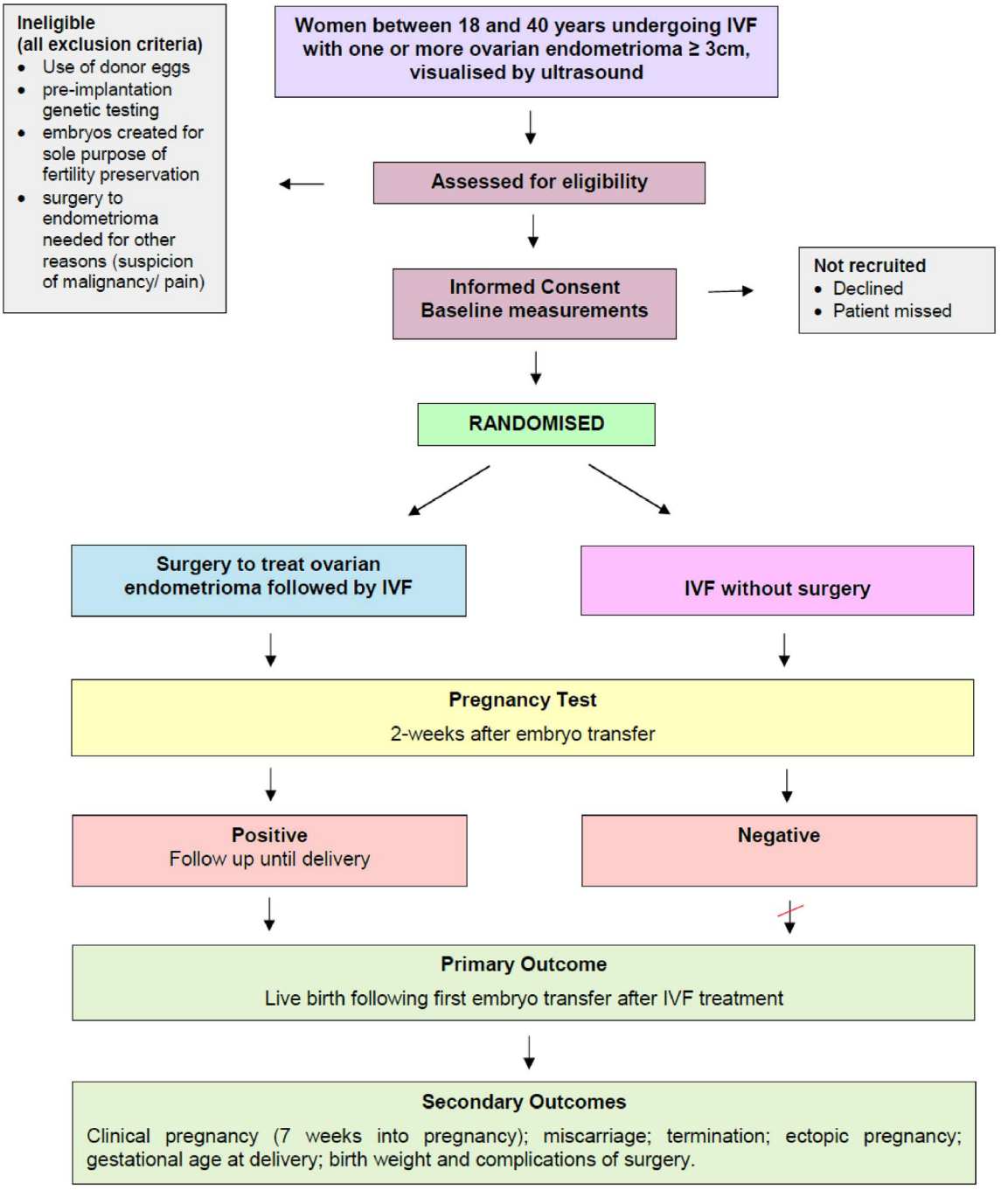

Figure 3 **Suggested trial design (with interim analysis for futility).**

analyses for futility based on clinical pregnancy rates after accrual of 30, 60 and 90 patients per arm at each stage. If the trial was not stopped for futility, then a final analysis for efficacy would take place on the full 240 participants using the primary outcome of live birth. We anticipate a 35% clinical pregnancy rate in the control group (Human Fertilisation and Embryology Authority, 2019; Miller *et al.*, 2019). For surgery to improve live-birth rates by 10%, one would expect to see at least a 12% improvement in CPR i.e. from 35 to 47%, to account for any pregnancy losses (Coomarasamy *et al.*, 2019). Interim futility analyses on CPR will allow us to stop such a trial if probability that the difference in the CPR is less than 12% is 90% or greater. The trial can be stopped as surgery for endometrioma is not likely to provide at least a 10% higher live birth than no surgery. In this way the scarce research resources can be maximised, at the same time a credible answer can be provided. A suggested trial design is illustrated in Figure 3.

