Endometrial scratch to increase live birth rates in women undergoing first-time in vitro fertilisation: RCT and systematic review.

METWALLY, Mostafa, CHATTERS, Robin, PYE, Clare, DIMAIRO, Munya, WHITE, David, WALTERS, Stephen, COHEN, Judith, YOUNG, Tracey, CHEONG, Ying, LAIRD, Susan <http://orcid.org/0000-0003-4020-9020>, MOHIYIDDEEN, Lamiya, CHATER, Tim, PEMBERTON, Kirsty, TURTLE, Chris, HALL, Jamie, TAYLOR, Liz, BRIAN, Kate, SIZER, Anya and HUNTER, Helen

Citation:

Endometrial scratch to increase live birth rates in women undergoing first-time in vitro fertilisation: RCT and systematic review

Mostafa Metwally, Robin Chatters, Clare Pye, Munya Dimairo, David White, Stephen Walters, Judith Cohen, Tracey Young, Ying Cheong, Susan Laird, Lamiya Mohiyiddeen, Tim Chater, Kirsty Pemberton, Chris Turtle, Jamie Hall, Liz Taylor, Kate Brian, Anya Sizer and Helen Hunter on behalf of the Endometrial Scratch Trial Research Team

Endometrial scratch to increase live birth rates in women undergoing first-time in vitro fertilisation: RCT and systematic review

Mostafa Metwally,[1]* Robin Chatters,[2] Clare Pye,[1]
Munya Dimairo,[2] David White,[2] Stephen Walters,[3]
Judith Cohen,[4] Tracey Young,[5] Ying Cheong,[6]
Susan Laird,[7] Lamiya Mohiyiddeen,[8] Tim Chater,[2]
Kirsty Pemberton,[2] Chris Turtle,[2] Jamie Hall,[2]
Liz Taylor,[1] Kate Brian,[9] Anya Sizer[9]
and Helen Hunter[10] on behalf of the Endometrial
Scratch Trial Research Team

[1]Assisted Conception Unit, Sheffield Teaching Hospitals NHS Foundation Trust,
   Sheffield, UK
[2]Sheffield Clinical Trials Research Unit (CTRU), School of Health and Related
   Research (ScHARR), The University of Sheffield, Sheffield, UK
[3]Design, Trials and Statistics, School of Health and Related Research (ScHARR),
   Sheffield, UK
[4]Hull Health Trials Unit, University of Hull, Hull, UK
[5]Health Economic and Decision Science, School of Health and Related Research
   (ScHARR), Sheffield, UK
[6]Faculty of Medicine, University of Southampton, Southampton, UK
[7]Faculty of Health and Wellbeing, Sheffield Hallam University, Sheffield, UK
[8]Saint Mary's Hospital, Manchester University NHS Foundation Trust, Manchester, UK
[9]Fertility Network UK, Greenwich, UK
[10]Department of Reproductive Medicine, Old St Mary's Hospital, Manchester, 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

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

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.

This report

Abstract

Endometrial scratch to increase live birth rates in women undergoing first-time in vitro fertilisation: RCT and systematic review

Mostafa Metwally,[1]* Robin Chatters,[2] Clare Pye,[1] Munya Dimairo,[2] David White,[2] Stephen Walters,[3] Judith Cohen,[4] Tracey Young,[5] Ying Cheong,[6] Susan Laird,[7] Lamiya Mohiyiddeen,[8] Tim Chater,[2] Kirsty Pemberton,[2] Chris Turtle,[2] Jamie Hall,[2] Liz Taylor,[1] Kate Brian,[9] Anya Sizer[9] and Helen Hunter[10] on behalf of the Endometrial Scratch Trial Research Team

[1]Assisted Conception Unit, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK
[2]Sheffield Clinical Trials Research Unit (CTRU), School of Health and Related Research (ScHARR), The University of Sheffield, Sheffield, UK
[3]Design, Trials and Statistics, School of Health and Related Research (ScHARR), Sheffield, UK
[4]Hull Health Trials Unit, University of Hull, Hull, UK
[5]Health Economic and Decision Science, School of Health and Related Research (ScHARR), Sheffield, UK
[6]Faculty of Medicine, University of Southampton, Southampton, UK
[7]Faculty of Health and Wellbeing, Sheffield Hallam University, Sheffield, UK
[8]Saint Mary's Hospital, Manchester University NHS Foundation Trust, Manchester, UK
[9]Fertility Network UK, Greenwich, UK
[10]Department of Reproductive Medicine, Old St Mary's Hospital, Manchester, UK

*Corresponding author Mmetwally@sheffield.ac.uk

Background: In vitro fertilisation is a widely used reproductive technique that can be undertaken with or without intracytoplasmic sperm injection. The endometrial scratch procedure is an in vitro fertilisation 'add-on' that is sometimes provided prior to the first in vitro fertilisation cycle, but there is a lack of evidence to support its use.

Objectives: (1) To assess the clinical effectiveness, safety and cost-effectiveness of endometrial scratch compared with treatment as usual in women undergoing their first in vitro fertilisation cycle (the 'Endometrial Scratch Trial') and (2) to undertake a systematic review to combine the results of the Endometrial Scratch Trial with those of previous trials in which endometrial scratch was provided prior to the first in vitro fertilisation cycle.

Design: A pragmatic, multicentre, superiority, open-label, parallel-group, individually randomised controlled trial. Participants were randomised (1 : 1) via a web-based system to receive endometrial scratch or treatment as usual using stratified block randomisation. The systematic review involved searching electronic databases (undertaken in January 2020) and clinicaltrials.gov (undertaken in September 2020) for relevant trials.

Setting: Sixteen UK fertility units.

ABSTRACT

Participants: Women aged 18–37 years, inclusive, undergoing their first in vitro fertilisation cycle. The exclusion criteria included severe endometriosis, body mass index ≥ 35 kg/m^2 and previous trauma to the endometrium.

Interventions: Endometrial scratch was undertaken in the mid-luteal phase of the menstrual cycle prior to in vitro fertilisation, and involved inserting a pipelle into the cavity of the uterus and rotating and withdrawing it three or four times. The endometrial scratch group then received usual in vitro fertilisation treatment. The treatment-as-usual group received usual in vitro fertilisation only.

Main outcome measures: The primary outcome was live birth after completion of 24 weeks' gestation within 10.5 months of egg collection. Secondary outcomes included implantation, pregnancy, ectopic pregnancy, miscarriage, pain and tolerability of the procedure, adverse events and treatment costs.

Results: One thousand and forty-eight (30.3%) women were randomised to treatment as usual ($n = 525$) or endometrial scratch ($n = 523$) and were followed up between July 2016 and October 2019 and included in the intention-to-treat analysis. In the endometrial scratch group, 453 (86.6%) women received the endometrial scratch procedure. A total of 494 (94.1%) women in the treatment-as-usual group and 497 (95.0%) women in the endometrial scratch group underwent in vitro fertilisation. The live birth rate was 37.1% (195/525) in the treatment-as-usual group and 38.6% (202/523) in the endometrial scratch group: an unadjusted absolute difference of 1.5% (95% confidence interval –4.4% to 7.4%; $p = 0.621$). There were no statistically significant differences in secondary outcomes. Safety events were comparable across groups. No neonatal deaths were recorded. The cost per successful live birth was £11.90 per woman (95% confidence interval –£134 to £127). The pooled results of this trial and of eight similar trials found no evidence of a significant effect of endometrial scratch in increasing live birth rate (odds ratio 1.03, 95% confidence interval 0.87 to 1.22).

Limitations: A sham endometrial scratch procedure was not undertaken, but it is unlikely that doing so would have influenced the results, as objective fertility outcomes were used. A total of 9.2% of women randomised to receive endometrial scratch did not undergo the procedure, which may have slightly diluted the treatment effect.

Conclusions: We found no evidence to support the theory that performing endometrial scratch in the mid-luteal phase in women undergoing their first in vitro fertilisation cycle significantly improves live birth rate, although the procedure was well tolerated and safe. We recommend that endometrial scratch is not undertaken in this population.

Trial registration: This trial is registered as ISRCTN23800982.

Funding: This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 26, No. 10. See the NIHR Journals Library website for further project information.

Contents

CONTENTS

x

List of tables

xvi

List of figures

List of abbreviations

A&E	accident and emergency	LBR	live birth rate
AD	absolute difference	LH	luteinising hormone
AE	adverse event	NEQAS	National External Quality Assessment Service
BMI	body mass index		
CENTRAL	Cochrane Central Register of Controlled Trials	OHSS	ovarian hyperstimulation syndrome
CI	confidence interval	ONS	Office for National Statistics
CINAHL	Cumulative Index to Nursing and Allied Health Literature	OR	odds ratio
		PI	Principal Investigator
CONSORT	Consolidated Standards of Reporting Trials	PIS	participant information sheet
		PPI	patient and public involvement
CRF	case report form	PRECIS-2	PRagmatic Explanatory Continuum Indicator Summary 2
CTRU	Clinical Trials Research Unit		
DMEC	Data Monitoring and Ethics Committee	PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
ES	endometrial scratch		
FET	frozen embryo transfer	PUL	pregnancy of unknown location
FSH	follicle-stimulating hormone	RCT	randomised controlled trial
GP	general practitioner	REC	Research Ethics Committee
HFEA	Human Fertilisation and Embryology Authority	RN	research nurse
		RR	relative risk
ICER	incremental cost-effectiveness ratio	SAE	serious adverse event
		SAP	statistical analysis plan
ICSI	intracytoplasmic sperm injection	SD	standard deviation
IPD	individual participant data	SET	single embryo transfer
IQR	interquartile range	SOP	standard operating procedure
IR	incidence rate	TAU	treatment as usual
IRR	incidence rate ratio	TIDieR	Template for Intervention Description and Replication
ITT	intention to treat		
IUI	intrauterine insemination	TMG	Trial Management Group
IVF	in vitro fertilisation	TSC	Trial Steering Committee

Plain English summary

The endometrial scratch is a simple procedure that involves 'scratching' the lining of the womb (the endometrium). Several small studies have shown that undertaking this before the first in vitro fertilisation cycle may improve live birth rates; however, other studies have contradicted this.

This large study was carried out to confirm whether or not having an endometrial scratch before the first in vitro fertilisation cycle would increase the number of women having a live birth compared with those having 'usual' in vitro fertilisation treatment (known as the 'control' group). We collected information about pregnancy, miscarriage, stillbirth, pain during the procedure and costs of treatment to find out if there were any meaningful differences.

A total of 1048 women aged between 18 and 37 years were randomly allocated to the two groups, so participants had a 50% chance of having the endometrial scratch. Women were followed up throughout their pregnancy to ascertain the outcome of their in vitro fertilisation cycle. Although the live birth rate was 1.5% higher in the endometrial scratch group (38.6%) than in the control group (37.1%), the difference was not large enough to show any benefit of having the procedure. Other outcomes did not differ significantly between the two groups. However, the procedure was safe and tolerable. We found that the cost of treatment was, on average, £316 per participant higher in the group that received endometrial scratch than in the control group; the difference was not large enough to show that receiving endometrial scratch was more cost-effective. We combined the results of this trial with those of previous trials that looked to answer a similar question, and found that, overall, the endometrial scratch procedure does not enhance the chances of achieving a live birth.

We conclude that endometrial scratch before first-time in vitro fertilisation does not improve the outcome of treatment, and we recommend that this procedure is not undertaken prior to a first cycle of in vitro fertilisation.

Scientific summary

Background

In vitro fertilisation (IVF), with or without intracytoplasmic sperm injection (ICSI), is a widely used reproductive technique in women unable to conceive naturally. An array of expensive 'add-on' treatments are often offered, many of which lack robust evidence to support their use. One such add-on is the endometrial scratch (ES) procedure, in which the endometrium is abraded prior to embryo transfer. ES is hypothesised to increase the chances of live birth by improving embryo implantation. However, in women undergoing their first IVF cycle, several controlled trials have identified contradictory evidence, with the most recent systematic review concluding that there is a lack of high-quality evidence. Many of these previous studies included both women undergoing their first IVF cycle and women undergoing subsequent cycles, and were therefore not powered to detect clinically worthwhile effects in those undergoing their first IVF cycle. A high-quality randomised controlled trial (RCT) is required to definitively determine if the ES procedure is effective and safe.

Objectives

The Endometrial Scratch Trial
This aimed to assess the clinical effectiveness, safety and cost-effectiveness of the ES procedure performed in the mid-luteal phase prior to the first IVF cycle, with or without ICSI.

Systematic review and meta-analysis
This aimed to synthesise evidence of the clinical effectiveness and safety of the ES, by combining the results of this trial with the results of similar RCTs in this population.

Qualitative substudy
The objective of the substudy was to understand the experiences of trial participants and fertility unit staff participating in the trial, including recruitment, receiving/delivering the ES, data collection methods and withdrawal from the ES procedure.

Methods

The Endometrial Scratch Trial

Design and setting
The trial was a pragmatic, two-arm, superiority, open-label, parallel-group, multicentre, individually randomised controlled trial undertaken at 16 fertility units in the UK, two of which were run privately.

Participants
Participants were women who:

- were aged 18–37 years (inclusive)
- were undergoing their first IVF cycle (with or without ICSI)
- were expected to be using fresh embryos and a single embryo transfer
- had a regular menstrual cycle, normal uterine cavity, good ovarian reserve and no relevant vaginal/uterine infections
- if randomised to receive ES, were willing to use a barrier method of contraception prior to the procedure.

Participants were excluded if they had received previous trauma to the endometrium, had a body mass index of ≥ 35 kg/m^2, were participating in another interventional fertility study or had grade 4 endometriosis.

Interventions
The ES procedure was performed by a doctor or nurse in the mid-luteal phase of the menstrual cycle preceding IVF. ES was performed by inserting a speculum into the vagina. A pipelle or similar device was then inserted into the cavity of the uterus and negative pressure was applied by withdrawal of the plunger. The sampler was rotated and withdrawn three or four times so that tissue appeared in the transparent tube.

Treatment-as-usual (TAU) participants received IVF treatment (with or without ICSI) in accordance with the usual care practice of their fertility unit. ES participants received the ES procedure followed by TAU.

Randomisation and blinding
Participants were randomised (1 : 1) via a web-based system to receive either TAU or ES using block randomisation stratified by fertility unit and planned treatment protocol. Only trial statisticians and a health economist were blinded.

Outcomes
The primary outcome was live birth after completion of 24 weeks' gestation within 10.5 months post egg collection. Secondary outcomes included implantation, pregnancy, miscarriage, ectopic pregnancy, multiple birth, preterm delivery, stillbirth, pain and tolerability of ES, safety, health resource use and treatment costs.

Sample size
The trial's recruitment target was 1044 participants (522 per group) to preserve a 90% power and 5% two-sided type 1 error, and to detect a 10% absolute difference (AD) in live birth rate (LBR), which was viewed to be of clinical importance to change practice. This assumed a 30% control LBR and 5% inflation to account for uncertainty around this LBR and dropouts.

Statistical analysis
Primary analysis was based on an intention-to-treat analysis population that included all randomised participants regardless of circumstances after randomisation. For the primary outcome, AD in LBRs between groups was the treatment effect of interest. Participants with missing data on live birth were assumed to have not achieved a live birth. Normal approximation to the binomial distribution was used to calculate the 95% confidence interval (CI) around the differences in LBRs and the associated p-value was calculated using Pearson's chi-squared test. Corresponding unadjusted odds ratio (uOR) and unadjusted relative risk (uRR) with 95% CIs were estimated using simple logistic regression and binomial generalised linear model with a log-link function, respectively. Sensitivity analysis was performed by adjusting for fixed stratification factors and potential prognostic factors.

Health economic evaluation
The cost-effectiveness analysis presented results as cost per extra live birth from the NHS and social care perspective in accordance with the National Institute for Health and Care Excellence guidelines. Women were asked to complete a questionnaire to collect their resource use data at baseline, 3 months post egg collection and 6 weeks post partum. A patient cost questionnaire sent at 3 months post egg collection requested information about the time taken to travel to appointments and loss of productivity. Unit costs were derived from appropriate national sources, which included NHS reference costs, Personal Social Service Research Unit costs and data from the Office for National Statistics. Results were presented in the net-benefit framework and allowed for uncertainty using bootstrapping and probabilistic sensitivity analysis.

Systematic review

Literature search

The electronic databases were searched in January 2020 and these included MEDLINE, EMBASE, Cumulative Index to Nursing and Allied Health Literature (CINAHL) and The Cochrane Library, including the Cochrane Database of Systematic Reviews (CDSR), the Cochrane Central Register of Controlled Trials (CENTRAL), Health Technology Assessment (HTA) and Database of Abstracts of Reviews of Effects (DARE) (archive only). ClinicalTrials.gov was searched in September 2020. The reference lists of included manuscripts and relevant systematic reviews were checked to identify any further potentially eligible studies.

Eligibility criteria

Eligible RCTs included women undergoing IVF for the first time that reported the effectiveness and/or safety of the ES procedure. No restrictions were imposed regarding the timing or method of the ES, the trial comparator, outcomes or date published. If there were insufficient details to allow extraction of study characteristics, we excluded reports published as abstracts only and reports published in languages other than English, including those where an abstract in English was provided.

Data extraction

One reviewer extracted data, which was checked by another reviewer. Two reviewers independently assessed within-study risks of bias using the Cochrane risk-of-bias tool (version 2.0).

Meta-analysis

This was undertaken in RevMan (v5.4; The Cochrane Collaboration, The Nordic Cochrane Centre, Copenhagen, Denmark) using random-effects models. Statistical heterogeneity was assessed using the I^2 statistic. Sensitivity analysis using a fixed-effects model was undertaken when there was little heterogeneity between trials.

Qualitative substudy

Participants sampling

Interviewees were purposively sampled from six of the trial sites, which were selected while ensuring variation in site type (NHS/private), size of centre and consent rate during the trial. Participants were selected ensuring variability in age, duration of infertility and live birth status. Participants who chose not to undergo ES were also interviewed.

Staff were sampled from the selected centres. Staff from other centres were also sampled because of the lack of staff availability at the selected centres.

Interviews and analysis

Potential interviewees were contacted by e-mail and consent was sought by telephone. Interviews were undertaken by telephone using a semistructured interview guide and were audio-recorded. Recordings were transcribed verbatim and analysed using an inductive framework approach. Staff and participant interviews were analysed separately. Two individuals coded four of the transcripts for the participant interviews and two for the staff interviews independently, and agreed on the final codes. The remaining transcripts were coded by one individual. All analysis was undertaken using NVivo (v12; QSR International, Warrington, UK).

Results

The Endometrial Scratch Trial

A total of 3454 potential participants were identified, of whom 1048 eligible women were randomised to either TAU ($n = 525$) or ES ($n = 523$), and followed up between July 2016 and October 2019. In the ES group,

86.6% (453/523) of women received the ES procedure as per protocol and 85.9% (449/523) subsequently received IVF, with only 9.2% (48/523) receiving IVF without ES. In the TAU group, 94.1% (494/525) of women received IVF, of whom only 1.0% (5/494) received the ES procedure outside the trial.

The median [interquartile range (IQR)] time from ES to embryo transfer was 34.0 (26.0–42.0) days ranging from 16.0 to 346.0 (outlier) days. A total of 99.8% (448/449) of women viewed the ES procedure as tolerable. The median (IQR) pain rating score within 30 minutes of the ES procedure and at 1 day and at 7 days after the ES procedure was 4.0 (2.0–6.0), 1.0 (0.0–3.0) and 0.0 (0.0–0.0), respectively.

The LBR was 37.1% (195/525) in the TAU group and 38.6% (202/523) in the ES group [unadjusted AD of 1.5% (95% CI –4.4% to 7.4%; $p = 0.621$), uRR of 1.04 (95% CI 0.89 to 1.21) and uOR of 1.06 (95% CI 0.83 to 1.37)]. Sensitivity analyses produced similar results even after adjusting for baseline covariates. The results were also generally consistent across prespecified subgroups. There were no statistical differences in the rates of secondary outcomes. Safety events were comparable across groups, although the proportions of babies of low or very low birthweight and small for gestational age were slightly lower in the ES group than in the TAU group. No neonatal deaths were reported. Only three severe congenital abnormalities were recorded, all in the TAU group.

The health economic analysis found that ES was more costly than TAU, with ES costing, on average, £316 more than TAU. However, the incremental cost-effectiveness ratio per successful live birth of £11.90 (95% CI –£134 to £127) was relatively low.

Systematic review and meta-analysis

Twelve trials were included. One trial contributed substantial heterogeneity to the meta-analysis of LBRs and pregnancy rates, and was therefore excluded, reducing heterogeneity to moderate or low levels. Meta-analysis showed no evidence of a significant effect of ES on LBRs [nine trials; odds ratio (OR) 1.03, 95% CI 0.87 to 1.22] and pregnancy rates (11 trials; OR 1.06, 95% CI 0.84 to 1.35) using a random-effects model, with consistent results when a fixed-effects model was used. Similar results were found for miscarriage rate using a random-effects model (10 trials; OR 0.96, 95% CI 0.57 to 1.63) and for multiple pregnancy rate (five trials; OR 1.09, 95% CI 0.68 to 1.74), with consistent results when a fixed-effects model was used. No significant effects on ectopic pregnancy rates were found (five trials; OR 0.66, 95% CI 0.17 to 2.51), but with high uncertainty. A meta-analysis was not undertaken for stillbirth or preterm delivery rates owing to the small number of studies reporting these outcomes.

Only six trials reported pain post procedure; few participants reported severe pain. Three studies reported a numerical measure of pain (from 0 to 10), with all three reporting moderate pain post procedure [with mean (standard deviation) scores of 4.1 (2.4), 4.2 (2.5) and 6.42 (2.35)]. Eight trials reported adverse events; however, seven trials recorded such events only in participants who received ES, limiting the conclusions that can be reached.

Qualitative substudy

Twenty-seven trial participants and seven staff members were interviewed. Both were generally happy with the recruitment process, and participants were prepared to receive the ES procedure. Three participants felt that recruitment was too informal. Some participants and staff stated that more information was required regarding the evidence base for or against ES. Eighteen of the interviewed participants discussed having positive preconceptions regarding the effect of the ES on the outcome of their IVF cycle, with staff, in some instances, appearing to have contributed to these. Some participants were unaware of the potential harms of participating in a RCT.

Participants' positive preconceptions meant that the recruitment process was challenging, as some participants felt demoralised when they were randomised to the TAU arm. Staff developed mechanisms to assist participants who felt this way. Five participants described the procedure as being more painful than they expected.

The participants who declined the ES procedure did so for personal reasons that could not have been prevented; however, the withdrawal of two participants could have been prevented by improved organisation at the fertility unit.

Conclusions

In this definitive trial, performing the ES procedure in the mid-luteal phase in women undergoing their first IVF cycle was found to be safe and well tolerated. However, the ES procedure did not result in a significant increase in the proportion of women achieving pregnancy and a live birth compared with TAU. Furthermore, it did not improve other secondary pregnancy outcomes, as their rates were comparable in the ES and TAU groups. The cost-effectiveness analysis found that ES may be cost-effective compared with TAU. However, this should be interpreted within the context of the clinical effectiveness analysis that showed that ES did not significantly improve LBR. Given the lack of consensus regarding cost-effectiveness thresholds for these outcomes, and high uncertainty in the cost-effectiveness estimates, we recommend that ES is not undertaken in this population.

A meta-analysis combining our trial results with the results of other trials undertaken in this population showed no significant clinical benefit of undertaking ES prior to the first IVF cycle to improve pregnancy outcomes.

A major strength of the study was the focus on one population in order to minimise heterogeneity. The trial recruited the target number of participants, with negligible contamination in the TAU arm. The response to treatment was good in both arms, with a high proportion of participants receiving a day 5 embryo transfer. The trial was open label and did not include a sham procedure in the TAU arm; however, this is unlikely to have biased the key trial results, as the key outcome measures are unlikely to have been influenced by a placebo effect. Nearly 10% of women randomised to receive ES did not receive it, which may have diluted the treatment effect; however, post hoc analysis excluding these participants produced consistent results and conclusions.

Despite identifying three safety outcomes (i.e. small for gestational age, low birthweight and very low birthweight) that suggested a potential positive effect on outcomes, we believe that future work in this area is unwarranted owing to the small number of events involved in these analyses and the likelihood that these findings are due to chance. An individual patient data meta-analysis is planned by Lensen and colleagues, the results of which should be received before carrying out any further research into the effectiveness of the ES procedure.

Trial registration

The trial is registered as ISRCTN23800982.

Funding

Chapter 1 Introduction

Background

Burden of fertility treatments

In vitro fertilisation (IVF) is an assisted reproductive technology widely used in women unable to reproduce naturally. IVF (in which the sperm and egg are mixed together in a Petri dish) can be undertaken with or without intracytoplasmic sperm injection (ICSI), in which the sperm is directly placed inside the egg; from hereon we refer to IVF, with or without ICSI, as 'IVF'. In 2018, about 54,000 patients in the UK underwent IVF.[1] Success rates are modest, but have been increasing over recent years, with an overall live birth rate (LBR) of 23% in the UK in 2018.[1] Undergoing IVF takes an emotional and financial toll on couples,[2] causing absence from work,[3] feelings of distress, lack of control and stigmatisation,[4] with repeated failure sometimes invoking subclinical emotional problems.[5] When treatment fails, there is often a desire to try one of a bewildering selection of expensive IVF 'add-ons' in order to improve the chances of success,[6] despite a lack of high-quality evidence to support their benefits or safety.[2,7,8] One such add-on is the use of local endometrial trauma, also known as endometrial scratch (ES), which can cost up to £325 per procedure and involves 'scratching' the lining of the womb (the endometrium) with a pipelle or similar device.[8]

Evidence for endometrial scratch

The use of ES to improve implantation rates in women undergoing assisted conception was first described in 2003.[9] Initially, ES was reported to positively affect the outcome of IVF in women with repeated implantation failure.[10] This was later also demonstrated in unselected populations (i.e. women who had undergone any number of previous IVF cycles),[11] including in a 2015 Cochrane review,[12] which reported a meta-analysis of randomised controlled trials (RCTs), finding that ES improves the chances of a live birth when undertaken in this unselected population [relative risk (RR) 1.37, 95% confidence interval (CI) 1.05 to 1.79]. However, a more recent systematic review has included subsequent RCTs and highlighted that the evidence for ES improving LBR in this population is unclear (RR of 0.95, 95% CI 0.64 to 1.41).[13] Subsequently, a large high-quality RCT by Lensen et al.[14] in an unselected population found that ES did not result in a higher LBR than 'usual' IVF. Despite this, the pros and cons of undertaking ES in an unselected population have been discussed since this publication, with proponents of the procedure arguing that it has low costs and a low level of harms, which, coupled with the possible benefits of taking a tissue sample at the same time as ES, warrant its use in selected populations.[11]

The chances of achieving a live birth decrease with successive cycles of IVF; therefore, the aforementioned evidence from unselected heterogeneous populations, in many cases including women who had undergone a previous IVF cycle, are not necessarily generalisable to the first-cycle population.[15] One systematic review by Vitagliano et al.[16] specifically focused on patients undergoing their first IVF cycle and included seven RCTs; it found that there was no evidence that ES followed by IVF, compared with IVF alone, increased the success of treatment, with a RR of live birth (or ongoing pregnancy if LBR was not reported) of 0.99 (95% CI 0.57 to 1.73; $p = 0.97$). Secondary outcomes (miscarriage, multiple pregnancy and ectopic pregnancy) were also not significantly altered by undertaking ES. Notably, the small sample sizes of the included studies resulted in uncertainty around the effects of ES in women undergoing their first IVF cycle, and the included trials were at either a high or an unclear risk of bias. Consequently, the authors concluded that a robust and definitive trial is required to assess the effect of ES on the chances of success of the first IVF cycle.[16]

Definitive studies of IVF 'add-ons' are challenging, as large numbers of patients are required to detect changes in LBRs with sufficient power, therefore making them expensive.[2] Consequently, studies in this

first-cycle population have been small and have included heterogeneous populations.[16] To date, four RCTs have specifically focused on assessing ES in women undergoing their first cycle of IVF,[17-20] all of which were included in the review undertaken by Vitagliano *et al.*[16] The two largest trials included 418 and 300 participants, and found significant increases in IVF success in women who received ES;[19,20] however, according to the Vitagliano *et al.*[16] review, in one study[20] the risk of bias was deemed to be high and in the other participants were not followed up until delivery.[19] Previous studies have also differed in the timing of the ES procedure, with Karimzade *et al.*[17] undertaking ES on the day of oocyte retrieval, and the majority of other studies undertaking ES in the menstrual cycle prior to the cycle in which embryo transfer was carried out. Notably, Karimzade *et al.*[17] identified that undertaking ES on the day of oocyte retrieval (compared with in the menstrual cycle prior to the cycle in which embryo transfer was carried out) resulted in a reduction in clinical pregnancy rates [12.3% vs. 32.9%, respectively; odds ratio (OR) 0.25, 95% CI 0.12 to 0.66; $p < 0.004$].

Three RCTs published since the Vitgaliano *et al.*[16] review included heterogeneous unselected populations but were not adequately powered to detect clinically important differences in women undergoing their first cycle.[14,21,22] In all three trials, outcomes for women undergoing their first cycle were not presented. However, for the unselected population, all three trials identified no significant differences in pregnancy rates or LBRs between women who received ES and those who received 'usual' IVF.[14,21,22] Of note, one of these trials stopped early following an unplanned futility analysis, as the clinical pregnancy rates between the ES and control (sham procedure followed by 'usual' IVF) groups were similar.[21]

The Human Fertilisation and Embryology Authority (HFEA) has set up a 'traffic-light' system to inform patients of the evidence supporting a selection of 'add-ons' that fertility units frequently offer and charge extra for – according to HFEA, the evidence to support the benefits of ES is currently unclear.[2]

Costs to the NHS and patients

Fertility treatment places a significant cost burden on the NHS and patients, with the cost to the NHS estimated to be around £68M per year (based on 2013 costs).[23] The number of first IVF cycles funded by the NHS has been decreasing over recent years; 58% of first cycles were funded by the NHS in 2017, reducing to 52% in 2018.[1] Given that the average patient undergoes three cycles of IVF,[1] and a single IVF cycle can cost £5000, the cost of IVF treatment is a significant burden to patients.[8] In addition, loss of productivity due to the emotional and physical effects of IVF treatment has been estimated at £540 (€596, converted in 2020) per woman per IVF cycle in one study.[3] The cost of IVF is increased further if the couple decides to undergo the ES procedure, despite a lack of evidence to suggest that it is effective or safe. ES is estimated to cost £325 per cycle, a cost that does not take into account the potential pain of receiving the procedure and the potential health-care contacts if the procedure is in fact unsafe. Therefore, any increase in LBRs in the first cycle of IVF will prevent patients from receiving subsequent IVF cycles, thereby reducing this burden.

The endometrial scratch intervention

Endometrial scratch is a routinely performed outpatient procedure in which a pipelle or similar device is inserted into the cavity of the uterus; negative pressure is applied by withdrawing the plunger and the sampler is rotated and withdrawn three or four times so that tissue appears in the transparent tube. The rotation and withdrawal of the plunger 'scratches' the endometrium. Risks were identified in a previous study when the procedure was undertaken on the day of oocyte retrieval.[17] The procedure is not known to be associated with any particular risks when undertaken in the menstrual cycle preceding that of IVF therapy, apart from pain during and shortly after the procedure.[12] As with any intrauterine procedure, there is a potential for intrauterine infection. However, women attending for fertility treatment are usually screened for serious vaginal infections, such as chlamydia, to minimise the risk of any spread of infection.

The exact mechanism by which ES may improve implantation is not yet known; however, it is known that implantation of the embryo is a complex process involving the release of several inflammatory mediators including uterine natural killer cells, leukaemia inhibitory factor and interleukin 15.[24] ES may lead to the release of inflammatory cells and mediators such as macrophages and dendritic cells, tumour necrosis factor alpha, interleukin 15, growth-regulated oncogene-α and macrophage inflammatory protein 1B.[25] ES has also been shown to cause the modulation of several endometrial genes that may be involved in membrane stability during the process of implantation, such as bladder transmembranal protein and adipose differentiation-related protein and mucin 1.[26]

Rationale for research, aims and objectives

In 2016, a survey undertaken by Lensen *et al.*[27] of clinicians based in the UK, Australia and New Zealand reported that ES was seen as potentially beneficial when undertaken prior to the first cycle of IVF by nine (6.3%) of the 143 surveyed clinicians, with a further 54 (38%) stating that they were neutral regarding its potential effect. Since this survey, a number of RCTs have reported conflicting evidence regarding the effect of ES on the outcome of IVF when undertaken prior to the first cycle of IVF, with both positive and negative results reported.[14,18,19,22,28] Given the conflicting evidence, and the lack of recent data regarding the provision of ES, it is unknown if this procedure is being widely delivered to women undergoing their first IVF cycle. However, anecdotal evidence collected by the Endometrial Scratch Trial team suggests that some fertility units in the UK are still offering this procedure to women undergoing their first cycle of IVF. Given the additional cost to patients of ES (around £325), this represents a significant burden.[8] Given the lack of evidence to support its use, ES may have a negative or unequivocal effect on pregnancy outcomes and may be unsafe or poorly tolerated in this population.

The Human Fertility and Embryology Authority states in its statistical report into multiple births that the risks associated with multiple births are the single biggest health risk associated with fertility treatment.[29] Multiple births carry risks to the health of both the mother and the babies. The birth of a healthy singleton child, born at full term, is therefore the safest outcome of fertility treatment for both mother and child, and this is best achieved through promoting the practice of single embryo transfer (SET).[30] Although SET may be associated with a slightly lower pregnancy rate in an IVF cycle using fresh (rather than frozen) embryos, when a patient has a surplus of embryos generated during the treatment cycle, and hence potentially more attempts at embryo replacement per cycle, the pregnancy rate is the same whether one or two embryos are replaced.[31] For this reason, most strategies of SET are limited to women for whom there is a reasonable chance of having surplus embryos available for cryopreservation. Through improving success rates, the ES procedure could have the potential to increase the uptake and implementation of the practice of SET, which would consequently have a large impact on the risks and costs associated with multiple pregnancies (which is the greatest risk of assisted conception treatment).[32]

The main trial aimed to assess the clinical effectiveness, safety and cost-effectiveness of the ES procedure in women aged between 18 and 37 years (inclusive) undergoing their first IVF cycle, with or without ICSI, using either antagonist or long protocols on the chances of achieving clinical pregnancy and live birth.

The objectives of the Endometrial Scratch Trial were to evaluate the:

- clinical effectiveness and safety of the ES procedure compared with 'usual' IVF treatment, with or without ICSI, in women undergoing their first IVF cycle by undertaking a RCT
- cost-effectiveness of the ES procedure. This was evaluated by collecting the costs to the NHS and patients of undergoing the intervention, including the cost of IVF care and other health-care contacts.

The secondary objectives were to:

- synthesise evidence of the clinical effectiveness and safety of the ES procedure by combining the results of this trial with other similar RCTs in this population
- understand the experiences of trial participants and fertility unit staff of participating in the trial, including recruitment, receiving/delivering ES, data collection methods and withdrawal from the ES procedure.

Chapter 2 Trial methods

Study design

This was a two-arm, pragmatic, individually randomised, superiority, open-label, parallel-group multicentre clinical trial that evaluated the clinical effectiveness, safety and cost-effectiveness of the ES procedure compared with no ES procedure, followed by usual IVF treatment, with or without ICSI. The pragmatic nature of the trial is explained using the PRagmatic Explanatory Continuum Indicator Summary 2 (PRECIS-2) tool (see *Appendix 1*).[33] Participants were randomised to one of these groups using a 1 : 1 allocation ratio stratified by fertility unit and planned IVF protocol (antagonist or long). Participants were recruited from 16 fertility units across the UK (England and Scotland) and followed up between 4 July 2016 and 24 October 2019.

Internal pilot summary

The main trial was designed with an internal pilot phase with stop–go criteria applied after 6 months of recruitment to assess the feasibility of recruitment and delivery of the ES procedure as per the protocol. The trial met the recruitment and intervention delivery outcomes (see *Appendix 2*), and therefore the trial continued as planned without any modifications to the trial protocol. Participants from the internal pilot phase were included in the final analysis. *Figure 1* is a flow chart of the study design.

Trial registration and ethics

The trial was registered on the ISRCTN registry (ISRCTN23800982) on 31 May 2016. The trial protocol was reviewed and ethics approval was granted by the National Research Ethics Service Berkshire South Central (reference 16/SC/0151).[34] Over the course of the trial, eight amendments were made to the protocol (consisting of seven substantial and one non-substantial amendment), all of which were also approved by the National Research Ethics Service Berkshire South Central (as detailed in *Appendix 3*).

Participant inclusion and exclusion criteria

Patients who met the following criteria were eligible for the trial. Parts of the text have been reproduced with permission from Pye *et al.*[35] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Inclusion criteria

- Women who were expected to be aged between 18 and 37 years (inclusive) at the time of egg collection.
- Women who were undergoing their first cycle of IVF, with or without ICSI treatment, using the antagonist or long protocol only.
- Women who were expected to receive treatment using fresh embryos.
- Women who were expected to be good responders to treatment, with:

  o ovulatory menstrual cycle (regular menstrual cycles defined by clinical judgement or with ovulatory levels of mid-luteal serum progesterone as defined by local laboratory protocols)
  o normal uterine cavity (assessed by transvaginal sonography at screening, and no endometrial abnormalities, such as suspected intrauterine adhesions, uterine septa, submucosal fibroids or intramural fibroids exceeding 4 cm in diameter, as assessed by the investigator, that would require treatment to facilitate pregnancy)
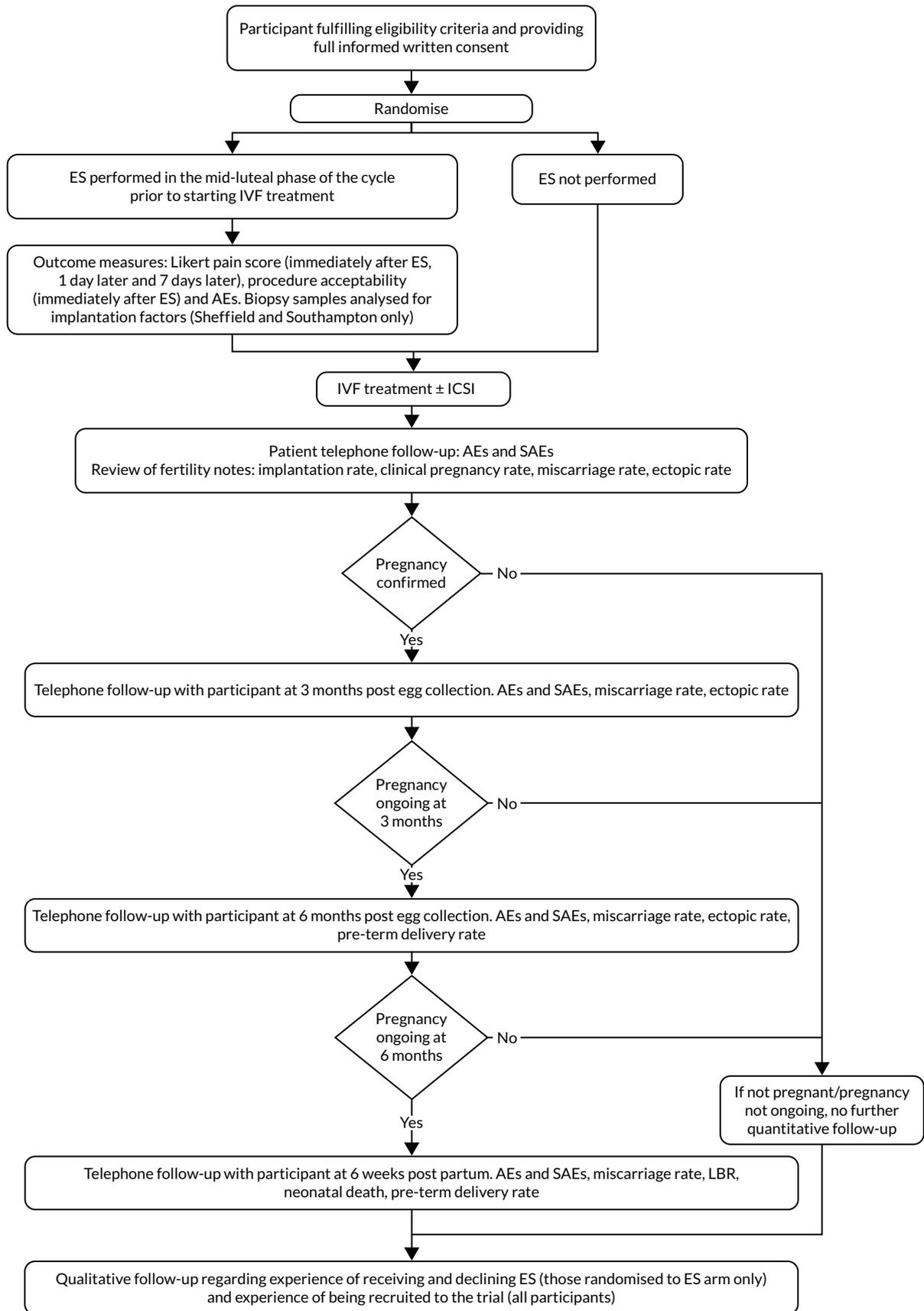
FIGURE 1 Flow chart of the study design. ±, with or without; AEs, adverse events; SAEs, serious adverse events.

    ○ expected good ovarian reserve, which was assessed clinically, biochemically [follicle-stimulating hormone (FSH) < 10 mIU/ml and normal follicular-phase oestradiol levels and/or normal anti-Müllerian hormone] and/or sonographically (antral follicle counts), and no history of previous radiotherapy or chemotherapy (all laboratory/ultrasound standards were based on local normal reference ranges)

    ○ SET expected.

- Women for whom local procedures were followed to exclude relevant vaginal/uterine infections prior to starting treatment.
- Women who were willing to use an appropriate method of barrier contraception if randomised to ES in the cycle in which the ES procedure is performed.
- Women who understood and were willing to comply with the protocol.

Exclusion criteria

The following women were excluded from the trial:

- women with a history of endometiral trauma or surgery (e.g. resection of submucous fibroid, intrauterine adhesions)
- women with a body mass index (BMI) $\geq$ 35 kg/m^2
- women with known grade 4 (severe) endometriosis
- women currently participating in any other fertility study involving medical/surgical intervention
- women who were expected to receive protocols other than antagonist or long (e.g. ultra-long protocol)
- women in whom an ES (or similar procedure, e.g. endometrial biopsy for the collection of natural killer cells) was planned (added to version 5 of the protocol, 20 July 2017; see *Appendix 3*)
- women previously randomised to this trial.

Participants for whom cycle programming was planned were initially excluded from the trial until 13 June 2016 (version 3 of the protocol), when this exclusion criterion was removed (see *Appendix 3*).

Recruitment of participants

Settings and locations where participants were recruited

During the internal pilot trial, recruitment took place at six IVF fertility units across England. In the main trial, recruitment sites were expanded to include an additional 10 IVF fertility units across England and Scotland. Two of these 16 fertility units were privately run outside the NHS. *Appendix 4* provides a list of sites and their involvement during and after the internal pilot phase.

Identification, screening and first approach

Participants, who were attending the IVF clinic prior to starting their first IVF cycle were identified by site staff [i.e. research nurses (RNs) and clinical trials assistants], who screened clinic lists to identify potential participants for the trial. Those who were identified and met the 'screening criteria' (aged 18–37 years inclusive, first IVF cycle, ovulatory menstrual cycle) were added to the screening log. A full 'eligibility' check was then undertaken, following consultation with the participants within the clinic, to confirm the remaining eligibility criteria.

Potential participants were approached about the trial by post by including an invitation to participate along with their initial clinic appointment letters, during routine clinic appointments before the start of their IVF or at patient information sessions, or they self-identified themselves after becoming aware of the trial from the trial website, recruitment video or posters that were placed within IVF units. Potential participants were approached at any time prior to the start of their IVF cycle, which was defined as the start of stimulation. Eligible participants were provided with written and verbal explanations about the trial; the participant information sheet (PIS) was provided to the individual either by post or in person,

depending on how the participant was first approached. Three different versions of the PIS were used owing to variation in trial procedures across the trial sites: one version for the Sheffield and Southampton sites (because of the tissue substudy at these sites; see Participant information sheet – Sheffield and Southampton sites), one for Dundee (where a postal consent process was used; see Participant information sheet – Dundee) and one for all other sites (see Participant information sheet – all sites, except Dundee, Sheffield and Southampton).

When participants had been approached and information about the trial provided, they were given time to consider their decision. Women were either asked about the trial at a subsequent routine visit to the fertility unit prior to the start of their IVF cycle, or contacted by telephone, e-mail or text message at a later date to ascertain their interest.

Informed consent

Informed consent was taken during a routine appointment prior to the start of the IVF cycle. The need to possibly delay the IVF cycle was discussed with participants, if randomisation was being undertaken close to the start of IVF, such that, if randomised to the ES arm, the IVF cycle would need to be delayed in order for ES to be delivered within the mid-luteal phase of the menstrual cycle prior to IVF. A decision to delay was made and agreed by both the participant and her fertility team prior to randomisation.

A written consent form (see Consent form – all sites except Sheffield and Southampton) was provided to the participant by a member of site staff [either the Principal Investigator (PI) or the RN], and, following the participant's agreement to participate, was signed by both the person taking consent and the participant. Participants recruited at the Sheffield and Southampton sites had the opportunity to consent to participate in a tissue substudy, in which tissue was collected from the endometrium and stored for future analysis (the results and detailed methods of which are not contained in this report); therefore, a different version of the consent form was used at these two sites (see Consent form – Sheffield and Southampton). Consent for this substudy was sought at the same time as consent for the trial. Participants were able to participate in the main trial without consenting to participate in the substudy, but not vice versa. Participants recruited at the Dundee site had the option to use a postal consent process, to lessen the burden of having to travel long distances to the fertility unit to consent and be randomised to the trial. In such cases, consent was obtained by post using an adapted consent form [see Postal consent form (Dundee only)], followed by consent in person post randomisation at the next routine appointment.

Description of trial arms

Endometrial scratch arm

Participants randomised to the ES arm received ES in the mid-luteal phase (defined as 5–7 days before the expected next period, or 7–9 days after a positive ovulation test) of the menstrual cycle prior to which IVF was being undertaken.

Endometrial scratch is a minor procedure performed in daily clinical practice in an outpatient setting. The procedure, which was already routinely undertaken at many of the fertility units for participants undergoing their second or subsequent IVF cycle, and, in some centres, was already offered to patients undergoing their first cycle of IVF, was undertaken as per local guidelines by a suitably trained doctor or nurse. When undertaking the procedure, a standard operating procedure (SOP) was followed (see Endometrial Scratch Standard Operating Procedure).

The participant was required to use a barrier method of contraception (if necessary) in the menstrual cycle in which the ES was performed. Before ES was undertaken, participants were asked if they had complied with this. If there was a risk of pregnancy (i.e. the participant reported unprotected sexual

intercourse prior to ES), the ES was not undertaken and was either rescheduled (and the IVF cycle delayed, as appropriate) or IVF commenced as planned without performing the ES procedure.

Participants were free to withdraw from receiving the ES at any point prior to or during the procedure. Those who did not receive the procedure remained in the trial and were followed up as described in *Procedures and assessments*.

At two centres (Sheffield and Southampton), endometrial biopsies were taken during the ES procedure (if the participant consented to do so). In order to collect the exact time point of menstruation at which the procedure was undertaken (and the tissue collected), ovulation kits were provided to participants who were randomised to the intervention arm at these centres. Once ovulation was detected [luteinising hormone (LH) surge detected], the participant contacted the fertility unit in order to schedule the ES, and the procedure was undertaken 7–9 days later. The tissue was harvested from the pipelle sampler that was inserted into the uterus to undertake the ES. Samples were collected and stored in formalin-containing specimen pots in the fertility units; the specimen pots will be stored for up to 10 years in liquid nitrogen. The tissue samples were anonymised and coded prior to storage and the liquid nitrogen Dewars were kept locked.

At all other fertility units, participants randomised to receive ES contacted the fertility clinic when their menstrual cycle preceding IVF began; the procedure was then scheduled 5–7 days prior to their next period. Women who failed to contact the site were telephoned by the RN.

To perform ES, a speculum was first inserted into the vagina and the cervix was exposed and cleaned. A pipelle or similar endometrial sampler (as per local procedures) was then inserted into the cavity of the uterus; negative pressure was applied by withdrawal of the plunger. The sampler was then rotated and withdrawn several times so that tissue appeared in the transparent tube. The sampler and speculum were then removed. If no tissue was seen in the transparent sampler, this was an indication that the sampler was not fully inside the uterine cavity and, therefore, the procedure was repeated. In total, the ES procedure took approximately 10–20 minutes. See *Appendix 5* for a description of the intervention using the Template for Intervention Description and Replication (TIDieR) checklist.[36]

Following ES, 'usual' IVF treatment was received as per local protocols and no further interventions were performed beyond usual care.

Treatment-as-usual arm
Participants randomised to the treatment-as-usual (TAU) arm received 'usual' IVF treatment in accordance with the local protocols at their fertility unit.