Discussion

Principal findings

Our results show that, despite strong preferences, over two-thirds of surveyed clinicians were willing to recruit to an RCT of surgery to endometriomas prior to IVF but were concerned about NHS surgical waiting lists and access to public funding for IVF. Women and their partners in an exploratory qualitative study, as part of feasibility study felt that the research question was important and were willing to participate in an RCT but were anxious to avoid a delay in starting IVF treatment. An improvement in live birth rate of at least 10% was the most desired outcome.

The low numbers of eligible participants pose a major barrier to recruitment to a conventional trial. We have suggested a non-traditional design which requires lower numbers of eligible participants to provide a credible answer in a limited time frame. Hence, we recommend a Bayesian group-sequential design with live birth in an IVF cycle as the primary outcome with interim futility analysis using clinical pregnancy rate.

Strengths and limitations of our study

This is the first study assessing the feasibility of a randomised trial comparing clinical effectiveness of surgery versus no surgery for endometrioma in women prior to undergoing IVF. Our team had representation from fertility specialists, endometriosis specialists, trial methodologists, international experts in the field, qualitative researchers and patient representatives. Our clinicians' survey had equal participation from endometriosis and fertility specialists.

Only 12 patients were interviewed in total. Participants consisted of women and partner: individually and jointly. Being a feasibility study, a formal power calculation was not needed.

This feasibility study had views of clinicians across the UK, but the views of patients were from a small geographical area in Scotland. Scotland has very good public funding for IVF; therefore, views may not be representative of other regions of the UK where funding is limited. We restricted to one centre for recruiting patients as process are similar across Scotland, it was felt that addition of multiple centres for exploratory qualitative study would add to the resources but not the benefit.

Meaning of our findings

This proposed study on surgery to endometrioma prior to IVF is unique in that although endometriosis affects a large proportion, but this specific question is only relevant to a small group of population. It is a treatment where public funding is restricted, outcome depends on multiple factors including male partner and the treatment spans over two specialist areas. This is why undertaking a traditional RCT would be difficult and possibly why it has not been conducted despite continued debate in both fertility and endometriosis the community.

The results from a survey of potential participants were predictable in terms of their views which pertain to any trial such as adequate honest, timely information and extra visits. Their own, and clinicians, views specific to this trial were also predictable in terms of impact of waiting lists and public funding on recruitment and retention in the trial. Very interesting is the minimum acceptable improvement in live birth which was quoted as 10%. There is existing literature in the field of IVF where both patients as well as clinicians would accept a lower improvement in success rates as important (Scotland *et al.*, 2007). A larger improvement here may be due to the invasiveness and complications associated with surgery, along with the delay to IVF. What couples would be willing to trade off to accept invasive nature of the surgery will need to be further evaluated by larger studies and methods such as Discrete Choice Experiments. This would ideally be done alongside the proposed trial.

There is a lack of feasibility studies in the area of Reproductive Medicine even though trials in this field are difficult, particularly the ones where treatment pathway is affected. Feasibility studies represent an attempt to get the question and design right prior to full funding so avoiding the trap of supporting projects that ask the wrong questions or not designed to be successfully completed. There is benefit overall in doing feasibility studies (Morgan *et al.*, 2018). The main purpose of a feasibility study is to 'de-risk' potential full trial funding. Given difficulties in recruiting to randomised trials, we would suggest doing feasibility studies prior to planning an RCT. This will be certainly more attractive to funders (Morgan *et al.*, 2018).

In terms of trial designs in fertility where the number of participants is limited but traditional trial designs require a large number to show a meaningful benefit, it may be that we need to think of adaptive designs to reduce uncertainty in clinical decision-making. Although live birth should be the outcome to prove the effectiveness of the intervention, there are possibilities where clinical pregnancy rates can be used to identify ineffectiveness of expensive and invasive interventions to save scarce research resources.

The Bayesian approach provides a formal framework for researchers to incorporate prior knowledge and current evidence to derive new probabilities for various hypotheses. Since the results are presented in terms of probability, clinicians can interpret and apply research findings to clinical practice directly (Wong and Warren, 2010).