Randomisation and concealment

Randomisation was undertaken by a doctor or nurse based at the fertility unit prior to the start of the IVF cycle.

Eligible and consented participants were randomised (1 : 1) to either ES or TAU via a web-based randomisation system hosted by Sheffield Clinical Trials Research Unit (CTRU) using block randomisation stratified by fertility unit (site) and planned treatment protocol (antagonist or long). Research staff at recruiting sites logged onto the web-based randomisation system to enter participant details, and the treatment allocation was revealed only when randomisation was requested.

Random permuted blocks of size two, four, and six were used to ensure that participants were allocated evenly to each site and in accordance with IVF treatment protocols during the trial. The trial statistician generated the sequence that was retained within the web-based randomisation system,

which had restricted access rights, but they did not have access to the generated sequence. Research staff based at the trial sites did not have access to the randomisation sequence. The block sizes were masked to the research team during the trial except for the trial statistician who generated the randomisation sequence but was not involved in the screening and randomisation of participants. The details of block sizes that were disclosed at the reporting stage were documented in a restricted access computer folder. Once a participant was randomised, re-randomisation of the same participant was not permitted under any circumstance.

Blinding and masking

The nature of the ES procedure makes it impracticable to blind participating women, clinicians, research staff and outcome assessors, and therefore this was an open-label RCT. However, the trial uses objective pregnancy outcomes that are unlikely to be affected by the placebo effect or assessment bias; for example, achieving a live birth is a definitive primary end point. Therefore, it was unnecessary to use a sham ES procedure in the TAU group. The trial statisticians and health economist were blinded to the treatment allocation throughout the trial.

Outcome measures

Primary and secondary outcomes

Primary outcome

Live birth was defined as live birth after 24 weeks' gestation within the 10.5 months post egg collection follow-up period. The denominator for calculating the LBR is the number of women randomised to each treatment group. Multiple live births from a single pregnant woman because of a multiple pregnancy contributed to one live birth in the numerator (i.e. a multiple live birth was counted as one live birth).

Secondary outcomes

- Acceptability and tolerability of ES procedure:

    - pain rating and tolerability (yes or no) within 30 minutes of the procedure
    - pain rating directly after 24 hours and 7 days post procedure.

- Pain was assessed on a rating scale of 0 (no pain) to 10 (worst pain imaginable) and was self-reported by women who underwent the ES procedure.
- Implantation, measured based on a positive serum beta-human chorionic gonadotropin or by a positive urine pregnancy test on approximately day 14 following egg collection. The denominator for calculating the implantation rate is the number of women randomised to each treatment group.
- Clinical pregnancy, measured based on an observation of viable intrauterine pregnancy with a positive heart pulsation seen on ultrasound at/after 8 weeks' gestation.
- Ectopic pregnancy, measured as a pregnancy outside the normal uterine cavity.
- Miscarriage, defined as a spontaneous pregnancy loss, including pregnancy of unknown location (PUL) prior to 24 weeks' gestation within the 10.5 months post egg collection follow-up period.
- Multiple birth, defined as the birth of more than one living fetus after completed 24 weeks' gestation.
- Stillbirth, defined as the delivery of a stillborn fetus showing no signs of life after 24 weeks' gestation within the 10.5 months post egg collection follow-up period.
- Preterm delivery, measured by a live birth after 24 weeks before 37 weeks' gestation within the 10.5 months post egg collection follow-up period.
- Treatment cycle characteristics detailing the participant's IVF cycle, including number of eggs retrieved, number of embryos generated 1 day after egg collection, quality of the embryos transferred, number of embryos replaced and day of embryo replacement.

Safety outcomes

Data on adverse events (AEs) were collected up to 6 weeks post partum in participants (see *Table 1*). In the fetus/baby, data on severe congenital abnormalities (detected both antenatally and postnatally) and neonatal deaths were also collected up to 6 weeks post partum (see *Table 1*). Definitions of the different events recorded can be found below.

In participating women

- Adverse events, defined as any untoward medical occurrence in a participant that does not necessarily have a causal relationship with the intervention.
- Expected AEs, defined as an AE related to pregnancy or IVF treatments that is expected in this patient population (listed in *Appendix 6*).
- Unexpected AEs, defined as an AE that does not meet the definition of an expected AE as defined above.
- Serious adverse event (SAEs), defined as any untoward medical occurrence that results in death; is life-threatening; requires participant hospitalisation or prolongation of existing hospitalisation (excluding events related to the birth or birth process); results in significant disability/incapacity; is a congenital anomaly/birth defect; or is an important medical event.
- Expected SAEs, defined as a SAE that is expected in the patient population or is a result of the routine care/treatment of a patient (listed in *Appendix 7*).
- Unexpected SAEs, defined as a SAE that does not meet the definition of an expected SAE as defined above. This includes maternal death, stillbirth, and any other SAE that is not foreseeable.

In the fetus/neonate

- Severe congenital abnormalities, defined as a congenital anomaly detected antenatally or postnatally. Common minor congenital anomalies as defined by the European Monitoring of Congenital Anomalies (EUROCAT)[37] minor anomaly exclusion list were excluded as unexpected SAEs. Those anomalies that were excluded were either minor (e.g. skin tags) or expected for the gestation (e.g. patent ductus arteriosus in babies born at < 37 weeks' gestation).
- Neonatal deaths, defined as the death of a baby within 6 weeks of life.
- Preterm delivery, defined as delivery of a live birth of gestational age ≥ 24 weeks and < 37 weeks.
- Very preterm delivery, defined as delivery of a live birth of gestational age < 24 weeks.
- Low birthweight, defined as delivery of a live birth with birthweight ≤ 2499 g or < 10th centile for a given gestational age and sex of baby.
- Very low birthweight, defined as delivery of a live birth with birthweight < 1500 g or < 5th centile for a given gestational age and sex of baby.
- Small for gestational age, defined as delivery of a live birth with birthweight < 10th centile for a given gestational age and sex of baby.
- Large for gestational age, defined as delivery of a live birth with birthweight > 95th centile for a given gestational age and sex of baby.

Health economic outcomes

- Health resource use of the participant measured at baseline and 3 and 10.5 months (or at 6 weeks post partum) post egg collection.
- Patient costs at 3 months post egg collection.

Both outcomes were collected using a bespoke questionnaire (see *Collection of health resource use and patient costs data*).

Procedures and assessments

An online database was utilised in order to collect participant data. Data collection proceeded as described in *Table 1* (see *Figure 1* for a study flow chart).

TABLE 1 Data collection time points and sources

Time point	Data type	Data	Source
Baseline (pre randomisation)	Medical history	Alcohol, smoking, drugs	Medical notes
		Pregnancy and medical history	Participant (in person)/ medical notes
	Fertility history	Expected treatment protocol and methods	Medical notes
		Ovulation markers	Medical notes
		Fertility history	Medical notes
	Health resource use		Questionnaire
	Eligibility assessment		Participant (in person)/ medical notes
	Demographics		Medical notes
Randomisation		Date and outcome of randomisation	Randomisation system
Intervention	Details of ES procedure	Medications taken prior to procedure	Participant (in person)
		Day of cycle, reason ES not received (if applicable), instrument used	Medical notes
	Pain and tolerability: within 30 minutes of procedure	Likert scale 0–10 and tolerability (yes/no)	Participant (in person)
	Pain: 1 day and 7 days post procedure	Likert scale 0–10	Participant (SMS message)
Treatment cycle	IVF cycle details	Details of down-regulation, stimulation, egg collection, embryo transfer	Medical notes
Approximately 2 weeks post egg collection	Pregnancy test	Test type and results (blood/urine)	Medical notes
Approximately 2 weeks post positive pregnancy test	Pregnancy scan	Outcome of scan (fetal heartbeats detected, intrauterine sacs)	Medical notes
		Pregnancy outcomes[a]	Medical notes
Between 2 and 4 weeks post egg collection[b]	AEs[c]		Participant (telephone/ in person)
3 months post egg collection	Pregnancy outcomes[a]		Participant (telephone)
	AEs[c]		Participant (telephone)
	Health resource use		Questionnaire
	Patient costs		Questionnaire

TABLE 1 Data collection time points and sources (*continued*)

Time point	Data type	Data	Source
6 months post egg collection	Pregnancy outcomes[a]		Participant (telephone)
	AEs[c]		Participant (telephone)
6 weeks post partum	Pregnancy outcomes[a]		Participant (telephone)
	AEs[d]		Participant (telephone)
	Weight and gestational age of baby when born		Participant (telephone)/ medical notes
	Neonatal death		Participant (telephone)

SMS, short message service.
a Pregnancy outcomes include all related events: miscarriage, termination, stillbirth, live birth.
b If participant did not proceed to egg collection, AEs were collected at a similar time point.
c AEs collected in the participating women and congenital abnormalities detected antenatally in the fetus.
d AEs collected in the participating women; neonatal deaths and postnatally detected congenital abnormalities collected in the neonate.

Baseline data collection and randomisation

Following consent, participants underwent baseline assessments prior to randomisation, which included height/weight, ethnicity, medical history and a fertility assessment, including ovarian reserve assessment.

Randomisation was then undertaken as described in *Randomisation and concealment*. If randomised to do so, the participant received ES (see *Description of trial arms*, *Endometrial scratch arm*), followed by 'usual' IVF. If randomised to TAU, 'usual' IVF was delivered (see *Description of trial arms*, *Treatment-as-usual arm*).

In vitro fertilisation cycle

Following randomisation, and ES (if randomised to this arm), information regarding the received treatment cycle was collected, including information about down-regulation and stimulation phases, egg collection and embryo transfer. During embryo transfer, details of whether the clinician detected fluid on the endometrium or blood on the tip of the catheter were collected. In Newcastle, data regarding fluid on the tip of the catheter could not be collected because ultrasound-guided embryo transfer was not undertaken at this site.

Owing to a perceived lack of equipoise in this patient group, fertility unit staff asked participants randomised to either arm if an ES had been received 'outside' the trial.

Embryo grading

The quality of each transferred embryo was recorded. Three stages of embryo development are commonly recognised: morula (day 3 of development), cleavage (day 1, 2 or 3) and blastocyst (day 5 or 6). Embryo grading was undertaken immediately prior to embryo transfer; embryos were subjectively assessed by visual microscopic inspection as per the fertility unit's protocols and recorded using established grading schemes for cleavage-stage embryos and blastocysts. Sites used either the National External Quality Assessment Service (NEQAS)[38] or Gardner blastocyst grading systems,[39] in accordance with their normal clinical practice (see *Appendix 8* for a list of embryo grading systems used at each trial site). The NEQAS system was updated in April 2017 and, therefore, both systems were used during the trial.[40]

For each grading system, we mapped embryo grading to a common quality rating scale created by a principal embryologist, to enable the comparison of embryo quality between treatment groups and sites. A separate quality scale was created for cleavage and blastocyst embryos. Both conversion algorithms can be found in *Appendix 9*.

For embryo selection at the blastocyst stage, three elements of the blastocyst were graded: the expansion score, trophectoderm and inner cell mass. The conversion algorithm for blastocysts does not take into account expansion score, as blastocysts undergo dynamic expansion, the status of which can change depending on what time on day 5 of development the grading was performed. As fertility units graded blastocysts at different time points and we did not record whether the grades provided were at selection or transfer, or a mixture of the two, it was cleaner to leave out the 'expansion' score. Cleavage-stage embryos, which are not rated using the Gardner blastocyst grading system, were rated using a different scale in both old and new NEQAS grading. Morula-stage embryos are not commonly graded.

Intervention delivery and post-intervention pain and tolerability

For those randomised to the ES, the details of the ES procedure were collected when the participant attended the unit to undergo the procedure. Pain was subjectively assessed within 30 minutes post ES, and then at 1 day and 7 days post ES, as described in *Outcome measures, Secondary outcomes*. This was collected in person directly after the procedure, at which time the participant took part in the pain assessment and was asked if the procedure was tolerable (yes/no) (see Directly post ES pain and tolerability CRF). At the 1- and 7-day post-ES time points, the Likert rating scale was recorded via short message service (SMS) (see One day and seven day pain rating CRF). If the participant failed to respond to the message, a reminder text message was sent. If there was still no response, a member of the fertility unit telephoned or e-mailed the participant to request this information. Part-way through the trial, the wording of the SMS was reviewed and changed because of concerns that participants were misreading the question posed.

Pregnancy test and scan

Pregnancy tests and scans were undertaken as per local procedures and details of these were collected by the fertility units. AE information for the period prior to the pregnancy test was also collected by the fertility unit. In the event that the participant's IVF cycle had ceased prior to the pregnancy test (i.e. no eggs collected or no embryos generated), the participant was contacted at a similar time point (i.e. 2 weeks post pregnancy test) to collect information on AEs.

Post-treatment cycle follow-up

Participants were followed up for at least 10.5 months post randomisation, or 6 weeks post partum, unless they did not become pregnant following their first IVF cycle (in which case, participants were followed up until the termination of their IVF cycle, e.g. IVF cycle not started, no eggs fertilised or pregnancy not achieved) or their pregnancy ended before full term (e.g. miscarriage or termination). Participants who were randomised to the trial but whose IVF treatment was then delayed were followed up for as long as the study timelines allowed in order to collect information regarding their treatment cycle and outcome, if treatment occurred.

Participants who were randomised to the trial but became spontaneously pregnant prior to embryo transfer were followed up in the same manner. However, the starting point for determining the timing of 3- and 6-month follow-ups was the date of the participant's last menstrual period, rather than the date of egg collection (which did not go ahead in cases of spontaneous pregnancy).

Participants who underwent frozen embryo transfer (FET) and had received no previous fresh/frozen embryo transfers were followed up until a maximum of 6 weeks post partum, if the timelines of the trial allowed. Participants who commenced the stimulation phase of IVF but did not complete IVF for any reason were classed as having received their first cycle of IVF, and, therefore, their involvement in the trial was complete.

Collection of pregnancy outcome data

Following a positive pregnancy scan, patients were discharged from the fertility unit; therefore, in order to collect information regarding the outcome of the pregnancy, participants were contacted by the fertility unit by telephone at three time points post discharge: 3 months post egg collection, 6 months post egg collection and 6 weeks post partum (in the event of a live birth at any gestation). Pregnancy outcome information (where applicable) and AEs were recorded at each time point.

Collection of health resource use and patient costs data

The health resource use questionnaire (at baseline, 3 months post egg collection and 6 weeks post partum) and patient costs questionnaire (at 3 months post egg collection) were sent to the participant by post or electronically if the pregnancy was ongoing (at 3 months) or if a live birth occurred (at 6 weeks post partum). Questionnaires were not sent in the event of a negative outcome (i.e. miscarriage or stillbirth) to minimise distress, as the questionnaire pertained to events/health-care contacts that occurred during the pregnancy. If a participant did not complete her questionnaire within approximately 1–2 weeks, she was contacted by text message, letter or e-mail to ask her to return the questionnaire. Questionnaires collected information for the preceding 3 months. Resource use included the cost of IVF treatment, visits to the assisted conception unit and, for those who conceived, the cost of antenatal and postnatal visits, delivery costs and any hospital stays not related to birth for both mother and baby. The resource use questionnaire collected information on contacts with the midwife and general practitioner (GP) visits. The resource use questionnaire was designed for this study from data collection tools developed in the School of Health and Related Research (Sheffield, UK) and those collated by the Database for Instruments for Resource Use Measurement – slightly adapted versions were used at the three time points of data collection (see Health resource use questionnaire – baseline, 3 months and 6 weeks post partum). Patients were sent a cost questionnaire 3 months post egg collection and asked to record time taken to travel to appointments and loss of productivity (see Patient costs questionnaire).

Safety outcomes

Adverse events and SAEs were collected at four follow-up time points, and were collected during the post-treatment cycle follow-up phase only if the participant had a positive pregnancy scan. Site staff were delegated the responsibility for recording AEs and SAEs electronically on the case report form (CRF). Unexpected SAEs were logged on the CRF and also sent by e-mail to Sheffield CTRU to expedite reporting.

In the participating women

Safety events (AEs and SAEs) in the participating women were classed as expected or unexpected (see *Outcome measures, Safety outcomes*). A list of expected AEs and expected SAEs can be found in *Appendix 6* and *Appendix 7*, respectively. The research sites made the assessment regarding whether an event was expected, in addition to recording the relatedness and seriousness of the event. The way in which these events were collected differed; only the presence of expected AEs was collected (i.e. a box was ticked on the CRF to signify that the participant had experienced the AE), whereas a 'full' AE report was collected for all other AEs and SAEs. AEs were not formally assessed for relatedness to trial procedures, but SAEs were. The sponsor and CTRU were responsible for reporting to relevant regulatory bodies, where appropriate. In January 2019, the trial protocol was amended to reflect the fact that events related to the birth of a baby or the process of birth were no longer to be classed as AEs or recorded during the trial (see *Appendix 3*).

In the fetus/neonate

Safety outcomes in the fetus/neonate were entered directly onto the CRF and were not subject to categorisation of seriousness or expectedness. A list of events recorded can be found in *Safety outcomes*. For outcomes regarding birthweight and size, the site entered gestational age at birth, date of birth, weight and sex into the CRF.

Withdrawal from the trial

Participants could voluntarily withdraw their consent at any time. If a participant was not contactable at a follow-up point, and three attempts to contact her had been made, attempts were made to review AEs in the participant's fertility notes. Contact was attempted again at the next follow-up point. If the participant could not be contacted at the final follow-up (6 weeks post partum), attempts were made to obtain the outcome of the pregnancy and other related information (sex of baby, birthweight and gestation) from the participant's medical records.

Sample size

The primary outcome was live birth, defined as a live birth after completing 24 weeks' gestation within the 10.5-month post egg collection (or 6 weeks' post partum) follow-up. The denominator for calculating the LBR is the number of women randomised to each treatment group. Available data from HFEA suggest a LBR of 32.8% (women < 35 years) and 27.3% (women aged 35–37 years). We therefore assumed a 30% LBR in the TAU group and that a 10% absolute increase to 40% in the ES group is of clinical and practical importance, which is equivalent to a RR of 1.33. We proposed a large effect size of 10% absolute difference (AD) in LBRs, as it was believed that an effect of such magnitude is needed to change clinical practice. There is a 5% AD in the LBR between women aged under 35 years and women aged 35–37 years, and the proposed effect size is less than that observed in the systematic reviews, in which the RR estimates ranged from 1.83 to 2.2.

To preserve at least 90% power of detecting a 10% AD in LBRs between treatment groups for a 5% two-sided type I error, the trial required 992 women (496 per group) with continuity correction. After adjusting for an expected follow-up dropout rate of 5% (due to anticipated difficulties of follow-up for participants who have been referred from NHS trusts other than the participating fertility unit), the trial required 1044 women (522 per group). It should be noted that a decision was later taken during the statistical analysis plan (SAP) development to include all dropouts in the prespecified primary analysis [see *Chapter 3, Analyses of the primary outcome (live birth)*] while recruiting the original targeted sample size.[41] This compensated for residual uncertainty in the TAU group around the assumed 30% LBR.

Interim analyses and stopping rules

This was a fixed sample size trial designed without any interim analysis, so no formal stopping rules based on efficacy data were applicable. Final analysis was planned after the recruitment of the target sample size of 1044 participants (522 per group). However, the trial had an internal pilot to assess the feasibility of recruitment and delivery of the ES procedure with stop/go criteria as detailed in *Appendix 2*. In addition, as part of trial oversight, the Trial Steering Committee (TSC) and Data Monitoring and Ethics Committee (DMEC) reviewed ongoing data, as detailed *in Trial management and oversight* to safeguard the welfare of participants and integrity of the trial.

Patient and public involvement

Service user involvement has been integral to the success of the trial. The patient and public involvement (PPI) co-applicants, patient and public advisory groups and service users have contributed to the development and progression of the trial throughout its course.

The RN leading on public involvement was involved with the trial throughout, including at the outline bid, and sought early engagement from couples waiting to commence IVF treatment. In addition to this, the Reproductive Health Research Public Advisory Panel at the Jessop Wing, Sheffield Teaching Hospitals NHS Foundation Trust, and two service users named as co-applicants also provided input.

Pre-funding preparation

All groups received the project plan and lay summary, and were asked about their perspective on the proposed research plan and research intervention, as well as the physiological and psychological impact of this new intervention on infertility treatment. The Chief Investigator and lead RN attended all meetings and were able to answer questions and resolve queries.

Feedback from PPI influenced the intervention design, data collection protocol and recruitment strategy, but it specifically influenced the follow-up of women who had failed to achieve a pregnancy

or who experienced a miscarriage. PPI members felt that it was unreasonable and insensitive to complete questionnaires and be contacted by research staff to collect AE data if there was no ongoing pregnancy.

Post-funding preparatory work

The PPI lead arranged frequent communications with the PPI co-applicants, service users and advisory panel members (by e-mail, teleconference and face-to-face meetings) to develop study protocols and participant documentation. Both PPI co-applicants reviewed participant-related documentation (including, but not limited to, the poster, PIS/consent, questionnaires and patient introduction video). These documents and the patient introduction video were updated in the light of their comments.

In late 2015, the E-Freeze trial[42] (funded by the National Institute for Health Research) and the HABSelect trial[43] (which had required an extension to its recruitment timeline) were endeavouring to recruit the same patient population in many of the same research sites as was the Endometrial Scratch Trial. The Trial Management Groups (TMGs) of each trial requested input from the Reproductive Health Research Public Advisory Panel (Jessop Wing, Sheffield Teaching Hospitals NHS Foundation Trust) to help support a potential way forward to ensure the success of the three trials. The panel recommended the creation of an overarching leaflet for all potential participants to explain, in lay terms, the differences in the three trials in an unbiased manner. This was approved by the regulatory bodies and was used to facilitate recruitment. It was made clear that, at the time of writing this leaflet, no particular study was known to be superior to another in increasing the chances of a live birth.

Throughout progression of the internal pilot trial

Throughout the course of the internal pilot trial, changes to the study protocol and documentation were reviewed by the PPI co-applicants, service users and advisory panel members. Discussions among these groups also included recruitment progress and ideas to improve the process of collecting follow-up data. During the design of the qualitative substudy, PPI members were instrumental in supporting its design by, again, recommending changes to the recruitment techniques and the patient-facing materials, specifically the interview topic guide for participants.

Report writing, academic paper preparation and dissemination

All PPI co-applicants, service users and advisory panel members have been involved in the interpretation of the results, and they recommended that the trial team consider how certain aspects of the results were worded, in order not to provide undue positivity to the results of the trial. The PPI co-applicants, service users and advisory panel members also provided input to the presentation of results and dissemination materials, and the production of the final report to the funder (including preparation of the *Plain English summary*), and will continue to be involved in dissemination activities and the preparation of academic papers.

Compliance monitoring

Intervention

Compliance with the intervention was ascertained through the record taken by the clinician or nurse of whether or not the participant had attended the clinic for the ES procedure and received the procedure per protocol. Any deviations from the ES procedure were noted and are reported.

Other

Any deviation from the trial protocol or Good Clinical Practice regulations[44] was recorded by staff at the site or at the CTRU.

Trial monitoring

Responsibility for trial monitoring was delegated to the CTRU and conducted in accordance with CTRU SOPs, using both on-site and central monitoring approaches.

On-site monitoring was undertaken at all sites at study set-up, after around 10 participants were recruited to the trial and at the mid-point of the recruitment phase. At each visit, the site file and study logs were reviewed for completeness. Source data verification was undertaken for at least 10% of the randomised participants. Fertility notes were reviewed to substantiate participant existence and eligibility. Monitoring reports were issued after each visit detailing the actions required. Study close-out was undertaken remotely.

Central monitoring included point of entry validation, verification of data and post-entry validation checks.

Trial management and oversight

Trial Management Group
The TMG was responsible for the day-to-day management and oversight of the trial, and included the Chief Investigator, clinical co-applicants, a statistician, the trial manager, the lead RN and a PPI representative. The TMG reviewed blinded trial reports without any outcome data during trial management meetings.

Trial Steering Committee
The TSC provided overall oversight of the trial and comprised two independent gynaecologists and an independent statistician. Members of the TMG (including the trial manager and Chief Investigator) also attended meetings, which occurred at least once every 6 months. Blinded reports without any outcome data were provided by the data management team based at the CTRU prior to each meeting.

Data Monitoring and Ethics Committee
The DMEC was responsible for the safety of trial participants and the scientific integrity of the trial, and comprised two independent clinicians (gynaecologists) and one independent statistician. Even though the trial was not designed with formal stopping rules, blinded reports (for open-session discussions) and unblinded reports consisting of safety and outcome data (for closed-door discussions) were provided on a monthly basis and prior to each 6-monthly meeting.