Live birth is suggested as an outcome, but we have to determine the time horizon as well. Ideally, cumulative live birth should be used but there is inconsistency in the definition (Maheshwari *et al.*, 2015) and this will take a long time to calculate. Hence, for initial trial live birth rate after first embryo transfer following randomisation could be used as an outcome. At the same time, follow-up data should be collected to determine cumulative live birth rate.

Conclusion

Clinicians and patients both support an appropriately powered RCT to answer the research question 'Should endometrioma be operated prior to IVF to improve the livebirth rate'? However, due to limited number of participants, complexity of public funding arrangements for IVF as well as waiting lists for surgery, traditional randomised trial designs would not be appropriate. We suggest using a group sequential Bayesian approach with interim analysis for futility for this as well as other trials in field of Reproductive Medicine. This would certainly reduce the uncertainty in decision-making where evidence is lacking currently and trials are difficult to design. We also highlight the need for close cooperation and communication between professionals to ensure recruitment and retention in this challenging study. Feasibility studies must be done prior to designing trials to minimise resource

wastage. Although live birth can be used in initial trial, arrangements should be made for follow-up to determine cumulative live birth rate. For a randomised trial comparing surgery versus no surgery prior to IVF in women with endometrioma, we recommend a Bayesian group-sequential design with live birth in an IVF cycle as the primary outcome with interim futility analysis using clinical pregnancy rate.

Acknowledgements

We thank all clinicians who filled the questionnaire, all patients and their partners who participated in interviews and focus groups, all women who helped to design WP3, theStaff at Aberdeen Fertility Centre who helped in recruiting for the qualitative study, Prof. Graeme McLennan (Director Clinical Trials Unit) and Prof. Shaun Treweek (Health Services Research Unit) in Aberdeen for their valuable suggestions in the design of the proposed trial, Dr Tarek Gelbaya for personal communication and Dr C Tomosseti for personal communication.

Authors' roles

A.M. conceived the idea; L.S., K.C., S.B.r., S.B.a., A.W.H., K.G., J.D., were co applicants. K.B. and G.B. were patient representatives on grant application and contributed intellectually throughout the conduct of the study. J.He. was a research fellow on grant and did the interviews, designed and administered the survey and did the day-to-day running of project. K.G. led the qualitative component. J.Hu did the statistical analysis. J.Hu. designed the example trial based on the findings. All authors contributed to initial and final drafts of the manuscript.

Funding

The study was funded by NHS Grampian pump priming endowment fund (RG14437-12).

Conflict of interest

S.B. is Editor-in-Chief of HROPEN, and A.W.H. was the Deputy Editor of HROPEN. Neither was involved in the review of this manuscript. L.S. reports grants from CSO and NIHR to do endometriosis research, outside the submitted work. K.C. reports grants from NIHR/HTA and CSO during the conduct of the study. J.He., A.W.H., J.D., S.Br., K.B., G.B., J.Hu. and K.G. report no conflict of interest.

References

Asakura K, Hamasaki T, Evans SR. Interim evaluation of efficacy or futility in group-sequential trials with multiple co-primary endpoints. *Biom J* 2017;**59**:703–731.

Berry DA. Bayesian clinical trials. *Nat Rev Drug Discov* 2006;**5**: 27–36.

Braakhekke M, Kamphuis EI, Dancet EA, Mol F, van der Veen F, Mol BW. Ongoing pregnancy qualifies best as the primary outcome measure of choice in trials in reproductive medicine: an opinion paper. *Fertil Steril* 2014;**101**:1203–1204.

Brooks J, McCluskey S, Turley E, King N. The utility of template analysis in qualitative psychology research. *Qual Res Psychol* 2015;**12**: 202–222.

Coomarasamy A, Devall AJ, Cheed V, Harb H, Middleton LJ, Gallos ID, Williams H, Eapen AK, Roberts T, Ogwulu CC *et al.* A randomized trial of progesterone in women with bleeding in early pregnancy. *N Engl J Med* 2019;**380**:1815–1824.

Demirol A, Guven S, Baykal C, Gurgan T. Effect of endometrioma cystectomy on IVF outcome: a prospective randomized study. *Reprod Biomed Online* 2006;**12**:639–643.

Dunselman GA, Vermeulen N, Becker C, Calhaz-Jorge C, D'Hooghe T, De Bie B, Heikinheimo O, Horne AW, Kiesel L, Nap A *et al.* ESHRE guideline: management of women with endometriosis. European Society of Human Reproduction and Embryology. *Hum Reprod* 2014;**29**:400–412.