Chapter 3 Statistical methods

This chapter details statistical methods and principles used to analyse outcomes relating only to the clinical effectiveness and safety objectives that are described in *Chapter 2, Outcome measures.* These methods are detailed in an accessible and approved pre-planned SAP version 2.0 and in the amendments in version 2.1.[41] This SAP also contains a history of amendments and whether or not they were made before or after blinded or/and unblinded review of the trial data. Any unplanned analyses conducted are highlighted in this chapter with the rationale for doing so. All analyses were performed in Stata® (v16.1) (StataCorp LP, College Station, TX, USA) software. For quality control, an independent statistician within the Sheffield CTRU reproduced the analyses of the primary and secondary outcomes (including sensitivity analyses) and safety outcomes in R (v3.6.0) (The R Foundation for Statistical Computing, Vienna, Austria) software.

Analysis populations

Intention to treat

The primary analysis was based on an intention-to-treat (ITT) analysis population that included all randomised participants with informed consent regardless of circumstances after randomisation, except if they withdrew consent and explicitly stated that their data should not be used. Treatment allocation was as randomised even if participants switched treatments or did not receive their randomly allocated treatment. The *Clinical effectiveness* section of this chapter explains how missing data were addressed under the worst- and best-case scenarios.

Per protocol

The purpose of the per-protocol analysis was to explore the effect of the ES procedure among those women who were viewed to have met key protocol requirements. The per-protocol population included women who complied with key protocol requirements, defined as those who:

- met the inclusion criteria as stipulated in the protocol
- received the allocated treatment, that is, excluding those who were allocated to TAU but underwent the ES procedure within or outside the trial, those who were allocated to ES but received only IVF, and those who did not receive any trial intervention
- did not achieve a spontaneous pregnancy before the delivery of the interventions
- completed the fertility treatment cycle and successfully generated embryos
- used contraception before the ES procedure (i.e. excluding those who failed to use contraception before ES and whose procedure could not be rescheduled)
- received treatment using fresh embryos (i.e. excluding those who underwent FET)
- were treated using only the antagonist or long protocols (i.e. excludes other protocols such as ultra-long).

Based on the advice of the TMG, since the delivery of the ES procedure is simple and unlikely to vary significantly across fertility units, per-protocol analysis also included women who were randomised to ES but known to have received the ES procedure outside the trial (if applicable).

Complete case

Complete-case analysis was used for sensitivity analysis on the primary analysis of the primary outcome. Complete-case population is the subset of the ITT population that included women with the known live birth outcome.

Safety

The analysis of all safety outcomes on participating women and born babies was based on the safety analysis population, which included randomised women who provided informed consent, and treatment assignment was according to the intervention received rather than the allocated intervention. Thus, women who were randomised to:

- ES but failed to receive the ES procedure before receiving IVF were reassigned to the TAU group
- TAU but received the ES procedure (within or outside the trial) before receiving IVF were reassigned to the ES group
- either intervention but failed to receive any were excluded.

Prespecified subgroups

The following subgroups were prespecified for subgroup analyses with an objective to explore the potential heterogeneity in the effect of the ES procedure across these subgroups:

- day of embryo transfer (day 2, 3, 4, 5 or 6)
- fertilisation method (IVF, ICSI or split ICSI)
- type of protocol (long treatment or antagonist)
- embryo transfer (single or double)
- nature of embryo used (frozen or fresh)
- history of miscarriages (0–2 or ≥ 3)
- cycle programming (yes/no).

In the ES group only, we also explored whether or not delaying the start of IVF after the ES procedure influences the chances of achieving a live birth.

Descriptive analyses for comparability between treatment groups

The demographics and characteristics of randomised women at baseline and treatment cycle characteristics during the trial were descriptively summarised by treatment group depending on the underlying distribution of variables without any statistical significance testing to assess comparability between the TAU and ES groups.[45-47] The numbers and percentages for categorical variables are reported. For continuous variables, the mean [standard deviation (SD)] and median [interquartile range (IQR)] as well as minimum and maximum values are reported. However, mean (SD) is reported in text for normally distributed variables, whereas the median (IQR) is used for skewed data. Any notable differences between treatment groups are discussed with implications, where applicable, in the results (see *Chapter 7*).

Clinical effectiveness

Analyses of the primary outcome (live birth rate)

The AD in LBRs was the primary measure of the treatment effect of interest. In the calculation of LBRs, we treated multiple live births per woman due to multiple pregnancies as a single event and the denominator was the number of women randomised. We estimated the 95% CI around the differences in LBRs using the normal approximation to the binomial distribution and calculated the associated *p*-value using Pearson's chi-squared test. A 5% two-sided significance level was used for the hypothesis test. To aid interpretation, in line with the Consolidated Standards of Reporting Trials (CONSORT) guidance,[48] we estimated the unadjusted OR with 95% CI using a simple logistic regression (binomial generalised linear model with a logit-link function) and the unadjusted RR with 95% CI using a simple

binomial generalised linear model with a log-link function. We performed a sensitivity analysis by adjusting for fixed stratification factors (i.e. site and planned treatment protocol) and potential prognostic factors [i.e. history of pregnancy (yes or no), age, BMI, duration of infertility and smoking status (yes or no)]. The adjusted OR and 95% CI were estimated using a multiple logistic regression model and adjusted RR (95% CI) using a binomial generalised linear model with a log-link function. The adjusted AD and associated 95% CI were post estimated via margins using the delta method after fitting a binomial generalised linear model with a log-link function.[49]

The primary approach to deal with missing data used the worst-case scenario by assuming that all women with an unknown live birth outcome, such as 'lost to follow-up', failed to achieve a live birth unless it was known otherwise (e.g. obtained from patients' medical notes). We performed a sensitivity analysis using the best-case scenario by assuming that women without pregnancy and/or live birth outcome data were pregnant and delivered a live birth. Furthermore, we also performed a sensitivity analysis using a complete case that included only women with the known live birth outcome (see *Complete case*).

Per-protocol analysis on the primary outcome was performed assuming a worst-case scenario using the same methods as described for the primary ITT analysis based on a per-protocol population, as defined in *Per protocol*. Following the presentation of trial results to the TMG, TSC and DMEC on 18 May 2020, an unplanned analysis was requested to aid the interpretation of the per-protocol results. This involved the estimation of the posterior probability of the treatment effect being at least 10% AD of clinical importance (equivalent to an OR of ≈ 1.5 given the observed 40.5% LBR in the TAU arm) given the observed data. We fitted an equivalent Bayesian multiple logistic regression model with all covariates that were included in the primary ITT analysis with normal priors on parameters with a mean of 0 and variance of 10^5. The posterior distribution of the treatment effect was estimated via Markov chain Monte Carlo sampling with a sample size of 200,000 and burn-in of 500.

Analyses of secondary outcomes

We performed unadjusted analyses on all secondary clinical outcomes in the same manner as the primary outcome (see *Analysis populations*). This covered ectopic pregnancy, implantation, clinical pregnancy, miscarriage, multiple birth, preterm delivery and stillbirth. We repeated this analysis on secondary outcomes that also relate to safety signals (except implantation and clinical pregnancy) in women who achieved successful implantation (i.e. with a positive pregnancy test). There was no correction for multiple hypotheses testing, as this was exploratory and all tests were performed at 5% two-sided significance level.

Subgroup analyses: primary and secondary outcomes

The purpose of the subgroup analyses was to explore whether the estimated treatment effects were consistent or there was potential heterogeneity across prespecified subgroups. The adjusted OR with associated 95% CI and interaction p-value were estimated using a multiple logistic regression model that included an interaction term between treatment and subgroup as well as fixed stratification factors (site and planned treatment protocol) and potential prognostic factors [history of pregnancy (yes/no), age, BMI, duration of infertility, smoking status (yes/no)]. However, for subgroups relating to the treatment protocol followed and the history of miscarriage, planned treatment protocol and history of pregnancy were not included as covariates, respectively (because the subgroup and covariate were highly correlated). The same approach to estimate the adjusted RR with associated 95% CI and interaction p-value using a Poisson generalised linear model with log-link function and robust standard errors was used. We did not estimate the adjusted AD with associated 95% CI because the post-estimation of margins using the delta method failed to converge.[49]

Subgroup analysis was only performed for the primary outcome (live birth) and selected secondary outcomes (clinical pregnancy and implantation). Other secondary outcomes with very small numbers of events were excluded because statistical models failed to converge; these were ectopic pregnancy, miscarriage, multiple birth, preterm delivery and stillbirth.

Safety analyses

Participating women

We used the Wilson score method to estimate the 95% CI around the proportion of women who tolerated the ES procedure based only on women with tolerability data.[50] Sensitivity analysis was performed assuming that women who received ES but without tolerability data did not tolerate the ES procedure. The distribution of pain rating scores within 30 minutes and at 1 day and at 7 days was displayed using a multiple box plot and was also summarised using means (SD), median (IQR) and minimum and maximum scores.

The expected AEs that were recorded during the trial were summarised by the treatment group using the numbers and proportions of women who experienced at least one event. Unexpected AEs and SAEs that occurred between receiving the intervention and the end of the trial were summarised by the number and proportion of women who reported at least one event and a total number of repeated events per treatment group. The AE category, frequency, seriousness, intensity, outcome and its relationship with ES procedure (in those who received ES) were summarised in the same manner.

The numbers of repeated events per woman were analysed using a negative binomial regression model (accounting for follow-up period) to estimate the incidence rate (IR) per treatment group and incidence rate ratio (IRR) with associated 95% CI. We also performed a sensitivity analysis by including all unexpected AEs and SAEs reported at any point during the trial. In addition, events that occurred between receiving ES and IVF were also summarised in the ES group only.

Born babies

In women with a positive pregnancy test, safety outcomes experienced by born babies (see *Chapter 2, Safety outcomes*) were summarised using the numbers and proportions per treatment group as well as the unadjusted AD in proportions between treatment groups with 95% CI estimated using the normal approximation to the binomial distribution.

The UK World Health Organization Neonatal and Infant Close Monitoring Growth Charts[51] were used to estimate the centiles of birthweight given gestational age and sex of baby using the *hbgd* R package [URL: https://hbgdki.github.io/hbgd/#growth-standards (accessed 4 February 2022)]. These centiles were used to classify babies at birth as low birthweight, very low birthweight, small for gestational age and large for gestational age, as described in *Chapter 2, Safety outcomes*.

Chapter 4 Economic evaluation methods

Another specific objective of the trial was to assess the cost-effectiveness of the ES procedure. This chapter details the methods utilised for the cost-effectiveness analysis, which aimed to compare ES with TAU. The results are presented in *Chapter 8* as incremental cost per extra live birth from an NHS and social care prospective. All analyses were undertaken using Stata (v16) software.

Analysis

The primary analysis aimed to present results as cost per extra live birth from an NHS and social care perspective in accordance with NICE guidelines.[52] Unit costs were derived from appropriate national sources and included NHS reference costs,[53] Personal Social Service Research Unit costs[54] and Office for National Statistics (ONS) data.[55]

Resource use
Resource use was collected as described in *Chapter 2, Collection of health resource use and patient costs data*. Unit costs were for 2018/19 and no discounting was applied as costs were presented over the time frame of the pregnancy, which was less than 1 year.

Primary cost-effectiveness outcome
We calculated the incremental cost-effectiveness ratio (ICER) as cost per percentage of successful pregnancies and the ICER as cost per successful pregnancies per 1000 population. To allow for uncertainty, a resampling method (known as bootstrapping) was used to obtain 95% bias-corrected CI around the ICER.[56] A total of 5000 simulations were used. Results are presented on the cost-effectiveness plane and on the cost-effectiveness acceptability curve.

Secondary cost-effectiveness outcomes
The cost-effectiveness of ES compared with TAU was examined for the following secondary outcomes: cost per clinical pregnancy, cost per ectopic pregnancy avoided, cost per miscarriage avoided, cost per multiple birth, cost per stillbirth avoided, cost per preterm delivery and cost per biochemical pregnancy. Analysis of variance was conducted to see whether or not there were any statistically significant differences in costs between the two treatment groups after allowing for the secondary outcome variable.

Prespecified subgroups
We explored the effectiveness differences in costs for different prespecified subgroups, as detailed in *Chapter 3, Prespecified subgroups*.

Sensitivity analysis
Resource use was collected at baseline, when participants were asked to recall their resource use over the previous 3 months. Analysis of covariance was undertaken to adjust for baseline costs and the ICER per successful live birth per 1000 population was calculated. A further analysis adjusting for baseline costs and any subgroups was also undertaken, as specified in *Prespecified subgroups*, and found differences in costs.

Triangulation based on labour and in vitro fertilisation resources reported by participants
The health resource use questionnaire was not designed to collect information on IVF or labour; this was collected directly from sites as part of the CRF. In the health resource use questionnaire, respondents were asked about visits to hospital, and some gave details of their IVF treatment and labour. Not all participants reported labour or IVF in the health resource use questionnaire; thus, if this questionnaire

was used alone, without the CRF data, the costs of labour and IVF would be under-reported. We assumed that those who did not report these resources in the questionnaire were also under-reporting other resources that they accessed during the study in the same way. As a sensitivity analysis, a simple triangulation exercise was carried out assuming that the information reported on the CRF was accurate, as it was recorded from patient notes. The percentage of participants who reported undergoing IVF and labour on the patient resource use questionnaire was reported and costs were inflated to reflect this under-reporting.

NHS and wider perspective

As part of the study, participants were asked to complete a patient cost questionnaire that asked them about how they travelled to their last fertility appointment, the cost and time taken to travel and what they would have been doing had they not been at the appointment (lost time/earnings). The questionnaire also asked whether or not a companion accompanied them and about any lost time/ earnings to their companion. Respondents indicated if they had travelled to their appointment by car, bus, train, taxi, foot or other mode of transport; they were also asked how far they travelled to their appointment. Respondents supplied details of the cost of car parking and any train, bus or taxi fares. The cost of car journeys was based on the UK Government's mileage allowance payments for car travel of £0.45 per mile[57] and multiplied by distance travelled to obtain a total cost for the journey.

Hourly rates of hours and earnings were taken from the ONS Earnings and Working Hours Survey for 2019.[58] It was assumed that any time away from usual activities had the same value for those in or not in paid employment, and the average hourly rate for the time spent at the appointment was applied to all participants. The cost of travel and time away from other activities was applied across all planned hospital visits plus seven additional visits for those who underwent IVF plus a further visit for those who became pregnant. These costs incurred by participants were added to the NHS costs to present a wider than NHS perspective in a sensitivity analysis.

Chapter 5 Qualitative methods

W e undertook a qualitative substudy, within a phenomenological framework, to provide context to the results of the trial and to improve the delivery and patient experience of future reproductive health studies. Qualitative interviews were undertaken with a sample of trial participants and site staff, which included doctors and nurses involved in recruitment and delivery of the trial, to explore their experience of recruitment, delivery of the intervention and delivery of the trial. The consolidated criteria for reporting qualitative research (COREQ) checklist was used when writing this report.[59]

Research questions and rationale for the qualitative interviews

This qualitative study aimed to answer the following research questions:

- What were the reasons for participants withdrawing from the intervention?
- What did those women who received the intervention think about the procedure and the information they received in preparation for the procedure?
- What were recruited trial participants' and research site staff's perceptions of the recruitment process and how did research sites deal with recruiting to two large randomised trials? How could the recruitment process be improved in future trials?
- What were recruited participants' perceptions of the data collection tools used during the trial, specifically text messages to collect pain and electronic questionnaires?

Reasons for non-receipt of the intervention were of particular interest. A total of 18 of 523 (3.4%) participants who were randomised to receive ES withdrew from the intervention (see *Figure 3*). During the trial, participants did not have to provide a reason for this, so, in all cases, the reason was not recorded on the trial database. Within this qualitative sub study, it was thought to be important to explore the reasons behind dropout from the ES, as this may identify ways in which ES, or the processes around introducing and scheduling the procedure, could be improved if the intervention were to be implemented in routine care.

Also of interest was the question of how research sites and participants dealt with recruitment being open at the same time to two large trials with overlapping eligibility criteria. The E-Freeze trial[42] was assessing the clinical and cost-effectiveness of a policy of freezing embryos followed by thawed FET, compared with a policy of fresh embryo transfer, in women undergoing IVF. The E-Freeze trial[42] and the Endometrial Scratch Trial were both open to recruitment at nine IVF centres: Sheffield, Southampton, Manchester, Birmingham, Liverpool, Homerton, Guy's and St Thomas' Hospital, Nottingham and Oxford. Both trials included women undergoing their first IVF cycle with women aged between 18 and 38 years inclusive, although the E-Freeze trial[42] allowed women aged up to 42 years to participate. The process that participants went through in order to decide which trial to participate in, and the site staff's involvement in this decision, was of interest in this qualitative substudy.

Interviewer characteristics

One woman [a research assistant with a PhD (Doctor of Philosophy)] and one man [RC, the trial manager with a BSc (Bachelor of Science)] conducted the interviews with all participants. Both had prior training and experience of undertaking qualitative interviews. Both researchers may have exhibited bias due to being involved in the set-up and design of the trial (RC) and the delivery and close-out of the trial (RC and the research assistant). RC was interested in the acceptability of the trial research procedures that he helped to create.

Relationship with participants

Staff

Both RC and AP worked on the trial; RC managed the trial from the start and AP became involved just prior to the start of this qualitative study. A relationship between the staff and RC was therefore already formed, which would also have been true for the research assistant and the staff, but to a lesser extent. Staff would have been aware of the interviewers' involvement in the trial (i.e. involvement in its design and delivery) and the reasons for the research.

Trial participants

The interviewers had no involvement in the clinical care of the trial participants, nor had they had previous contact with them. Participants would have been aware of the reasons for the research and may or may not have been aware of the interviewers' involvement in the trial in which they participated.

Trial participant sampling

We used purposive sampling in order to achieve the maximum variation in characteristics across trial participants and staff. Equal presentation was sought across these categories, except where otherwise stated.

Six sites were selected in order to recruit trial participants, using the following attributes:

- site type (NHS site/privately owned)
- involvement in E-Freeze trial[42] (also recruited to E-Freeze/did not recruit to E-Freeze)
- size of centre (number of IVF cycles undertaken per year)
- consent rate of centre (the proportion of patients consented relative to those screened).

The six sites selected were Sheffield, Leicester, Leeds, Guy's and St Thomas' Hospital, Wrightington and Oxford. Participants who had recently completed the trial (i.e. the site had logged the participant as having completed their involvement in the trial) and had not withdrawn from trial follow-up were then sampled from these sites. The following attributes were used:

- intervention status (ES received, TAU) – more participants who had received the ES procedure were sought, as a higher proportion of questions related to receiving the intervention
- age ($\leq$ 30 years, > 30 years)
- duration of infertility (< 36 months, $\geq$ 36 months)
- live birth status within trial follow-up period (live birth within 12 months of randomisation, no live birth within 12 months of randomisation).

In addition to the trial participants sampled as above, we invited to interview additional participants who were randomised to receive the ES procedure but had chosen not to undergo it. Those participants who withdrew from the trial as a whole (i.e. from the follow-up process), in addition to the intervention, were not approached.

Staff sampling

Individuals (PIs and RNs) from the six centres selected for the trial participant interviews were initially sampled for the staff interviews. Owing to low staff availability and/or willingness to participate in the interviews, additional sites were selected on the basis that the RN was still in post and had worked on the trial for a substantial period of the recruitment phase.

Ethics

This qualitative substudy was approved by Berkshire South Central Research Ethics Committee (REC) as a substantial amendment to the 'main' trial protocol (see *Appendix 3*).

Approach and recruitment

Trial participants

Trial participants were initially approached by letter or e-mail; a copy of the PIS (see Participant information sheet for the qualitative substudy – trial participant version) for this qualitative substudy was included. If no response was received, trial participants were telephoned approximately 1 week after the letter was sent, to ascertain their interest in participating. If interested, the interview was scheduled at a mutually convenient date and time. A letter or e-mail was sent to the individual to confirm the appointment time.

Staff

Initial interest to be interviewed was sought by e-mail or telephone – a PIS (see Participant information sheet for the qualitative substudy – staff version) was sent to the individual. If interested, a mutually convenient date and time for the interview was scheduled.

Consent

Consent was obtained by telephone immediately prior to the interview. In order to gain consent, the researcher read the statements from the consent form (see Qualitative substudy consent form) and asked the interviewee to state whether or not they agreed with them. Once all statements had been agreed to, the interviewee stated their name and the date was stated by the researcher. The interview did not commence if not all of the statements had been agreed by the interviewee. This conversation regarding consent was audio-recorded as a separate file to the interview itself. A completed copy of the consent form was posted to the interviewee for their records.

Data collection

A semistructured interview schedule was used to guide the interviews; one schedule was used for the trial participant interviews (see Semistructured interview guide – participants) and another was used for the staff interviews (see Semistructured interview guide – staff). The schedules were piloted in one interview per schedule, and adjustments were made following discussions between AP and RC. The pilot interviews were included in the analysis. Input to the pilot and adjusted schedule was sought from a PPI member (AS) for the participant interviews and from the lead RN (CP) for the staff interviews. The participant interviews were undertaken prior to the staff interviews in order to allow themes from the participant interviews to be discussed with staff; therefore, the staff interview schedule was constructed after the participant interviews had been completed. Prompts were used throughout the interviews to provide cues to participants who required them. Repeat interviews were not carried out.

Interviews were audio-recorded, with the interviewee's consent, and were transcribed verbatim. There were no other individuals present at the interviews except for, in a few instances, the participant's child. The interviewer recorded field notes after each interview. Interviews with trial participants lasted from 4.5 to 21.5 minutes [mean 14.1 (SD 6.4) minutes, median 13.5 (IQR 10.4–19.0) minutes]. Interviews with staff lasted from 19.5 to 57.6 minutes [mean 38.0 (SD 12.1) minutes, median 41.9 (IQR 25.1–46.5) minutes].

The trial participant interview schedule included questions regarding the participant's experience of recruitment, their preconceptions of the ES procedure and their experience of undergoing the ES procedure and undertaking the trial's data collection procedures. The staff interview schedule included questions regarding the staff member's experience of recruitment, their preconceptions of the ES procedure, scheduling and delivering the ES procedure and, in general, their experience of delivering the trial at their site.

All interviews were undertaken over the telephone. Data saturation was discussed separately for staff and trial participant interviews.

Data analysis

An inductive thematic analysis was used to analyse the data. We used Braun and Clark's reflexive thematic analysis framework,[60] which involves six phases of analysis: familiarisation, initial coding, theme construction, reviewing themes, defining themes and producing the report. Two pairs of researchers developed the coding structure separately by coding a small number of transcripts (four for the participant interviews and two for the staff interviews). NVivo (v12; QSR International, Warrington, UK) was used to undertake all analyses. RC and AP or RC and PK then met and resolved any differences in coding in order to develop the final list of codes. All other transcripts were then coded by one researcher, either AP or RC for the trial participant interviews, and either RC or PK for the staff interviews. Coding trees were formed of overarching themes that were related to the questions being asked: declining ES (trial participants only), experience of receiving/delivering ES, preconceptions of ES, recruitment, and data collection techniques. RC undertook all the analysis steps after coding, including the formation of themes and writing the analysis sections of the report. Interviewees did not provide feedback on the findings or transcripts.

Chapter 6 Systematic review methods

Rationale and objectives for the systematic review

Numerous systematic reviews have been undertaken to assess the clinical effectiveness of the ES procedure in the first IVF cycle. However, previous reviews have been unable to conclude the effect of ES on the first cycle of IVF, because of a lack of definitive, high-quality evidence.[10,12,16] Our definitive trial results (presented in *Chapter 7*) provide this evidence, and, therefore, we undertook a systematic review of the available evidence, aiming to combine the results of the Endometrial Scratch Trial with previously conducted RCTs in women undergoing the ES procedure prior to first-time IVF, with or without ICSI. Doing so should reduce the uncertainty around the effect of ES on key IVF outcomes, which has limited the conclusions of previous reviews.

Specific components of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist relevant to this report (i.e. introduction, methods, results and discussion) were used to construct this chapter and the corresponding results section.[61]

Protocol and registration

The systematic review was registered with PROSPERO (CRD42018111139; URL: www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID = 111139) on 18 October 2018.

Eligibility criteria

Types of participants
Studies involving women undergoing IVF, with or without ICSI, for the first time were considered for the review.

Studies that include participants who are undergoing intrauterine insemination (IUI) or ovulation induction (or other treatments not classed as IVF) and/or their second or subsequent IVF cycle will be excluded from this review, unless separate outcome data can be extracted for the subset of participants who have undergone their first IVF cycle.

Types of interventions
Studies evaluating the effectiveness or efficacy of the ES procedure, or a similar intervention (endometrial biopsy), were considered for the review.

No restrictions were placed on the type of comparator (e.g. sham procedure, other intervention or usual treatment).

Types of studies
Reports of randomised feasibility trials, pilot RCTs and RCTs were included in this review, as long as the outcomes of IVF of interest were reported (see *Outcome measures*). No date limits were set.

We excluded reports published as meeting abstracts only and/or reports published in languages other than English if the methodological details reported in the abstract were insufficient to allow extraction of study characteristics.

Subgroups

When analysing the results of the included studies, subgroup analyses by the following prespecified attributes were attempted, where data allowed:

- age
- duration of infertility
- IVF protocol (antagonist/long).

Where possible, we also subgrouped analyses by the timing of the ES procedure within the menstrual cycle, which was not prespecified.

Outcome measures

No restrictions were placed on the types of outcome measures considered. However, the outcomes of interest were live birth, pregnancy, miscarriage, ectopic pregnancy, multiple birth, stillbirth, preterm delivery, pain of the procedure and safety events in the mother and newborn.

Information sources

A comprehensive search was undertaken to systematically identify relevant RCTs. The search strategy comprised the following elements:

- searching key electronic databases for published journal articles
- searching electronic databases and web sites for grey literature
- scrutiny of bibliographies of relevant reviews and retrieved papers.

The search strategy was adapted across databases. Language and date restrictions were not being applied at this stage. Searches in the major databases were restricted by study type (i.e. RCTs and systematic reviews). An example MEDLINE search strategy is provided in *Appendix 10*.

ClinicalTrials.gov was searched using combinations of keywords: "endometrial biopsy" AND "infertility", "endometrial biopsy" AND "subfertility", "endometrial hysteroscopy" AND "infertility", "endometrial hysteroscopy" AND "subfertility".

The electronic databases were searched on 10 January 2020 and clinicaltrials.gov was searched on 21 September 2020.

The following electronic databases and websites were searched.

Electronic database sources

- MEDLINE(R) In-Process & Other Non-Indexed Citations and MEDLINE(R) (Ovid) 1948 to present
- EMBASE (Ovid) 1980 to present
- Cumulative Index to Nursing and Allied Health Literature (CINAHL) 1981 to present
- The Cochrane Library, including the Cochrane Database of Systematic Reviews (CDSR) 2005 to present, Cochrane Central Register of Controlled Trials (CENTRAL) 1898 to present, Health Technology Assessment (HTA) 1989–October 2016 (currently not being updated) and Database of Abstracts of Reviews of Effects (DARE) 1994–April 2015 (archive only)

Grey literature and internet sources

- ClinicalTrials.gov (www.clinicaltrials.gov/).

Study selection

Titles and abstracts of citations identified by the searches were screened for potentially relevant studies by two reviewers (JH and LR), with differences resolved by discussion with RC. Full texts were screened and checked by the same two reviewers (JH and LR), with differences also resolved by discussion with RC.

Data collection

One reviewer performed data extraction of each included study. For all included trials, extracted data were checked against the original article by a second reviewer. Any disagreements were resolved through discussion. Where studies consisted of duplicate reports (parallel publications), the most recent and relevant report was used as the main source, and additional reports checked for extra information.

Attempts were made to contact authors for missing data (see *Chapter 10*). Owing to the constraints on time, it was not possible to contact all the study authors for each of the items of missing data, for example risk-of-bias assessment.

Data items for extraction
The following data items were extracted from the included studies:

- study characteristics, such as country of origin, recruitment period, number of treatment arms, number randomised, and ITT and per-protocol samples
- participant characteristics, such as average age, average duration of infertility, egg source (i.e. donor eggs) and IVF protocol used
- intervention and control group details, such as timing of ES, ES technique, device used, co-interventions, and control group condition
- outcome data: live birth, ongoing pregnancy, clinical pregnancy, biochemical pregnancy, implantation, miscarriage, ectopic pregnancy, multiple pregnancy, multiple birth, pain of ES, stillbirth, preterm delivery and AEs.

Quality assessment strategy

The methodological quality of the included RCTs was assessed using the Cochrane Collaboration risk-of-bias assessment criteria at an outcome level:[62] process (domain 1), deviations from intended interventions (domain 2), missing outcome data (domain 3), measurement of the outcome (domain 4), selection of the reported result (domain 5) and overall bias (domain 6). Each domain is considered 'low risk', 'some concern' or 'high risk' based on answers to a series of signalling questions. The risk of bias was assessed for each reported outcome. The assessment was undertaken independently by two reviewers (i.e. PK, RC, AP or JH). Discrepancies were resolved by a third reviewer who had not been involved in the previous assessments of that study (RC or JH). Studies were graded with an overall risk of bias of 'high', 'low' or 'unclear'.

Synthesis of results

Analyses were undertaken using RevMan v5.2 (The Cochrane Collaboration, The Nordic Cochrane Centre, Copenhagen, Denmark). A random-effects model was primarily used, with a fixed-effects model used for sensitivity analysis when evidence of heterogeneity was low. We used ORs as a measure of between-group treatment effects of interest. The Mantel–Haenszel method was used to pool the effect estimates within meta-analyses.

Safety and pain data were synthesised narratively, owing to heterogeneity in the way in which these outcomes were reported.

One four-armed trial was included in the review, involving two different time points of ES, and two different time points of a sham procedure.[18] As suggested by Cochrane,[63] a technique of dealing with this is to combine the two time points, so that the proportion of participants fulfilling each outcome was summed across the two intervention and control time points, resulting in one proportion for the intervention overall, and another for the control overall. This was deemed appropriate as other trials included in the review also combined multiple time points of ES delivery within the same trial arm (e.g. Lensen *et al.*,[14] where ES was undertaken from day 3 of the menstrual cycle prior to IVF, to day 3 of the IVF cycle).

The I^2 statistic was used to measure statistical heterogeneity, which was interpreted in line with the *Cochrane Handbook for Systematic Reviews of Interventions*:[63]

- I^2 statistic of 0–40% = heterogeneity may not be important
- I^2 statistic of 30–60% = may mean moderate heterogeneity
- I^2 statistic of 50–90% = may mean substantial heterogeneity
- I^2 statistic of 75–100% = considerable heterogeneity.

Chapter 7　Clinical effectiveness results

This chapter reports the results to address the objective concerning clinical effectiveness and safety of the ES procedure stated in *Chapter 1, Rationale for research, aims and objectives*. The statistical methods used are summarised in *Chapter 3* and detailed in an accessible SAP (v2.1).[41] The reporting follows the CONSORT 2010 guidelines for individually randomised parallel-group trials,[64] extensions for pragmatic trials[65] and for reporting harms.[66]

Screening and flow of trial participants

Appendix 11 displays trial recruitment, which was completed 3 months ahead of schedule. Eligibility screening of participants and their follow-up from randomisation to the end of the study is presented in *Figure 2*. In summary, 3454 participants were identified, of whom 3204 (92.8%) were contacted; the study was discussed with 2486 (72.0%). For various reasons (stated in *Figure 2*), mainly lack of interest and ineligibility, 1392 (7.2%) women were not interested in taking part. Of those screened, 1072 (31.0%) were eligible and willing to take part; however, seven of these were not randomised because they withdrew consent ($n = 1$), were lost to follow-up ($n = 2$) or were no longer eligible ($n = 4$) before randomisation. Thus, 1048 (30.3%) women were randomised with informed consent to receive TAU ($n = 525$) or ES ($n = 523$). Of note, there were 1.9% (10/525) and 1.1% (6/523) women in the TAU and ES group, respectively, who were spontaneously pregnant after randomisation but prior to receiving any interventions.

Characteristics of study participants

We describe the characteristics of randomised participants to assess whether or not they are representative of the expected study population and whether or not they are similar between treatment groups to aid the interpretation of results.

Table 2 summarises the demographics and characteristics of randomised participants by treatment group and also the totals across treatment groups. In general, randomised participants were very similar, on average, across treatment groups. Women were recruited from 16 fertility units across the UK with four major fertility units (Sheffield, Leeds, Liverpool and Manchester) contributing 510 (48.7%) participants. The majority of women [932 (88.9%)] were white. The overall mean (SD) age and BMI was 32.5 (3.4) years and 24.5 (3.3) kg/m^2, respectively.

Only 24 (2.3%) participants were current cigarette smokers. No participants self-reported current use of recreational drugs; however, this could be associated with reporting bias. The majority [743 (70.9%)] of participants had no history of a previous pregnancy; 212 (20.2%) participants had had only one previous pregnancy, and only 50 (4.8%) had had a previous live birth. Previously, only 49 (4.7%) participants gave birth to a fetus with a gestational age of at least 24 weeks (either a live birth or stillbirth) and 169 (16.1%) participants had had at least one miscarriage. The median (IQR) duration of infertility was 2.8 (2.0–3.5) years.

Protocol non-compliances relating to participant ineligibilities

Only 10 out of 1048 (0.9%) women randomised (four in the TAU group and six in the ES group) were found to be ineligible at some point during the trial, for reasons that are summarised in *Appendix 12*. In five of these cases, there were randomisation errors at sites: randomisation of women with a BMI of 35 kg/m^2 or an expected age at egg collection of > 38 years or who had suffered previous endometrial trauma.
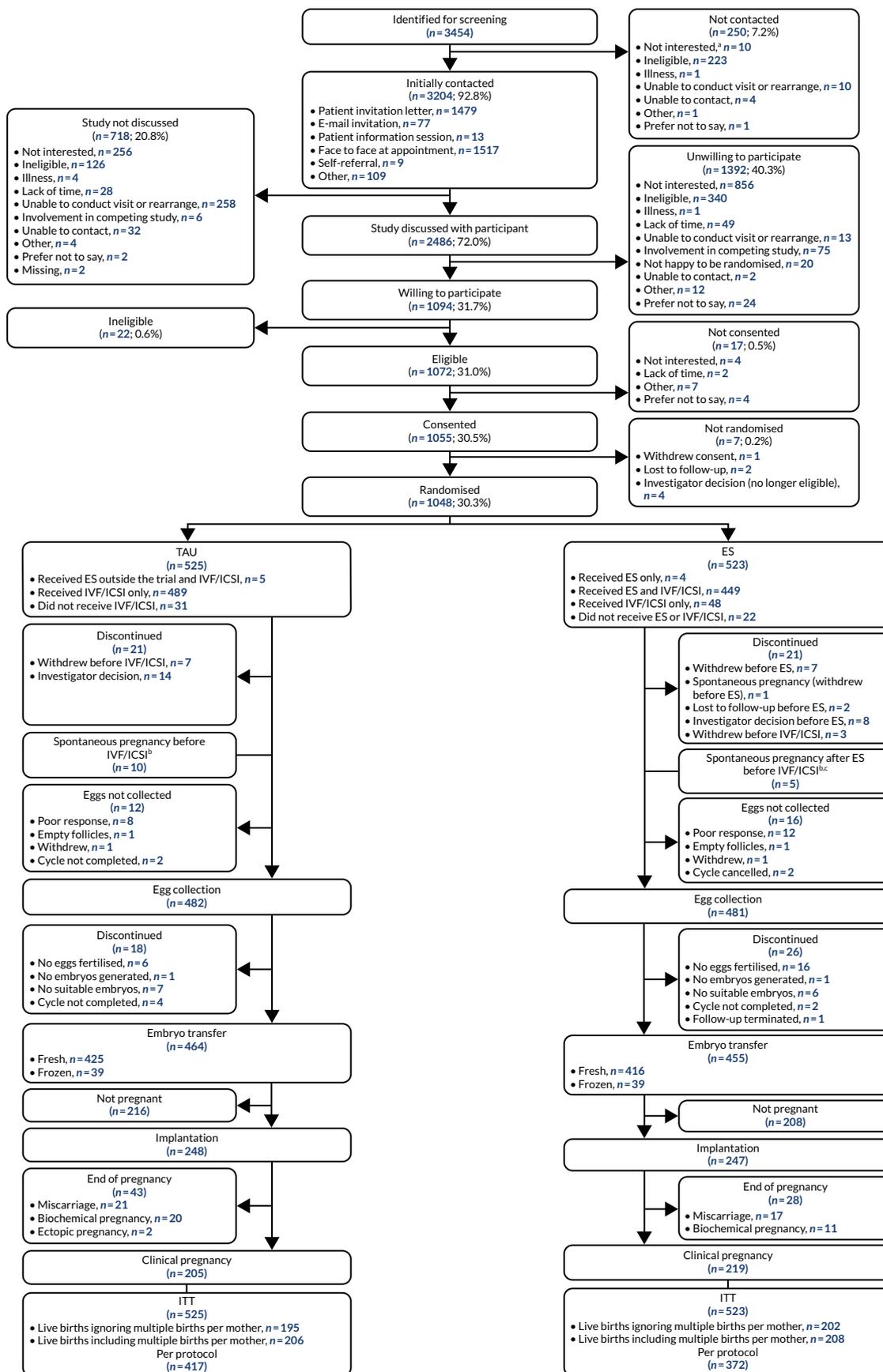
FIGURE 2 The CONSORT flow chart: screening and follow-up of study participants. a, Participant had already stipulated prior to being approached that they did not want to take part in any research while undergoing their IVF cycle. b, Note that spontaneous pregnancies were followed up. c, One spontaneous pregnancy resulted in ectopic pregnancy.

TABLE 2 Baseline demographics and characteristics of trial participants (ITT population)

Demographic or characteristic	TAU group (N = 525)	ES group (N = 523)	Total (N = 1048)
Fertility unit, n (%)			
Birmingham Women's and Children's NHS Foundation Trust	32 (6.1)	31 (5.9)	63 (6.0)
NHS Tayside (Dundee)	25 (4.8)	26 (5.0)	51 (4.9)
Gateshead Health NHS Foundation Trust	31 (5.9)	30 (5.7)	61 (5.8)
Guy's and St Thomas' NHS Foundation Trust	28 (5.3)	27 (5.2)	55 (5.2)
Homerton University Hospital NHS Foundation Trust	14 (2.7)	15 (2.9)	29 (2.8)
Leeds Teaching Hospitals NHS Trust	61 (11.6)	61 (11.7)	122 (11.6)
University Hospitals of Leicester NHS Trust	25 (4.8)	26 (5.0)	51 (4.9)
Liverpool Women's NHS Foundation Trust	54 (10.3)	56 (10.7)	110 (10.5)
Manchester University NHS Foundation Trust	62 (11.8)	63 (12.0)	125 (11.9)
Newcastle Upon Tyne Hospitals NHS Foundation Trust	28 (5.3)	28 (5.4)	56 (5.3)
Nurture Fertility (Nottingham)	3 (0.6)	3 (0.6)	6 (0.6)
Oxford Fertility	27 (5.1)	26 (5.0)	53 (5.1)
Sheffield Teaching Hospitals NHS Foundation Trust	78 (14.9)	75 (14.3)	153 (14.6)
South Tees Hospitals NHS Foundation Trust	15 (2.9)	15 (2.9)	30 (2.9)
University Hospital Southampton NHS Foundation Trust	28 (5.3)	29 (5.5)	57 (5.4)
Wrightington, Wigan and Leigh NHS Foundation Trust	14 (2.7)	12 (2.3)	26 (2.5)
BMI (kg/m^2)			
	(n = 525)	(n = 523)	(n = 1048)
Mean (SD)	24.5 (3.4)	24.5 (3.3)	24.5 (3.3)
Median (IQR)	24.1 (22.0–27.1)	24.2 (21.9–27.1)	24.1 (22.0–27.1)
Minimum, maximum	17.3, 35.0	16.8, 34.9	16.8, 35.0
Expected age at egg collection (years)			
	(n = 525)	(n = 523)	(n = 1048)
Mean (SD)	32.4 (3.4)	32.6 (3.4)	32.5 (3.4)
Median (IQR)	32.6 (30.2–35.0)	33.0 (30.4–35.3)	32.8 (30.3–35.1)
Minimum, maximum	21.5, 38.0	21.4, 38.1	21.4, 38.1
Actual age at egg collection (years)[a]			
	(n = 482)	(n = 481)	(n = 963)
Mean (SD)	32.4 (3.4)	32.7 (3.3)	32.5 (3.4)
Median (IQR)	32.5 (30.1–35.0)	33.0 (30.4–35.2)	32.8 (30.3–35.1)
Minimum, maximum	21.4, 38.8	21.4, 38.1	21.4, 38.8
Ethnicity, n (%)			
White[b]	472 (89.9)	460 (88.0)	932 (88.9)
Asian/Asian British[c]	31 (5.9)	47 (9.0)	78 (7.4)
Mixed/multiple ethnic groups[d]	7 (1.3)	9 (1.7)	16 (1.5)
Black/African/Caribbean/Black British[e]	7 (1.3)	3 (0.6)	10 (1.0)
Other ethnic group[f]	5 (1.0)	4 (0.8)	9 (0.9)
Prefer not to say	2 (0.4)	0 (0.0)	2 (0.2)
Unknown	1 (0.2)	0 (0.0)	1 (0.1)

continued

TABLE 2 Baseline demographics and characteristics of trial participants (ITT population) (*continued*)

Demographic or characteristic	TAU group (*N* = 525)	ES group (*N* = 523)	Total (*N* = 1048)
Current cigarette smoker,[g] *n* (%)	13 (2.5)	11 (2.1)	24 (2.3)
Number of cigarettes per day			
	(*n* = 13)	(*n* = 10)	(*n* = 23)
Mean (SD)	6.2 (5.5)	8.2 (5.6)	7.1 (5.5)
Median (IQR)	3.0 (2.0–10.0)	10.0 (2.0–10.0)	7.0 (2.0–10.0)
Minimum, maximum	1.0, 17.0	1.0, 20.0	1.0, 20.0
Alcohol drinker, *n* (%)	286 (54.5)	278 (53.2)	564 (53.8)
Alcohol intake (units per week)[h]			
	(*n* = 279)	(*n* = 274)	(*n* = 553)
Mean (SD)	4.4 (3.2)	4.6 (4.1)	4.5 (3.7)
Median (IQR)	4.0 (2.0–6.0)	3.0 (2.0–6.0)	4.0 (2.0–6.0)
Minimum, maximum	1.0, 18.0	1.0, 20.0	1.0, 20.0
Current recreational drug user, *n* (%)	0 (0.0)	0 (0.0)	0 (0.0)
History of fertility treatment (not IVF), *n* (%)	89 (17.0)	109 (20.8)	198 (18.9)
Fertility treatment received, *n* (%)			
	(*n* = 89)	(*n* = 109)	(*n* = 198)
IUI	55 (61.8)	66 (60.6)	121 (61.1)
Clomifene citrate (Clomid, SANOFI, Paris, France)	31 (34.8)	40 (36.7)	71 (35.9)
IUI and Clomid	0 (0.0)	2 (1.8)	2 (1.0)
Donor insemination	1 (1.1)	1 (0.9)	2 (1.0)
Tamoxifen citrate (Tamoxifen, Wockhardt UK Ltd, Wrexham, UK) and cabergoline (Dostinex, Pfizer, New York, NY, USA)	1 (1.1)	0 (0.0)	1 (0.5)
Had other significant medical conditions, *n* (%)	139 (26.5)	122 (23.3)	261 (24.9)
History of any previous pregnancies	150 (28.6)	155 (29.6)	305 (29.1)
Planned method of fertilisation, *n* (%)			
IVF	319 (60.8)	316 (60.4)	635 (60.6)
ICSI	206 (39.2)	207 (39.6)	413 (39.4)
Planned treatment protocol, *n* (%)			
Antagonist	313 (59.6)	308 (58.9)	621 (59.3)
Long protocol	212 (40.4)	215 (41.1)	427 (40.7)
Planned cycle programming, *n/N* (%)	131/313 (41.9)	126/308 (40.9)	257/621 (41.4)
Cycle programming details			
	(*n* = 131)	(*n* = 126)	(*n* = 257)
Oral contraception	70 (53.4)	67 (53.2)	137 (53.3)
Progestogens	54 (41.2)	52 (41.3)	106 (41.2)
Oral oestrogen	7 (5.3)	7 (5.6)	14 (5.4)
Duration of infertility (years)[i]	*n* = 525	*n* = 523	*n* = 1048
Mean (SD)	3.1 (1.7)	3.1 (1.9)	3.1 (1.8)
Median (IQR)	2.8 (2.0–3.7)	2.8 (2.0–3.5)	2.8 (2.0–3.5)
Minimum, maximum	0.0, 15.0	0.0, 18.0	0.0, 18.0

TABLE 2 Baseline demographics and characteristics of trial participants (ITT population) (*continued*)

Demographic or characteristic	TAU group (N = 525)	ES group (N = 523)	Total (N = 1048)
Previous pregnancies, *n* (%)			
0	375 (71.4)	368 (70.4)	743 (70.9)
1	103 (19.6)	109 (20.8)	212 (20.2)
2	34 (6.5)	33 (6.3)	67 (6.4)
3	10 (1.9)	4 (0.8)	14 (1.3)
4	2 (0.4)	5 (1.0)	7 (0.7)
≥ 5	1 (0.2)	4 (0.8)	5 (0.5)
Previous miscarriages, *n* (%)			
0	442 (84.2)	437 (83.6)	879 (83.9)
1	65 (12.4)	68 (13.0)	133 (12.7)
2	12 (2.3)	12 (2.3)	24 (2.3)
≥ 3	6 (1.1)	6 (1.1)	12 (1.1)
Previous terminations, *n* (%)			
0	479 (91.2)	471 (90.1)	950 (90.6)
1	41 (7.8)	48 (9.2)	89 (8.5)
≥ 2	5 (1.0)	4 (0.8)	9 (0.9)
Previous stillbirths, *n* (%)			
0	522 (99.4)	520 (99.4)	1042 (99.4)
1	3 (0.6)	3 (0.6)	6 (0.6)
Previous live births, *n* (%)			
0	501 (95.4)	497 (95.0)	998 (95.2)
1	21 (4.0)	22 (4.2)	43 (4.1)
≥ 2	3 (0.6)	4 (0.8)	7 (0.7)
Previous ectopic pregnancies, *n* (%)			
0	508 (96.8)	504 (96.4)	1012 (96.6)
1	13 (2.5)	13 (2.5)	26 (2.5)
≥ 2	4 (0.8)	6 (1.1)	10 (1.0)
Parity, *n* (%)[j]			
0	501 (95.4)	498 (95.2)	999 (95.3)
1	21 (4.0)	22 (4.2)	43 (4.1)
≥ 2	3 (0.6)	3 (0.6)	6 (0.6)

a Only in women with successful egg collection.
b Ethnicity classification: English/Welsh/Scottish/Northern Irish/British, Irish, Gypsy or Irish Traveller, and any other white background.
c Ethnicity classification: Indian, Pakistani, Bangladeshi, Chinese, and any other Asian background.
d Ethnicity classification: white and black Caribbean, white and black African, white and Asian, and any other mixed/multiple ethnic groups background.
e Ethnicity classification: African, Caribbean and any other black/African/Caribbean/black British background.
f Ethnicity classification: Arab, and any other ethnic group.
g Current smoker relates to smoking cigarettes and not vaping.
h A few alcohol drinkers had missing alcohol intake data.
i Twelve participants had zero duration of infertility (these were women seeking treatment or couples in same-sex relationships without any known fertility problems). Parity is defined as the number of times a woman gave birth to a fetus (either a live or stillbirth) with a gestational age of at least 24 weeks.
j Parity could not be ascertained in eight births or stillbirths without gestational age: TAU (*n* = 3) and ES (*n* = 5).

In other cases, it was not clear whether the medical condition that would have made a participant ineligible developed after randomisation or if it was missed during screening prior to randomisation. However, the number of ineligibility errors was negligible relative to the number of women randomised and the number was similar in both treatment groups.

Notably, three women in the ES group were found to have a hydrosalpinx during the trial. This was not a stated exclusion criterion in the protocol and sites were not prompted to report this during data collection, so the numbers could be slightly higher.

Uptake of the interventions

In general, the uptake of the ES procedure (followed by IVF) in the ES group and of IVF in the TAU group was reasonably satisfactory.