Goldman B, LeBlanc M, Crowley J. Interim futility analysis with intermediate endpoints. *Clin Trials* 2008;**5**:14–22.

Gsponer T, Gerber F, Bornkamp B, Ohlssen D, Vandemeulebroecke M, Schmidli H. A practical guide to Bayesian group sequential designs. *Pharm Stat* 2014;**13**:71–80.

Hampson LV, Whitehead J, Eleftheriou D, Brogan P. Bayesian methods for the design and interpretation of clinical trials in very rare diseases. *Stat Med* 2014;**33**:4186–4201.

Human Fertilisation and Embryology Authority. *Fertility treatment 2017: trends and figures* 2019. https://www.hfea.gov.uk/media/2894/fertility-treatment-2017-trends-and-figures-may-2019.pdf (accessed on 12 November 2020).

Harb HM, Gallos ID, Chu J, Harb M, Coomarasamy A. The effect of endometriosis on in vitro fertilisation outcome: a systematic review and meta-analysis. *BJOG* 2013;**120**:1308–1320.

Hamdan M, Dunselman G, Li TC, Cheong Y. The impact of endometrioma on IVF/ICSI outcomes: a systematic review and meta-analysis. *Hum Reprod Update* 2015;**21**:809–825.

Horne AW, Saunders PTK, Abokhrais IM. Hogg L, on behalf of the Endometriosis Priority Setting Partnership Steering Group, top ten endometriosis research priorities in the UK and Ireland. *The Lancet* 2017;**389**:2191–2192.

Maheshwari A, McLernon D, Bhattacharya S. Cumulative live birth rate- time for a consensus? *Hum Reprod* 2015;**30**: 2703–2707.

Miller D, Pavitt S, Sharma V, Forbes G, Hooper R, Bhattacharya S, Kirkman-Brown J, Coomarasamy A, Lewis S, Cutting R *et al.* Physiological, hyaluronan-selected intracytoplasmic sperm injection for infertility treatment (HABSelect): a parallel, two-group, randomised trial. *Lancet* 2019;**393**:416–422.

Muzii L, Di Tucci C, Di Feliciantonio M, Marchetti C, Perniola G, Panici PB. The effect of surgery for endometrioma on ovarian reserve evaluated by antral follicle count: a systematic review and meta-analysis. *Hum Reprod* 2014;**29**:2190–2198.

Morgan B, Hejdenberg J, Hinrichs-Krapels S, Armstrong D. Do feasibility studies contribute to, or avoid, waste in research? *PLoS One* 2018;**13**:e0195951.

Raffi F, Metwally M, Amer S. The impact of excision of ovarian endometrioma on ovarian reserve: a systematic review and meta-analysis. *J Clin Endocrinol Metab* 2012a;**97**:3146–3154.

Raffi F, Shaw RW, Amer SA. National survey of the current management of endometriomas in women undergoing assisted reproductive treatment. *Hum Reprod* 2012b;**27**:2712–2719.

Redwine DB. Ovarian endometriosis: a marker for more extensive pelvic and intestinal disease. *Fertil Steril* 1999;**72**:310–315.

Scotland GS, McNamee P, Peddie VL, Bhattacharya S. Safety versus success in elective single embryo transfer: women's preferences for outcomes of in vitro fertilisation. *BJOG* 2007;**114**: 977–983.

Somigliana E, Berlanda N, Benaglia L, Viganò P, Vercellini P, Fedele L. Surgical excision of endometriomas and ovarian reserve: a systematic review on serum antimüllerian hormone level modifications. *Fertil Steril* 2012;**98**:1531–1538.

Tomassetti C, Adamson D, Arici A, Canis M, Hompes P, Hummelshoj L, Mol BW, Nardo L, Rombauts L, D'Hooghe T EndoART: a proposed randomized controlled trial on endometriomas in assisted reproductive technologies, comparing the effect of no intervention, surgery, and prolonged GnRH downregulation on pregnancy rates. *J Endometr Pelvic Pain Disord* 2018;**10**:158–173.

Young JM, O'brien C, Harrison JD, Solomon MJ. Clinical trials in head and neck oncology: an evaluation of clinicians' willingness to participate. *Head Neck* 2006;**28**:235–243.

Wong AYL, Warren S. Kawchuk, A new statistical trend in clinical research - Bayesian statistics. *Physical Therapy Reviews* 2010;**15**:372.