Endometrial scratch procedure and in vitro fertilisation in the endometrial scratch group and time to in vitro fertilisation after endometrial scratch procedure

Figure 3 summarises the uptake of the ES procedure and IVF among those randomised to the ES group, including reasons for failing for the 13.4% (70/523) of participants who did not receive the intervention. Of those randomised to ES, 86.6% (453/523) received the ES procedure within the trial. All participants who received ES received it as per protocol.
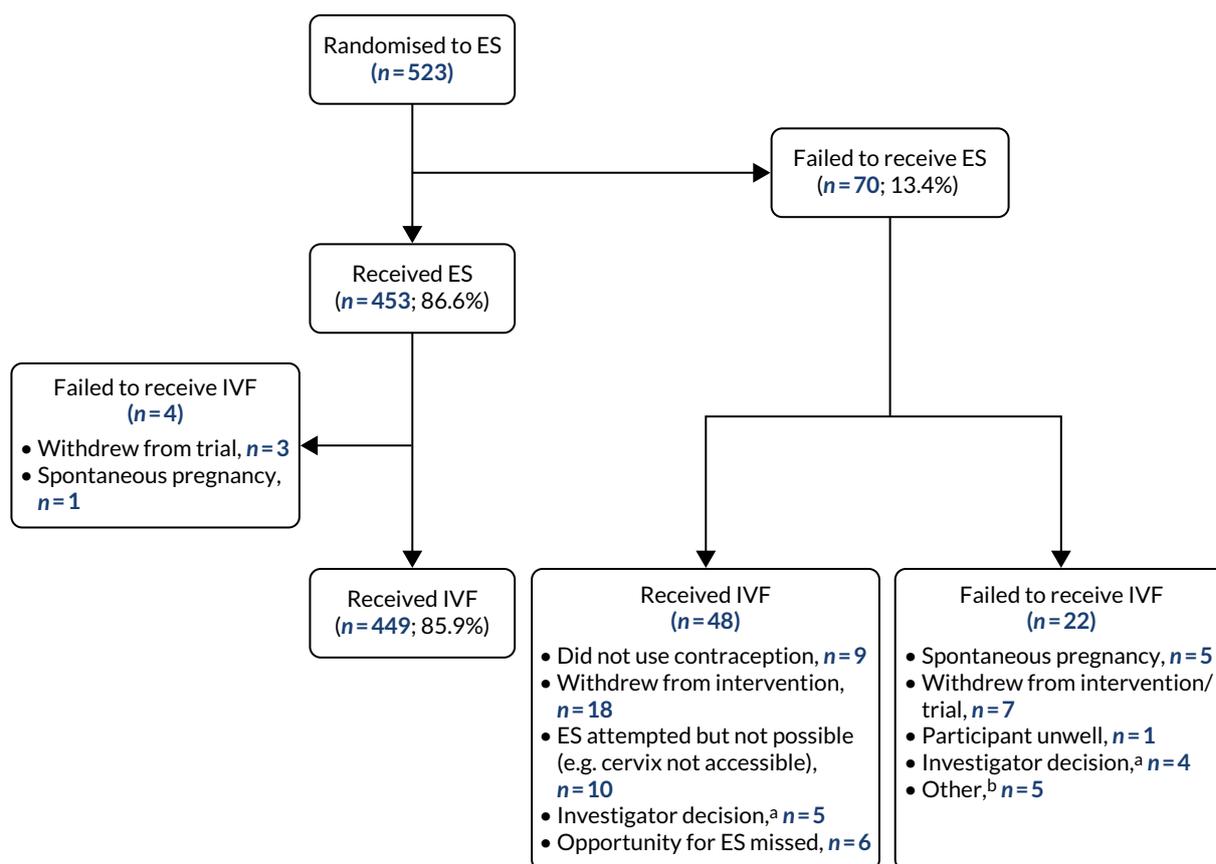
FIGURE 3 Uptake of the ES procedure and IVF in the ES group. a, Investigator decisions (*n* = 9): deemed too risky to perform ES as the participant had a previous severe infection following IUI (*n* = 1); risk of perforation due to stenosis of the cervical canal (*n* = 1); lost to follow-up (*n* = 1); recently had surgery for adhesions (*n* = 1); had partial septate uterus on the scan (*n* = 1); identified adenomyosis within the uterus (*n* = 1); had leuprorelin acetate (Prostap, Takeda UK Ltd, London, UK) (*n* = 2); participant had a scratch procedure in the past year which was noted when an endometrial biopsy was taken during exploratory surgery (*n* = 1). b, Other reasons (*n* = 5): lost to follow-up (*n* = 1), never returned to commence treatment (*n* = 3) and BMI was too high to start treatment (*n* = 1).

Health Technology Assessment 2022 Vol. 26 No. 10

Overall, 95.0% (497/523) of those randomised to ES received IVF; 85.9% (449/523) received IVF after receiving ES procedure and 9.2% (48/523) received IVF without undergoing ES procedure.

Among the 449 women in the ES group who received the ES and IVF, the median (IQR) time to the start of IVF post ES procedure was 15 (9–22) days, ranging from 2 to 126 days (*Figure 4a*). The mean time (SD) was 17.1 (11.5) days. In summary, the majority of women started their FSH drugs within 30 days of the ES procedure. It should be noted that the date when FSH drugs were started was used as a proxy for the start of IVF after the ES procedure. *Figure 4b* also displays the distribution of time from ES procedure to embryo transfer.
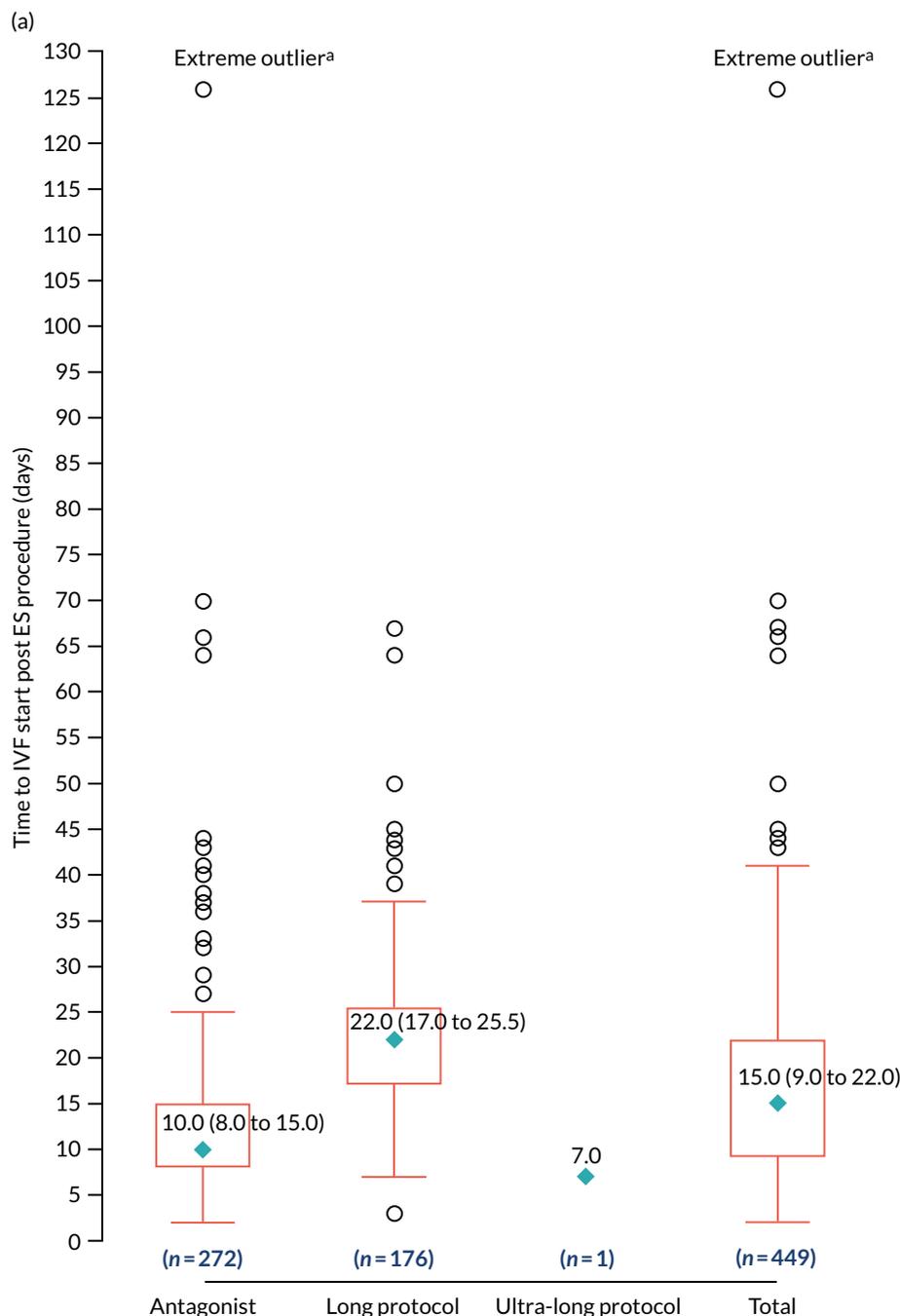
FIGURE 4 Distribution of time from ES procedure to (a) start of IVF and (b) embryo transfer. a, Extreme outlier: participant developed a tubal ovarian abscess and required surgery before starting IVF. (*continued*)
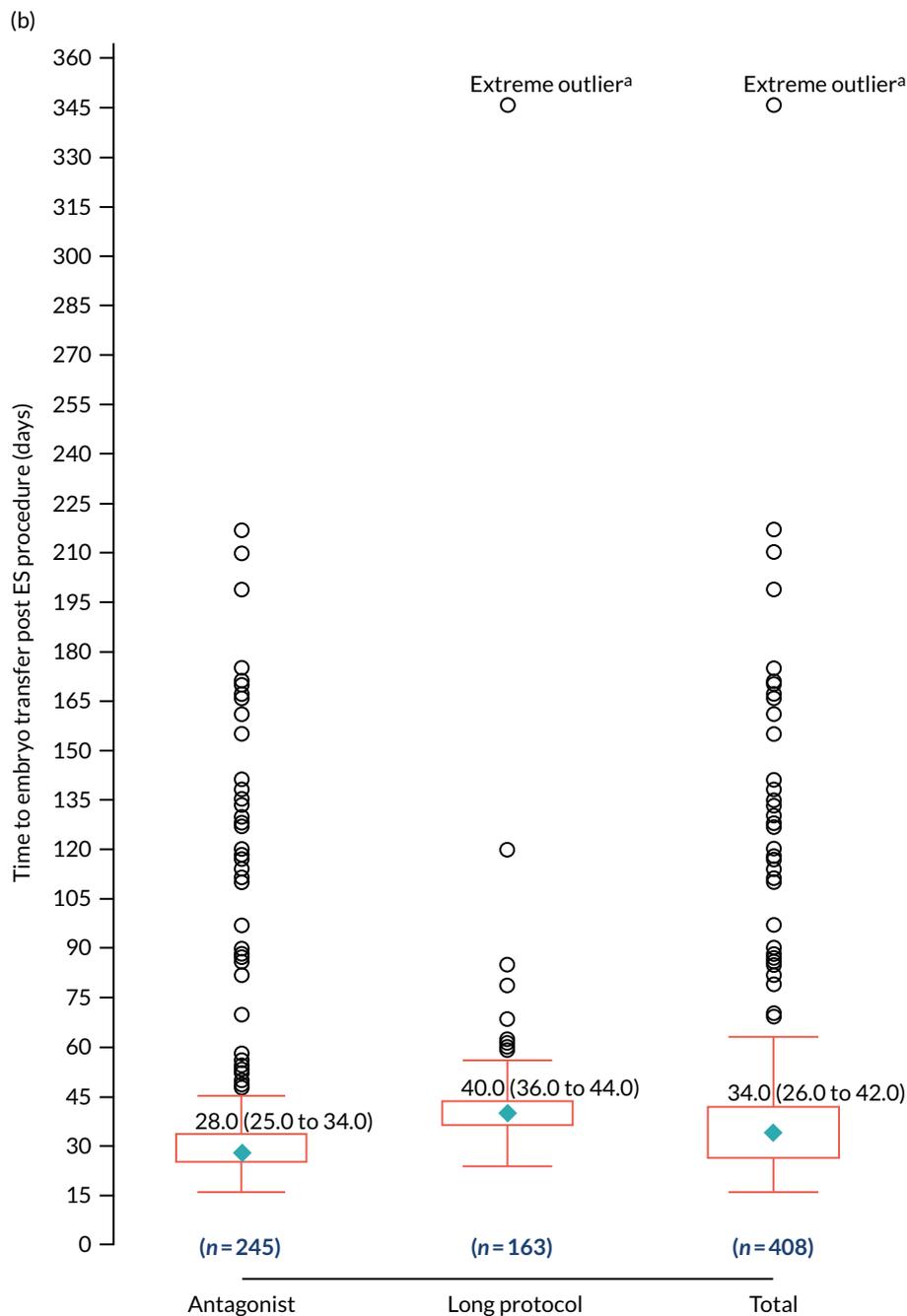
FIGURE 4 Distribution of time from ES procedure to (a) start of IVF and (b) embryo transfer. a, Extreme outlier: participant developed a tubal ovarian abscess and required surgery before starting IVF.

In vitro fertilisation in the treatment-as-usual group

Of those allocated to TAU, 94.1% (494/525) received IVF (*Figure 5*). Only 5.9% (31/525) of participants did not receive IVF, for reasons that are detailed in *Figure 5*. Only 1.0% (5/494) of IVF recipients received ES outside the trial.

Treatment cycle characteristics

Table 3 summarises the treatment cycle characteristics of women before egg collection. In general, these characteristics are very similar across treatment groups. In summary, 94.6% (991/1048) of women in total received IVF: 94.1% (494/525) in the TAU group and 95.0% (497/523) in the ES group. In total, of those that received IVF, 588 (59.3%) and 402 (40.6%) were treated using the antagonist
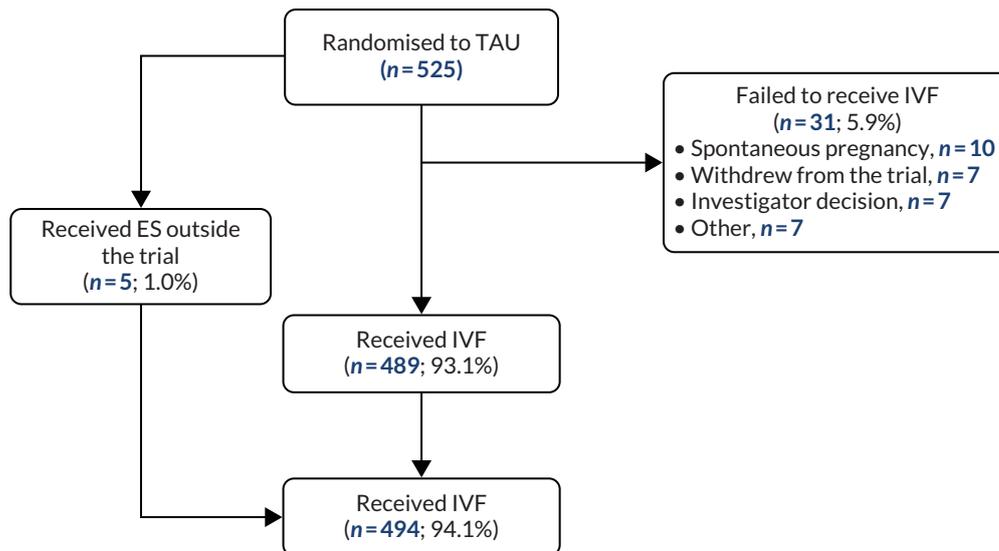
FIGURE 5 Uptake of IVF in the TAU group.

TABLE 3 Treatment cycle characteristics before egg collection

Characteristic	TAU group (N = 494)	ES group (N = 497)	Total (N = 991)
Treatment protocol followed, n (%)			
Antagonist	297 (60.1)	291 (58.6)	588 (59.3)
Long protocol	197 (39.9)	205 (41.2)	402 (40.6)
Ultra-long protocol	0 (0.0)	1 (0.2)	1 (0.1)
FSH drug used, n (%)			
Follitropin alfa (GONAL-f, Merck, Kenilworth, NJ, USA)	153 (29.1)	158 (30.2)	311 (29.7)
Menotrophin BP (Merional, Ferring Pharmaceuticals Ltd, Saint-Prex, Switzerland)	127 (24.2)	118 (22.6)	245 (23.4)
Menotrophin (MENOPUR, Ferring Pharmaceuticals Ltd)	127 (24.2)	135 (25.8)	262 (25.0)
Follitropin alfa [Bemfola, Gedeon Richter (UK) Ltd, London, UK]	46 (8.8)	46 (8.8)	92 (8.8)
Urofollitropin (Fostimon, Pharmasure Ltd, Watford, UK)	8 (1.5)	7 (1.3)	15 (1.4)
Menotrophin (Meriofert, Pharmasure Ltd)	27 (5.1)	24 (4.6)	51 (4.9)
Follitropin alfa (Ovaleap, Theramex UK Ltd, London, UK)	4 (0.8)	4 (0.8)	8 (0.8)
Follitropin delta (Rekovelle, Ferring Pharmaceuticals Ltd)	1 (0.2)	1 (0.2)	2 (0.2)
Bemfola and Menopur	1 (0.2)	0 (0.0)	1 (0.1)
Gonal F and Menopur	0 (0.0)	2 (0.4)	2 (0.2)
Merional and Fostimon	0 (0.0)	1 (0.2)	1 (0.1)
Class of FSH drug used, n (%)			
Recombinant	204 (38.9)	209 (40.0)	413 (39.4)
Purified urinary	289 (55.0)	285 (54.5)	574 (54.8)
Purified urinary and recombinant combination	1 (0.2)	2 (0.4)	3 (0.3)
Purified urinary combination	0 (0.0)	1 (0.2)	1 (0.1)

continued

TABLE 3 Treatment cycle characteristics before egg collection (*continued*)

Characteristic	TAU group (N = 494)	ES group (N = 497)	Total (N = 991)
Planned treatment protocol changed, *n* (%)			
No	488 (98.8)	489 (98.4)	977 (98.6)
Yes	6 (1.2)	8 (1.6)	14 (1.4)
Trigger, *n* (%)			
hCG	468 (94.7)	456 (91.8)	924 (93.2)
Agonist	18 (3.6)	28 (5.6)	46 (4.6)
Not administered	8 (1.6)	13 (2.6)	21 (2.1)
Number of days of FSH (days)	*n* = 494	*n* = 496	*n* = 990
Mean (SD)	10.9 (1.9)	11.0 (2.2)	10.9 (2.0)
Median (IQR)	11.0 (10.0–12.0)	11.0 (10.0–12.0)	11.0 (10.0–12.0)
Minimum, maximum	5.0, 19.0	3.0, 30.0	3.0, 30.0
Performed cycle programming, *n* (%)	124 (23.6)	118 (22.6)	242 (23.1)
Details of cycle programming	*n* = 124	*n* = 118	*n* = 242
Oral contraception, *n* (%)	64 (51.6)	63 (53.4)	127 (52.5)
Progestogens, *n* (%)	53 (42.7)	48 (40.7)	101 (41.7)
Oral oestrogen, *n* (%)	7 (5.6)	7 (5.9)	14 (5.8)
Taking any medications that contraindicate IVF, *n* (%)	0 (0.0)	0 (0.0)	0 (0.0)

hCG, human chorionic gonadotropin.
Recombinant: Gonal F, Bemfola, Rekovelle and Ovaleap. Purified urinary: Merional, Menopur, Fostimon and Meriofert. Purified urinary and recombinant combination: Bemfola and Menopur, Gonal F and Menopur. Purified urinary FSH combination: Merional and Fostimon.

and long protocols, respectively. Only one woman received the ultra-long protocol in the ES group (a protocol deviation). The treatment protocol was only changed in the treatment of 14 (1.4%) women (TAU = 6, ES = 8). Just over half of the FSH drugs administered were purified urinary drugs. No one was taking any medications known to contraindicate IVF.

Table 4 summarises the treatment cycle characteristics of women at egg collection. In general, these characteristics are very similar across treatment groups. Among those who received IVF, egg collection was successful in 482 (97.6%) women in the TAU group and in 481 (96.8%) women in the ES group. Poor response to treatment drugs was the main reason for unsuccessful egg collection. The mean (SD) number of eggs collected was 12 (6.4) in the TAU group and 12 (6.6) in the ES group. Successful egg fertilisation was very high in both groups. IVF and ICSI split fertilisation method was used in only 22 (2.3%) women (seven in the TAU group and 15 in the ES group). Fresh sperm were predominantly used (90.9%). Embryo generation rates were the same and distribution of the number of embryos generated was similar across treatment groups.

Table 5 summarises the treatment cycle characteristics of women at embryo transfer. These characteristics are broadly very similar across treatment groups. The majority of embryo transfers were fresh embryos (91.5%), single embryos (82.2%) and day 5 transfers (78.2%). Hyperstimulation was the main reason for not transferring embryos. Other reasons among 18 women are detailed as footnotes in *Table 5*. Only 52 (5.7%) embryo transfers were viewed as difficult (TAU = 24, ES = 28).

TABLE 4 Treatment cycle characteristics at egg collection

Characteristic	TAU group (N = 494)	ES group (N = 497)	Total (N = 991)
Eggs collected, n (%)	482 (97.6)	481 (96.8)	963 (97.2)
Eggs not collected with reasons, n (%)	12 (2.4)	16 (3.2)	28 (2.8)
Empty follicles, n	1	1	2
Poor response, n	8	12	20
Participant withdrew from IVF/trial, n	1	1	2
Other,[a] n	2	2	4
Number of eggs collected	n = 482	n = 481	n = 963
Mean (SD)	12.0 (6.4)	12.0 (6.6)	12.0 (6.5)
Median (IQR)	11.0 (8.0–15.0)	11.0 (8.0–16.0)	11.0 (8.0–15.0)
Minimum, maximum	1.0, 39.0	1.0, 44.0	1.0, 44.0
Successful egg fertilisation, n/N (%)	476/482 (98.8)	465/481 (96.7)	941/963 (97.7)
Method of fertilisation, n (%)			
IVF	264 (55.5)	241 (51.8)	505 (53.7)
ICSI	205 (43.1)	209 (44.9)	414 (44.0)
IVF and ICSI split	7 (1.5)	15 (3.2)	22 (2.3)
Sperm used, n (%)			
Fresh	434 (91.2)	421 (90.5)	855 (90.9)
Frozen	42 (8.8)	44 (9.5)	86 (9.1)
Type of sperm used, n (%)			
Ejaculate	432 (90.8)	429 (92.3)	861 (91.5)
Donor	24 (5.0)	20 (4.3)	44 (4.7)
PESA	6 (1.3)	8 (1.7)	14 (1.5)
TESA	2 (0.4)	3 (0.6)	5 (0.5)
TESE	10 (2.1)	3 (0.6)	13 (1.4)
MicroTESE	2 (0.4)	2 (0.4)	4 (0.4)
Method of fertilisation changed since plan, n (%)	26 (5.5)	28 (6.0)	54 (5.7)
Sperm quality, n	23	23	46
Other,[b] n	3	5	8
Embryos generated day 1 after fertilisation, n (%)	475 (99.8)	464 (99.8)	939 (99.8)
Number of embryos generated	n = 475	n = 464	n = 939
Mean (SD)	7.0 (4.6)	7.0 (4.5)	7.0 (4.5)
Median (IQR)	6.0 (4.0–9.0)	6.0 (4.0–9.0)	6.0 (4.0–9.0)
Minimum, maximum	1.0, 28.0	1.0, 31.0	1.0, 31.0

MicroTESE, microsurgical testicular sperm extraction; PESA, percutaneous epididymal sperm aspiration; TESA, testicular sperm aspiration; TESE, testicular sperm extraction.
a Other reasons for not collecting eggs (n = 4): cycle cancelled due to no maturing follicles (n = 1), cycle abandoned after 11 days as tachycardia and thyroid functional tests abnormal and high risk with general anaesthetic (n = 1), cycle cancelled as the participant has ovarian hyperstimulation syndrome (n = 1) and microsurgical testicular sperm extraction proved unsuccessful due to no sperm being retrieved (n = 1).
b Other reasons for the change in fertilisation method (n = 8): low sperm concentration (n = 1), no sperm and donor back-up used (n = 1), the patient suffered a bereavement so plan was changed to create embryos and freeze them (n = 1), prolonged ovarian stimulation resulted in the cancellation of planned surgical sperm retrieval and treatment converted to IVF using donor sperm (n = 1), the patient could not produce semen and semen was produced at home so there was a time delay with sperm production (n = 1), poor response to sperm asseparation (n = 1), patient choice (n = 1) and hyperstimulation (n = 1).

TABLE 5 Treatment cycle characteristics at embryo transfer

Characteristic[a]	TAU group (N = 475)	ES group (N = 464)	Total (N = 939)
Embryos transferred, n (%)	464 (97.7)	455 (98.1)	919 (97.9)
Fresh, n (%)	425 (91.6)	416 (91.4)	841 (91.5)
Frozen, n (%)	39 (8.4)	39 (8.6)	78 (8.5)
Reasons for not transferring fresh embryos, n (%)	n = 50	n = 48	n = 98
Abnormal uterine cavity	2 (4.0)	3 (6.3)	5 (5.1)
Hyperstimulation	32 (64.0)	30 (62.5)	62 (63.3)
No suitable embryos to transfer	7 (14.0)	6 (12.5)	13 (13.3)
Other[b]	9 (18.0)	9 (18.8)	18 (18.4)
Difficult embryo transfer, n (%)	24 (5.2)	28 (6.2)	52 (5.7)
Number of embryos transferred, n (%)			
Single	374 (80.6)	381 (83.7)	755 (82.2)
Double	90 (19.4)	74 (16.3)	164 (17.8)
Day of embryo transfer, n (%)	n = 464	n = 455	n = 919
2	27 (5.8)	29 (6.4)	56 (6.1)
3	57 (12.3)	68 (14.9)	125 (13.6)
4	11 (2.4)	6 (1.3)	17 (1.8)
5	367 (79.1)	352 (77.4)	719 (78.2)
6	2 (0.4)	0 (0.0)	2 (0.2)
Type of catheter used, n (%)	n = 464	n = 455	n = 919
Cook K Jet (Cook Medical, Limerick, Ireland)	81 (17.5)	90 (19.8)	171 (18.6)
Guardia™ AccessET Embryo Transfer Catheter (Cook Medical)	137 (29.5)	122 (26.8)	259 (28.2)
Wallace® Classic (CooperSurgical®, Målov, Denmark)	79 (17.0)	70 (15.4)	149 (16.2)
Wallace® Sure-Pro® (CooperSurgical®)	120 (25.9)	131 (28.8)	251 (27.3)
Labotect (Labotect GMBH, Gottingen, Germany)	2 (0.4)	1 (0.2)	3 (0.3)
Rocket Embryon® (Rocket Medical plc, Watford, England)	21 (4.5)	19 (4.2)	40 (4.4)
Wallace® SureView (CooperSurgical®)	24 (5.2)	22 (4.8)	46 (5.0)
Blastocyst 1, n (%); quality	355 (76.5)	345 (75.8)	700 (76.2)
	n = 464	n = 455	n = 919
Excellent	87 (18.8)	102 (22.4)	189 (20.6)
Very good	92 (19.8)	109 (24.0)	201 (21.9)
Good	95 (20.5)	74 (16.3)	169 (18.4)
Fair and freezable	35 (7.5)	28 (6.2)	63 (6.9)
Fair	21 (4.5)	15 (3.3)	36 (3.9)
Poor	13 (2.8)	7 (1.5)	20 (2.2)
Early blastocyst	12 (2.6)	10 (2.2)	22 (2.4)

TABLE 5 Treatment cycle characteristics at embryo transfer (*continued*)

Characteristic[a]	TAU group (N = 475)	ES group (N = 464)	Total (N = 939)
Blastocyst 2, *n* (%); quality	30 (6.5)	22 (4.8)	52 (5.7)
	n = 464	*n* = 455	*n* = 919
Excellent	1 (0.2)	0 (0.0)	1 (0.1)
Very good	2 (0.4)	3 (0.7)	5 (0.5)
Good	5 (1.1)	2 (0.4)	7 (0.8)
Fair and freezable	9 (1.9)	3 (0.7)	12 (1.3)
Fair	4 (0.9)	3 (0.7)	7 (0.8)
Poor	5 (1.1)	3 (0.7)	8 (0.9)
Early blastocyst	4 (0.9)	8 (1.8)	12 (1.3)
Cleavage 1, *n* (%); quality	77 (16.6)	94 (20.7)	171 (18.6)
	n = 464	*n* = 455	*n* = 919
Excellent	5 (1.1)	7 (1.5)	12 (1.3)
Good	34 (7.3)	41 (9.0)	75 (8.2)
Fair	2 (0.4)	4 (0.9)	6 (0.7)
Poor	15 (3.2)	15 (3.3)	30 (3.3)
Very poor	2 (0.4)	1 (0.2)	3 (0.3)
Slow	7 (1.5)	10 (2.2)	17 (1.8)
Arrested development	1 (0.2)	0 (0.0)	1 (0.1)
Ungraded[c]	11 (2.4)	16 (3.5)	27 (2.9)
Cleavage 2, *n* (%); quality	43 (9.3)	40 (8.8)	83 (9.0)
	n = 464	*n* = 455	*n* = 919
Excellent	4 (0.9)	1 (0.2)	5 (0.5)
Good	9 (1.9)	10 (2.2)	19 (2.1)
Fair	1 (0.2)	2 (0.4)	3 (0.3)
Poor	16 (3.4)	12 (2.6)	28 (3.0)
Slow	5 (1.1)	10 (2.2)	15 (1.6)
Arrested development	1 (0.2)	0 (0.0)	1 (0.1)
Not graded (Gardner blastocyst grading system)	7 (1.5)	5 (1.1)	12 (1.3)
Morula 1 (not graded), *n* (%)	32 (6.9)	16 (3.5)	48 (5.2)
Morula 2 (not graded), *n* (%)	17 (3.7)	12 (2.6)	29 (3.2)
Blood presence on tip of catheter, *n* (%)	34 (7.3)	27 (5.9)	61 (6.6)
Fluid in the endometrium, *n* (%)	3 (0.6)	4 (0.9)	7 (0.8)
Vulsellum used, *n* (%)	0 (0.0)	5 (1.1)	5 (0.5)

a Other reasons that fresh embryos were not transferred (*n* = 18): participant request to freeze embryos owing to sudden family bereavement (*n* = 1); participant's request to freeze embryos (*n* = 1); reason for freeze is unknown (*n* = 1); personal issues (*n* = 1); unable to do the transfer (*n* = 1); acutely anteverted anteflexed uterus, scarring, adhesions at endocervix (*n* = 1); hydrosalpinges requiring intervention (*n* = 3); risk of ovarian hyperstimulation syndrome (*n* = 3); planned FET that was a protocol violation (*n* = 1); raised progesterone (*n* = 3); endometrial polyp found after IVF cycle started (*n* = 2); and breaking down of a relationship (*n* = 1).
b The denominator is embryo generated.
c Grading was not done in sites that used the Gardner blastocyst grading system.

Acceptability of the endometrial scratch procedure

Pain relief medication was administered prior to the ES procedure in 78.8% (357/453) of women who went on to receive ES, of whom 86.0% (307/357) received paracetamol, 13.4% (48/357) ibuprofen, 0.3% (1/357) aspirin, 3.4% (12/357) co-codamol and 3.4% (12/357) other medications.

Complete tolerability data were available for 449 (99.1%) of the 453 women who underwent ES, of whom 99.8% (448/449) viewed the ES procedure as tolerable (95% CI 98.7% to 100.0%). Tolerability was similar (98.9%; 95% CI 97.4% to 99.5%) even if assuming that the additional four women with incomplete data did not tolerate the ES procedure.

The pain was subjectively assessed on a rating scale of 0 (no pain) to 10 (worst pain imaginable). *Figure 6* shows the distribution of pain rating scores within 30 minutes and at 1 day and 7 days after the ES procedure, stratified by whether or not ambiguous pain scores were included. Descriptive summaries of the distribution of these pain ratings, including the number of participants included, are provided in *Appendix 13*.

In summary, the ES was viewed as a well-tolerated procedure. The distributions of pain ratings with or without ambiguous scores were very similar. Furthermore, exploratory analysis (not presented here) showed that the ratings within 30 minutes were also consistent across the CRF versions used (minor wording clarifications were made early on during the trial).

The effectiveness of endometrial scratch procedure on the primary outcome (live birth)

Table 6 summarises the unadjusted effect of ES on the chances of achieving a live birth. For the primary ITT analysis set, the LBR was 37.1% (195/525) in the TAU group and 38.6% (202/523) in the ES group, translating to a 1.5% absolute increase in the LBR attributed to the ES intervention (95% CI −4.4% to 7.4%; $p = 0.621$). In other words, the chance of having a live birth was only 1.04 times more (in relative terms) in women who received the ES procedure than in women who received TAU. The CI includes zero and does not include the assumed targeted 10% AD prespecified in the sample size calculation (see *Chapter 2, Sample size*) and viewed as an effect likely to change practice.

In summary, LBRs are very similar between treatment groups and we can conclude that there was no reliable statistical evidence to support the theory that ES procedure before IVF increases the chance of
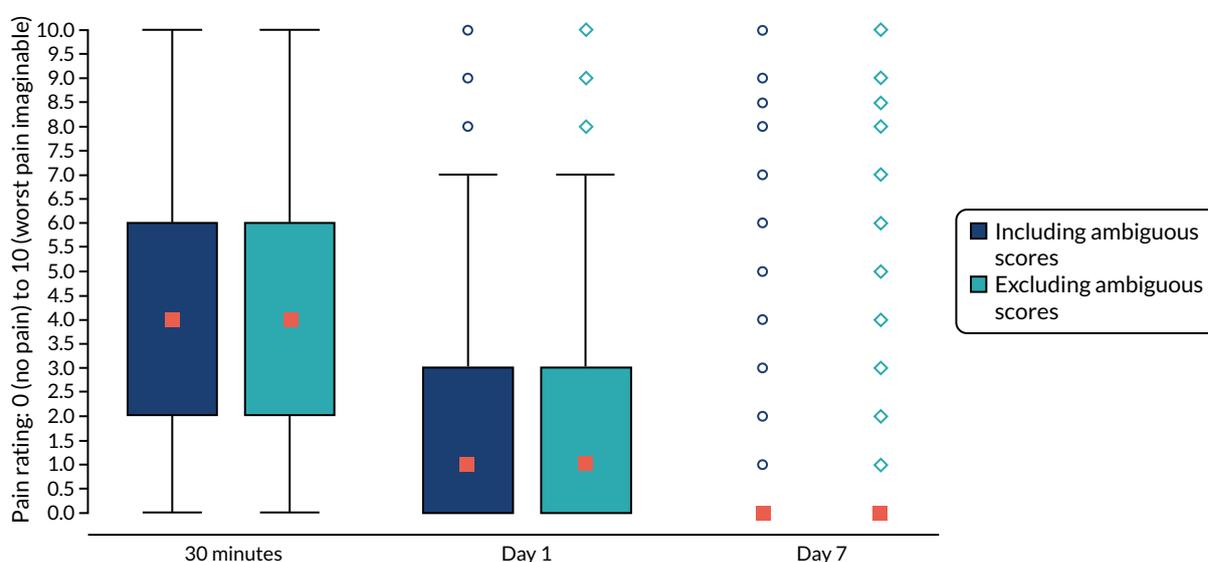
FIGURE 6 Distribution of pain rating of ES procedure.

TABLE 6 The effect of ES on achieving a live birth (primary outcome)

Primary outcome	TAU group, n/N (%)	ES group, n/N (%)	Unadjusted treatment effect (95% CI)			
			Risk difference	OR	RR	p-value
LBR (worst case)	195/525 (37.1)	202/523 (38.6)	1.5% (−4.4% to 7.4%)	1.06 (0.83 to 1.37)	1.04 (0.89 to 1.21)	0.621
Sensitivity analysis						
LBR (best case)	197/525 (37.5)	205/523 (39.2)	1.7% (−4.2% to 7.6%)	1.07 (0.84 to 1.38)	1.04 (0.90 to 1.22)	0.578
LBR (complete case)	195/523 (37.3)	202/520 (38.8)	1.6% (−4.3% to 7.5%)	1.07 (0.83 to 1.37)	1.04 (0.89 to 1.22)	0.604

Best-case scenario assumes that patients who were lost to follow-up or had an unknown pregnancy outcome at the end of the trial, for any reason, became pregnant and had a live birth outcome. The complete-case analysis includes only those women with known pregnancy and a live birth outcome at the end of the trial.

having a live birth compared with IVF in women undergoing IVF for the first time. Sensitivity analysis results using the best-case scenario and complete cases are very similar to the primary results and so are the conclusions (see *Table 6*).

Table 7 presents results of the effect of ES on the primary outcome adjusted for fixed stratification (site and planned treatment protocol) and potential prognostic factors (age, BMI, history of previous pregnancy, duration of infertility and current smoking status). In summary, the unadjusted (see *Table 6*) and adjusted (see *Table 7*) results are very similar and consistent with a lack of evidence to support the theory that performing ES procedure before IVF increases the chances of having a live birth in women undergoing IVF for the first time.

Per-protocol analysis of the primary outcome (live birth)

We explored the effect of the ES intervention on achieving a live birth (primary outcome) among those who have complied with the key components of the protocol as defined in *Chapter 3, Analysis populations*.

Of those randomised, 79.4% (417/525) in the TAU group and 71.1% (372/523) in the ES group met the criteria for inclusion in the per-protocol analysis population, assuming few women with missing live birth outcome data had a negative pregnancy outcome (worst case). Thus, the proportion of women who met the per-protocol criteria was slightly higher in the TAU group than in the ES group. *Table 8* reports both the unadjusted and adjusted effects of the ES procedure on live birth. The LBR was

TABLE 7 Sensitivity analysis: adjusted effect of ES on achieving a live birth (primary outcome)

Primary outcome	TAU group, n/N (%)	ES group, n/N (%)	Adjusted treatment effect (95% CI)[a]			
			Risk difference	OR	RR	p-value
LBR (worst case)	195/525 (37.1)	202/523 (38.6)	1.3% (−4.4% to 7.1%)	1.07 (0.83 to 1.39)	1.04 (0.89 to 1.21)	0.655
LBR (best case)	197/525 (37.5)	205/523 (39.2)	1.5% (−4.3% to 7.3%)	1.09 (0.84 to 1.40)	1.04 (0.89 to 1.21)	0.608
LBR (complete case)	195/523 (37.3)	202/520 (38.8)	1.4% (−4.4% to 7.2%)	1.08 (0.84 to 1.39)	1.04 (0.89 to 1.21)	0.637

a Adjusted for fixed stratification factors (site and planned treatment protocol) and potential prognostic factors (history of pregnancy, yes or no; age; BMI; duration of infertility; smoking status, yes or no).
Best-case scenario assumes that patients who were lost to follow-up or had an unknown pregnancy outcome at the end of the trial, for any reason, became pregnant and had a live birth outcome. Complete-case analysis includes only those women with known pregnancy and a live birth outcome at the end of the trial.

TABLE 8 Effect of ES on live birth based on per-protocol analysis population

LBR	TAU group, n/N (%)	ES group, n/N (%)	Unadjusted treatment effect (95% CI)[a]			p-value	Adjusted treatment effect (95% CI)[b]			p-value
			AD	OR	RR		AD	OR	RR	
PP (worst case)	169/417 (40.5)	164/372 (44.1)	3.6% (−3.3% to 10.5%)	1.16 (0.87 to 1.54)	1.09 (0.92 to 1.28)	0.312	3.4% (−3.4% to 10.2%)	1.16 (0.86 to 1.54)	1.08 (0.92 to 1.27)	0.323
PP (best case)	170/417 (40.8)	164/372 (44.1)	3.3% (−3.6% to 10.2%)	1.15 (0.86 to 1.52)	1.08 (0.92 to 1.27)	0.346	3.2% (−3.6% to 10.0%)	1.14 (0.86 to 1.53)	1.08 (0.92 to 1.27)	0.360
PP (complete case)	169/416 (40.6)	164/372 (44.1)	3.5% (−3.4% to 10.4%)	1.15 (0.87 to 1.53)	1.09 (0.92 to 1.28)	0.326	3.3% (−3.5% to 10.1%)	1.15 (0.86 to 1.54)	1.08 (0.92 to 1.27)	0.342

PP, per protocol.
a Unadjusted for any covariates except the intervention.
b Adjusted for fixed stratification factors (site and planned treatment protocol) and potential prognostic factors (age, BMI, duration of infertility, history of pregnancy and current smoking status).
n, number of women who had at least one live birth; N, the total number of women.

40.5% (169/417) in the TAU group and 44.1% (164/372) in the ES group. This translates to only a 3.6% increase in live births in favour of ES compared with TAU (95% CI −3.3% to 10.5%; $p = 0.312$). In relative terms, the chance of achieving a live birth is only 1.09 times more (95% CI 0.92 to 1.28) in women who received ES procedure than in those who receive TAU. Although this effect is not statistically significant and could not rule out the lack of effect of ES in improving LBR, it could be of potential clinical importance, as the upper limit of the CI includes a 10% targeted AD. However, based on post hoc analysis [described in *Chapter 3, Analyses of the primary outcome (live birth)*], the observed effect (a 3.6% increase) suggests that such large differences in the LBRs are unlikely, as there is only a 2.9% probability of the effect being at least 10% given the observed 3.6% increase in the LBR. As a result, we did not find enough evidence to support the effect of ES in increasing the chances of achieving a live birth; however, we could not rule out the potential beneficial effect of ES in women who met per-protocol criteria. For comparability, *Figure 7* displays the effect of ES on the live birth based on both ITT and per-protocol analysis populations.

In summary, the adjusted and unadjusted results (see *Table 8*) are very similar, consistent with the fact that the characteristics of participants at baseline and treatment cycle were, on average, well matched between the treatment groups (see *Tables 3–5*).

Subgroup analysis results on the effect of ES on the primary outcome (live birth)

We explored the heterogeneity in the effect of the ES procedure on the primary outcome across prespecified subgroups detailed in *Chapter 3, Prespecified subgroups*. The interpretation of these subgroup results requires caution as the trial was not adequately powered to address this objective and the interaction tests were not controlled for multiple hypothesis tests. Therefore, these results should be used for hypothesis generation guided by biological plausibility and consistency. *Figure 8* displays LBRs, the estimate of the treatment effect within each subgroup and the interaction test. Furthermore, we display the overall results for the primary outcome including per-protocol analysis for comparability.

In summary, the effect of ES appears consistent across subgroups; however, we could not rule out potential benefits when embryos are transferred on day 5 or in women who had cycle programming performed using oral contraception, progestogens or oral oestrogen.
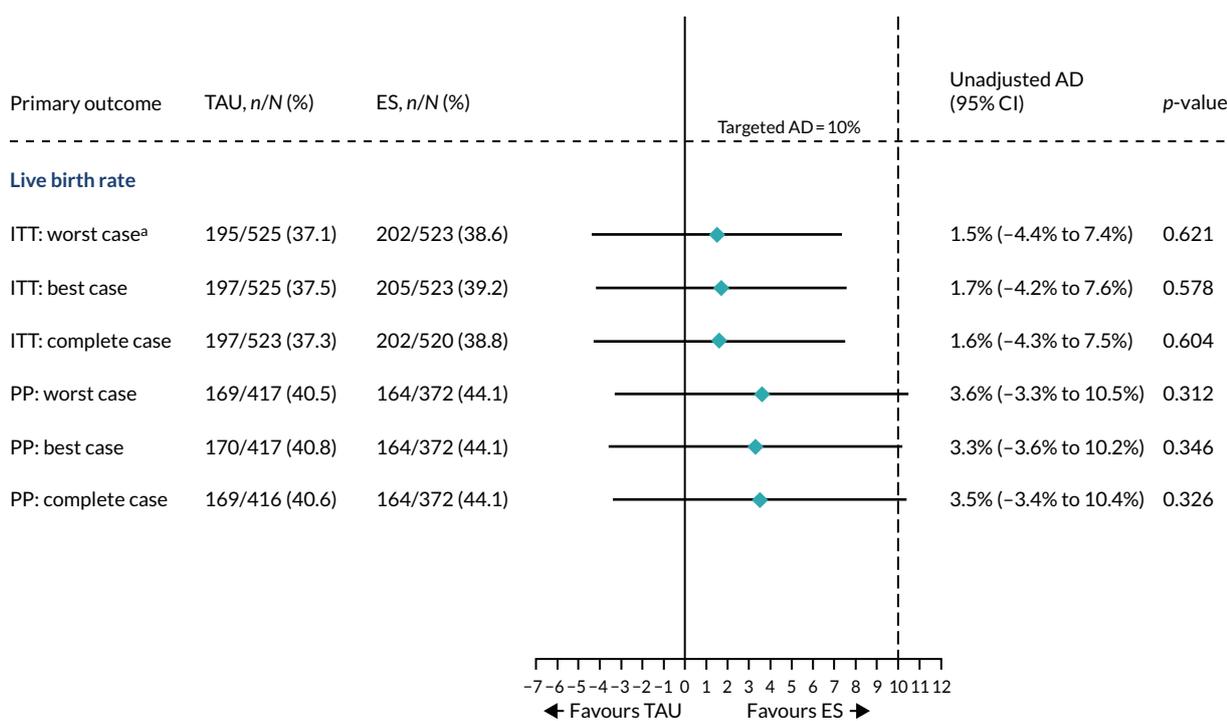
FIGURE 7 Effect of ES on live birth based on ITT and per protocol populations. a, Primary ITT results. PP, per protocol.

Characteristic	TAU, n/N (%)	ES, n/N (%)		Adjusted OR (95% CI)	Interaction test p-value
Treatment protocol					
Antagonist	118/297 (39.7)	116/291 (39.9)		1.02 (0.73 to 1.43)	
Long protocol	70/197 (35.5)	83/205 (40.5)		1.24 (0.82 to 1.88)	0.479
Fertilisation method					
IVF	105/264 (39.8)	110/241 (45.6)		1.30 (0.91 to 1.87)	
ICSI	80/205 (39.0)	84/209 (40.2)		1.08 (0.72 to 1.62)	
IVF and ICSI split	3/7 (42.9)	5/15 (33.3)		0.57 (0.08 to 3.90)	0.603
Embryo transferred					
Single	157/374 (42.0)	169/381 (44.4)		1.13 (0.84 to 1.51)	
Double	31/90 (34.4)	30/74 (40.5)		1.34 (0.70 to 2.57)	0.633
Nature of embryo transferred					
Fresh	173/425 (40.7)	183/416 (44.0)		1.16 (0.88 to 1.53)	
Frozen	15/39 (38.5)	16/39 (41.0)		1.27 (0.50 to 3.21)	0.857
Day of embryo transfer					
2	8/27 (29.6)	5/29 (17.2)		0.56 (0.15 to 2.03)	
3	16/57 (28.1)	17/68 (25.0)		0.89 (0.40 to 2.02)	
4	6/11 (54.5)	2/6 (33.3)		0.37 (0.04 to 3.13)	
5	158/367 (43.1)	175/352 (49.7)		1.32 (0.98 to 1.78)	0.334
Cycle programming					
No	73/173 (42.2)	61/173 (35.3)		0.77 (0.49 to 1.20)	
Yes	45/124 (36.3)	55/118 (46.6)		1.55 (0.92 to 2.62)	0.045
History of miscarriage					
0–2	191/519 (36.8)	200/517 (38.7)		1.09 (0.85 to 1.41)	
≥ 3	4/6 (66.7)	2/6 (33.3)		0.26 (0.02 to 3.07)	0.256
Overall LBR					
ITT: worst case	195/525 (37.1)	202/523 (38.6)		1.06 (0.83 to 1.37)	N/A
ITT: best case	197/525 (37.5)	205/523 (39.2)		1.07 (0.84 to 1.38)	N/A
ITT: complete case	195/523 (37.3)	202/520 (38.8)		1.07 (0.83 to 1.37)	N/A
PP: worst case	169/417 (40.5)	164/372 (44.1)		1.16 (0.87 to 1.54)	N/A
PP: best case	170/417 (40.8)	164/372 (44.1)		1.15 (0.86 to 1.52)	N/A
PP: complete case	169/416 (40.6)	164/372 (44.1)		1.15 (0.87 to 1.53)	N/A

0.00 0.25 0.50 0.75 1.00   1.50   2.00   2.50   3.00   3.50   4.00

◄ Favours TAU   Favours ES ►

FIGURE 8 Forest plot of subgroup results relating to the primary outcome (live birth). Embryos transferred on day 6 were excluded because there were only two women in the TAU group and none in the ES group (see *Table 5*) and no live births were recorded. N/A, interaction test not applicable; PP, per protocol.

We found similar results in OR and RR measures of the treatment effect, respectively (see *Appendices 14* and *15*). Results on the AD scale are not presented as the models failed to converge owing to a small number of participants and events within subgroups.

The effects of endometrial scratch on the secondary outcomes

Appendix 16 summarises the effect of ES on the secondary outcomes based on the ITT analysis population assuming the worst-case scenario. Results of secondary outcomes that also signal the safety of the ES procedure in women with successful implantation (e.g. miscarriages and stillbirths) are presented in *Appendix 17*. These results are displayed in forest plots using different treatment effect scales: using AD (*Figure 9*), ORs (*Figure 10*) and RRs (see *Appendix 18*).

Directional interpretation (favours TAU/favours ES) depends on the outcome.

Secondary outcome	TAU, n/N (%)	ES, n/N (%)		Unadjusted AD (95% CI)	p-value
ITT					
Implantation rate	258/525 (49.1)	253/523 (48.4)		−0.8 (−6.8 to 5.3)	0.804
Clinical pregnancy rate	213/525 (40.6)	223/523 (42.6)		2.1 (−3.9 to 8.0)	0.497
Miscarriage rate	43/525 (8.2)	32/523 (6.1)		−2.1 (−5.2 to 1.0)	0.193
Multiple birth rate	11/525 (2.1)	6/523 (1.1)		−0.9 (−2.5 to 0.6)	0.224
Preterm delivery rate	20/525 (3.8)	14/523 (2.7)		−1.1 (−3.3 to 1.0)	0.301
Ectopic pregnancy rate	2/525 (0.4)	1/523 (0.2)		−0.2 (−0.8 to 0.5)	0.565
Stillbirth rate	1/525 (0.2)	1/523 (0.2)		0.0 (−0.5 to 0.5)	0.998
Pregnancy women (ITT)					
Miscarriage rate	43/258 (16.7)	32/253 (12.6)		−4.0 (−10.1 to 2.1)	0.199
Multiple birth rate	11/258 (4.3)	6/253 (2.4)		−1.9 (−5.0 to 1.2)	0.233
Preterm delivery rate	20/258 (7.8)	14/253 (5.5)		−2.2 (−6.5 to 2.1)	0.314
Ectopic pregnancy rate	2/258 (0.8)	1/253 (0.4)		−0.4 (−1.7 to 0.9)	0.574
Stillbirth rate	1/258 (0.4)	1/253 (0.4)		0.0 (−1.1 to 1.1)	0.989

FIGURE 9 Forest plot of effects of ES on secondary outcomes (AD scale). Directional interpretation (favours TAU/favours ES) depends on the outcome.

Secondary outcome	TAU, n/N (%)	ES, n/N (%)		Unadjusted OR (95% CI)	p-value
ITT					
Implantation rate	258/525 (49.1)	253/523 (48.4)		0.97 (0.76 to 1.24)	0.804
Clinical pregnancy rate	213/525 (40.6)	223/523 (42.6)		1.09 (0.85 to 1.39)	0.497
Miscarriage rate	43/525 (8.2)	32/523 (6.1)		0.73 (0.45 to 1.17)	0.193
Multiple birth rate	11/525 (2.1)	6/523 (1.1)		0.54 (0.20 to 1.48)	0.224
Preterm delivery rate	20/525 (3.8)	14/523 (2.7)		0.69 (0.35 to 1.39)	0.301
Ectopic pregnancy rate	2/525 (0.4)	1/523 (0.2)		0.50 (0.05 to 5.54)	0.565
Stillbirth rate	1/525 (0.2)	1/523 (0.2)		1.00 (0.06 to 16.09)	0.998
Pregnancy women (ITT)					
Miscarriage rate	43/258 (16.7)	32/253 (12.6)		0.72 (0.44 to 1.19)	0.199
Multiple birth rate	11/258 (4.3)	6/253 (2.4)		0.55 (0.20 to 1.50)	0.233
Preterm delivery rate	20/258 (7.8)	14/253 (5.5)		0.70 (0.34 to 1.41)	0.314
Ectopic pregnancy rate	2/258 (0.8)	1/253 (0.4)		0.51 (0.05 to 5.64)	0.574
Stillbirth rate	1/258 (0.4)	1/253 (0.4)		1.02 (0.06 to 16.39)	0.989

FIGURE 10 Forest plot of effects of ES on secondary outcomes (OR scale). Directional interpretation (favours TAU/favours ES) depends on the outcome.

In summary, the rates of all secondary outcomes were similar between the TAU group and the ES group, and the differences between groups were not statistically significant; for example, the clinical pregnancy rate was 40.6% (213/525) in the TAU group and 42.6% (223/523) in the ES group, resulting in only a 2.1% absolute increase in favour of the ES procedure (95% CI –3.9% to 8.0%; $p = 0.497$). There were no PULs recorded during the trial. Ectopic pregnancies and stillbirths were very rare in both groups. Multiple birth and preterm delivery rates were very low and similar across treatment groups. Biochemical pregnancies were recorded in 20 out of 525 (3.8%) participants in the TAU group and 11 out of 523 (2.1%) participants in the ES group relative to the number of women randomised.

Subgroup analysis of secondary outcomes

We explored heterogeneity in the treatment effect of the ES procedure on selected secondary outcomes (*Tables 9* and *10*), excluding those with a small number of events (i.e. ectopic pregnancy, stillbirth, multiple birth and preterm delivery). These were excluded because the numbers of events within each subgroup were insufficient for models to converge and for making a reasonable exploratory inference. It should be noted that implantation and clinical pregnancy outcomes are highly correlated.

TABLE 9 Subgroup analysis results on the effect of ES on implantation

Characteristic/subgroup	TAU group, n/N (%)	ES group, n/N (%)	Adjusted OR (95% CI)	Interaction test, p-value
Treatment protocol				
Antagonist	154/297 (51.9)	140/291 (48.1)	0.87 (0.63 to 1.21)	
Long protocol	94/197 (47.7)	107/205 (52.2)	1.19 (0.80 to 1.79)	0.239
Fertilisation method				
IVF	135/264 (51.1)	130/241 (53.9)	1.12 (0.78 to 1.61)	
ICSI	109/205 (53.2)	112/209 (53.6)	1.06 (0.71 to 1.57)	
IVF and ICSI split	4/7 (57.1)	5/15 (33.3)	0.32 (0.05 to 2.12)	0.439
Embryo transferred				
Single	211/374 (56.4)	210/381 (55.1)	0.96 (0.71 to 1.29)	
Double	37/90 (41.1)	37/74 (50.0)	1.49 (0.79 to 2.83)	0.218
Nature of embryo transferred				
Fresh	227/425 (53.4)	224/416 (53.8)	1.01 (0.77 to 1.33)	
Frozen	21/39 (53.8)	23/39 (59.0)	1.53 (0.61 to 3.85)	0.401
Day of embryo transfer[a]				
2	10/27 (37.0)	8/29 (27.6)	0.75 (0.24 to 2.38)	
3	21/57 (36.8)	24/68 (35.3)	0.93 (0.43 to 1.99)	
4	7/11 (63.6)	2/6 (33.3)	0.30 (0.04 to 2.62)	
5	210/367 (57.2)	213/352 (60.5)	1.16 (0.85 to 1.57)	0.559
Cycle programming				
No	96/173 (55.5)	74/173 (42.8)	0.61 (0.39 to 0.94)	
Yes	58/124 (46.8)	66/118 (55.9)	1.42 (0.85 to 2.39)	0.014
History of miscarriage				
0–2	254/519 (48.9)	251/517 (48.5)	0.99 (0.77 to 1.27)	
≥ 3	4/6 (66.7)	2/6 (33.3)	0.31 (0.03 to 3.63)	0.358

a Day 6 excluded because there were only two women in the TAU group.

TABLE 10 Subgroup analysis results on the effect of ES on clinical pregnancy

Characteristic/subgroup	TAU group, n/N (%)	ES group, n/N (%)	Adjusted OR (95% CI)	Interaction test, p-value
Treatment protocol				
Antagonist	127/297 (42.8)	128/291 (44.0)	1.07 (0.77 to 1.50)	
Long protocol	78/197 (39.6)	91/205 (44.4)	1.22 (0.81 to 1.83)	0.636
Fertilisation method				
IVF	114/264 (43.2)	116/241 (48.1)	1.23 (0.86 to 1.77)	
ICSI	88/205 (42.9)	98/209 (46.9)	1.23 (0.83 to 1.84)	
IVF and ICSI split	3/7 (42.9)	5/15 (33.3)	0.54 (0.08 to 3.67)	0.701
Embryo transferred				
Single	172/374 (46.0)	188/381 (49.3)	1.18 (0.88 to 1.57)	
Double	33/90 (36.7)	31/74 (41.9)	1.28 (0.67 to 2.44)	0.815
Nature of embryo transferred				
Fresh	189/425 (44.5)	199/416 (47.8)	1.16 (0.88 to 1.53)	
Frozen	16/39 (41.0)	20/39 (51.3)	1.79 (0.71 to 4.51)	0.373
Day of embryo transfer[a]				
2	9/27 (33.3)	6/29 (20.7)	0.61 (0.18 to 2.08)	
3	18/57 (31.6)	17/68 (25.0)	0.71 (0.32 to 1.58)	
4	6/11 (54.5)	2/6 (33.3)	0.40 (0.05 to 3.36)	
5	172/367 (46.9)	194/352 (55.1)	1.42 (1.05 to 1.92)	0.169
Cycle programming				
No	82/173 (47.4)	68/173 (39.3)	0.75 (0.48 to 1.15)	
Yes	45/124 (36.3)	60/118 (50.8)	1.83 (1.08 to 3.09)	0.010
History of miscarriage				
0–2	209/519 (40.3)	220/517 (42.6)	1.11 (0.86 to 1.43)	
≥ 3	4/6 (66.7)	2/6 (33.3)	0.30 (0.03 to 3.52)	0.299

a Day 6 excluded because there were only two women in the TAU group.

The effect of in vitro fertilisation delay after endometrial scratch procedure

Figure 11 shows a negligible predicted change in the probability of achieving a live birth with increasing delay in starting IVF after ES procedure. This is adjusted for other covariates: age, BMI, site, treatment protocol followed, duration of infertility, history of pregnancy and current smoking status. This picture is similar even without accounting for these covariates. It should be noted that uncertainty increases with delay in IVF after ES procedure as the number of participants decreases.

Effect of endometrial scratch on safety outcomes

We present expected and unexpected AEs and SAEs reported in women as well as safety events reported in born babies based on a safety analysis set defined in *Chapter 3, Safety analyses*.
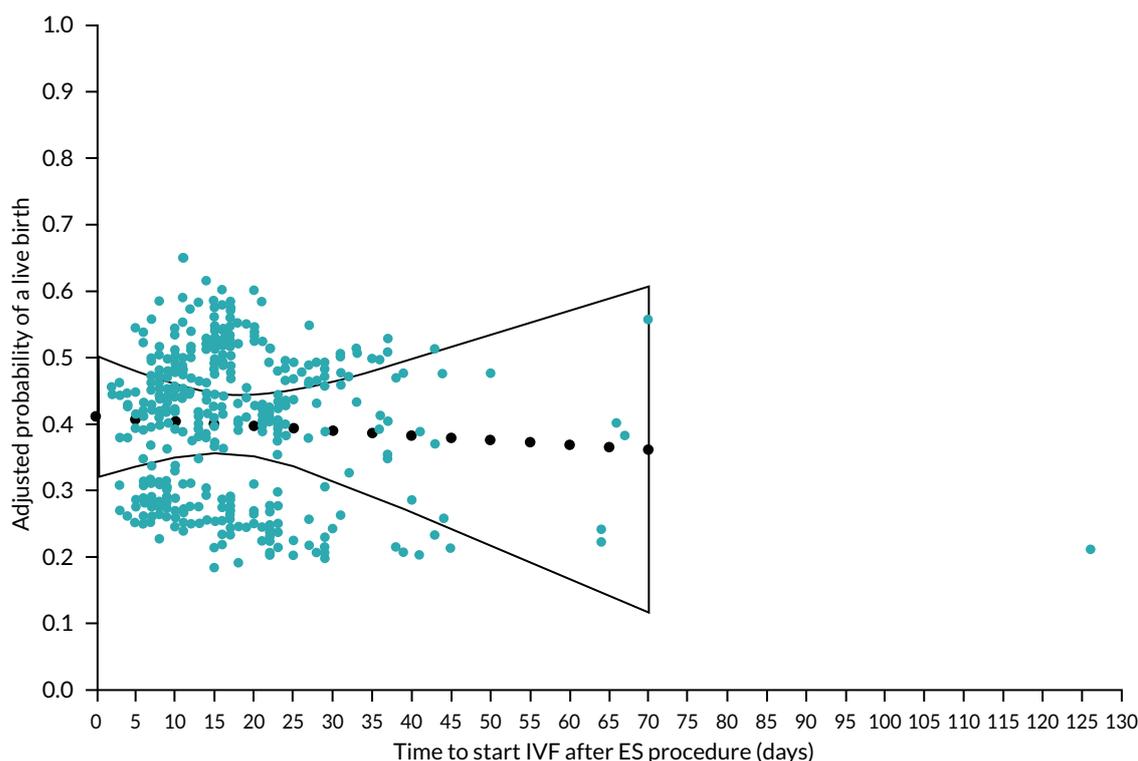
FIGURE 11 Adjusted association between delay in IVF after ES procedure and chances of a live birth.

Expected adverse events in participating women

Expected AEs that were reported at any point during the trial among women who received the study treatments are summarised in *Table 11*. In general, the prevalence of these expected AEs was very similar between the treatment groups. For instance, the proportion of women who reported at least one expected AE was 174 (32.4%) in the TAU group and 144 (31.4%) in the ES group. The most common expected AEs were abdominal pain, nausea and vaginal bleeding; however, these were only reported in < 14% of women.

Unexpected adverse events and serious adverse events in participating women

The incidences of unexpected AEs, expected SAEs and unexpected SAEs reported after the delivery of the interventions were similar between treatment groups (*Table 12*). For example, 16.9% (91/537) of participants in the TAU group and 19.0% (87/458) of participants in the ES group reported at least one unexpected AE. The total unexpected AEs including repeated events per participant were 120 in the TAU group and 114 in the ES group, and women were followed up for a total of 255.7 person-years in the TAU group and 248.6 person-years in the ES group. This translates to an incidence of 47 per 100 person-years in the TAU group and 48 per 100 person-years in the ES group. Thus, the incidences of unexpected AEs are 1.02 times (95% CI; 0.76 to 1.38) more in the ES group than in the TAU group.

In total, there were 9.3% (50/537) and 8.1% (37/458) SAEs with an IR of 22 per 100 person-years and 19 per person-years in the TAU group and ES group, respectively. This equates to an IR of 0.88 times (95% CI 0.58 to 1.34) in the ES group compared with the TAU group.

There were no deaths reported during the entire trial. Only 1 out of 458 (0.2%) SAEs in the ES group was viewed as possibly related to the ES procedure. Most of the SAEs resulted in inpatient or prolonged hospitalisation and were rated as isolated and of moderate intensity; however, these incidences were very similar between treatment groups. Only one vaginal bleeding incident was reported in the ES group, but this event did not occur within 48 hours of the ES procedure.

TABLE 11 Expected AEs during the entire trial

AEs	TAU group (N = 537), n (%)	ES group (N = 458), n (%)
At least one expected AE	174 (32.4)	144 (31.4)
At least one expected AE category		
Abdominal pain	44 (8.2)	37 (8.1)
Anaemia	22 (4.1)	14 (3.1)
Back pain/sciatica	4 (0.7)	5 (1.1)
Bloating	8 (1.5)	4 (0.9)
Cholestasis	2 (0.4)	4 (0.9)
Clicky hip	1 (0.2)	2 (0.4)
Cold/flu	1 (0.2)	1 (0.2)
Conjunctivitis	1 (0.2)	0 (0.0)
Constipation	10 (1.9)	10 (2.2)
Cough	6 (1.1)	2 (0.4)
Diarrhoea	6 (1.1)	2 (0.4)
Dizziness/feeling faint	12 (2.2)	8 (1.7)
Epistaxis	4 (0.7)	2 (0.4)
Fall	8 (1.5)	5 (1.1)
Fatigue/tiredness	3 (0.6)	5 (1.1)
Gestational diabetes	7 (1.3)	10 (2.2)
Group B Streptococcus infection	4 (0.7)	1 (0.2)
Headache/migraine	35 (6.5)	24 (5.2)
Hyperemesis	1 (0.2)	1 (0.2)
Hypertension	4 (0.7)	4 (0.9)
Itchy skin	5 (0.9)	3 (0.7)
Mild OHSS	19 (3.5)	18 (3.9)
Nausea	73 (13.6)	50 (10.9)
Vomiting	39 (7.3)	25 (5.5)
Palpitations	13 (2.4)	6 (1.3)
Pelvic girdle pain or symphysis pubis dysfunction or hip pain/pelvis	11 (2.0)	6 (1.3)
Pre-eclampsia	4 (0.7)	4 (0.9)
Proteinuria	1 (0.2)	1 (0.2)
Rash	2 (0.4)	2 (0.4)
Reduced fetal movement	20 (3.7)	13 (2.8)
Reflux/heartburn/indigestion	2 (0.4)	2 (0.4)
Urinary tract infection	6 (1.1)	4 (0.9)
Vaginal bleeding	50 (9.3)	33 (7.2)
Vaginal discharge	9 (1.7)	0 (0.0)
Vaginal infection	5 (0.9)	3 (0.7)
Viral infection	3 (0.6)	2 (0.4)

OHSS, ovarian hyperstimulation syndrome.

TABLE 12 Unexpected AEs and SAEs after the delivery of interventions

AE classification	TAU group (N = 537), n (%)	n/TPy (IR)	ES group (N = 458), n (%)	n/TPy (IR)	IRR (95% CI)
Unexpected AEs	91 (16.9)	120/255.7 (0.47)	87 (19.0)	114/248.6 (0.48)	1.02 (0.76 to 1.38)
SAEs	50 (9.3)	58/255.7 (0.22)	37 (8.1)	49/248.6 (0.19)	0.88 (0.58 to 1.34)
Unexpected SAEs	25 (4.7)	27/255.7 (0.10)	22 (4.8)	26/248.6 (0.10)	1.00 (0.56 to 1.78)
Expected SAEs	28 (5.2)	31/255.7 (0.12)	20 (4.4)	23/248.6 (0.09)	0.77 (0.43 to 1.37)
Unexpected AE category					
Pain	20 (3.7)	21	13 (2.8)	17	
OHSS	5 (0.9)	5	7 (1.5)	7	
GI-related issues	10 (1.9)	12	10 (2.2)	11	
Infection/inflammation	18 (3.4)	18	21 (4.6)	23	
Bleeding/blood-related events	34 (6.3)	38	21 (4.6)	22	
Urinary-related issues	6 (1.1)	6	2 (0.4)	2	
Cardiac-related issues	6 (1.1)	6	13 (2.8)	14	
Vaginal bleeding-related issues[a]	0 (0.0)	0	1 (0.2)	1	
Pregnancy-specific issues	9 (1.7)	10	8 (1.7)	9	
Placenta-related issues	7 (1.3)	8	4 (0.9)	4	
Skin-related issues	5 (0.9)	5	0 (0.0)	0	
Events in baby/fetus	5 (0.9)	5	4 (0.9)	4	
Mental health	1 (0.2)	2	1 (0.2)	1	
Neuro-related issues	1 (0.2)	1	1 (0.2)	1	
Uterine abnormality	0 (0.0)	0	4 (0.9)	4	
Other	13 (2.4)	14	21 (4.6)	21	
Seriousness of SAE					
Life-threatening	5 (0.9)	5	2 (0.4)	2	
Inpatient hospitalisation	36 (6.7)	40	29 (6.3)	37	
Prolongs hospitalisation	12 (2.2)	12	8 (1.7)	10	
Persistent or significant disability/incapacity	1 (0.2)	1	0 (0.0)	0	
Frequency of SAE					
Isolated	43 (8.0)	48	34 (7.4)	39	
Intermittent	4 (0.7)	4	5 (1.1)	8	
Continuous	2 (0.4)	2	1 (0.2)	1	
Missing[b]	0 (0.0)	0	1 (0.2)	1	
Unknown	4 (0.7)	4	0 (0.0)	0	
Intensity of SAE					
Mild	18 (3.4)	18	7 (1.5)	7	
Moderate	29 (5.4)	32	23 (5.0)	31	
Severe	8 (1.5)	8	6 (1.3)	10	
Missing	0 (0.0)	0	1 (0.2)	1	

TABLE 12 Unexpected AEs and SAEs after the delivery of interventions (*continued*)

| AE classification | TAU group | | ES group | | |
	(N = 537), n (%)	n/TPy (IR)	(N = 458), n (%)	n/TPy (IR)	IRR (95% CI)
Outcome of SAE					
Recovered	45 (8.4)	50	32 (7.0)	40	
Improved	7 (1.3)	7	5 (1.1)	7	
Unchanged	1 (0.2)	1	1 (0.2)	1	
Missing[b]	0 (0.0)	0	1 (0.2)	1	
Relationship to ES					
Possible	N/A	N/A	1 (0.2)	1	
Unlikely	N/A	N/A	10 (2.2)	11	
Unrelated	N/A	N/A	25 (5.5)	35	
Missing[b]	N/A	N/A	1 (0.2)	1	
Not assessable	N/A	N/A	1 (0.2)	1	

GI, gastrointestinal; N/A, not applicable; n/TPy (IR), number of all repeated events/total follow-up time in years (IR per person-year); OHSS, ovarian hyperstimulation syndrome.
a Did not occur within 24 hours of the ES procedure.
b The frequency, intensity, outcome and relationship to ES procedure of one unexpected SAE that resulted in inpatient hospitalisation could not be ascertained (SAE was identified from the final trial questionnaire and the site could not contact the participant to ascertain the details).

For sensitivity analysis, *Appendix 19* is a replication of *Table 12* that includes all unexpected AEs and AEs reported at any point during the trial (before and after receiving the interventions), as there is a possibility that down-regulation could have happened before the interventions in few unknown cases. In summary, there were very few events reported before the interventions so *Table 12* and *Appendix 19* are very similar.

Unexpected adverse events and serious adverse events in women between endometrial scratch and in vitro fertilisation

Only 11 (2.4%) women reported unexpected AEs that occurred after the delivery of ES but before receiving IVF procedure (*Table 13*). In addition, there was only one unexpected SAE, which was unlikely to be related to the ES procedure.

Serious adverse events in born babies

We present SAEs reported in born babies among women who had successful embryo implantation (with a positive pregnancy test) based on treatment as received. Successful implantation was reported in 270 and 226 women who received IVF only and ES procedure followed by IVF, respectively (*Table 14*). Three (1.1%) severe congenital abnormalities were reported, all in the TAU group. There were no neonatal deaths reported during the entire trial. Four other SAEs reported in the ES group only were congenital brain abnormality at 12 weeks (termination recommended as condition not compatible with life), lumbosacral myelomeningocele with an Arnold–Chiari malformation, hypoplastic left heart syndrome and full pulmonary cardiac resuscitation (baby recovered).

In summary, the rates of live births with low birthweight, very low birthweight and small for gestational age were higher in the TAU group than in the ES group, which appears to favour the ES procedure (see *Table 14*). For example, 14.1% (38/270) of mothers in the TAU group and 5.8% (13/226) in the ES group gave birth to at least one low birthweight baby resulting in an absolute reduction in proportions

TABLE 13 Unexpected AEs between ES and IVF (ES group only)

AE classification	ES group (N = 458)	
	n (%)	Total events
Unexpected AE	11 (2.4)	13
SAE	1 (0.2)	1
Unexpected SAE	1 (0.2)	1
Expected SAE	0 (0.0)	0
Unexpected AE category		
Pain	2 (0.4)	2
GI-related issues	2 (0.4)	2
Infection/inflammation	1 (0.2)	1
Cardiac-related issues	1 (0.2)	1
Mental health	1 (0.2)	1
Other	9 (2.0)	9
Seriousness of SAE		
Inpatient hospitalisation	1 (0.2)	1
Frequency of SAE		
Isolated	1 (0.2)	1
Intensity of SAE		
Moderate	1 (0.2)	1
Outcome of SAE		
Improved	1 (0.2)	1
Relationship to ES		
Unlikely	1 (0.2)	1
GI, gastrointestinal.		

TABLE 14 Serious adverse events in born babies

SAE classification	TAU group (N = 270)		ES group (N = 226)		AD (95% CI)
	n (%)	Total events	n (%)	Total events	
Severe congenital abnormality	3 (1.1)	3	0 (0.0)	0	−1.1% (−2.4% to 0.1%)
Neonatal death	0 (0.0)	0	0 (0.0)	0	N/A
Other[a]	0 (0.0)	0	4 (1.8)	4	1.8% (0.1% to 3.5%)
Preterm delivery[b]	22 (8.1)	30	12 (5.3)	16	−2.8% (−7.2% to 1.5%)
Very preterm delivery[b]	1 (0.4)	1	0 (0.0)	0	−0.4% (−1.1% to 0.4%)
Low birthweight[b]	38 (14.1)	44	13 (5.8)	17	−8.3% (−13.5% to −3.2%)
Very low birthweight[b]	14 (5.2)	17	3 (1.3)	3	−3.9% (−6.9% to −0.8%)
Small for gestational age[b]	19 (7.0)	19	6 (2.7)	6	−4.4% (−8.1% to −0.7%)
Large for gestational age[b]	19 (7.0)	19	13 (5.8)	13	−1.3% (−5.6% to 3.0%)

a Congenital brain abnormality at 12 weeks, lumbosacral myelomeningocele with an Arnold–Chiari malformation, hypoplastic left heart syndrome, and full pulmonary cardiac resuscitation.
b These terms are defined in *Chapter 2, Safety outcomes, In the fetus/neonate.*

of 8.3% (95% CI 3.2% to 13.5%) in favour of the ES procedure. These results should be interpreted with caution because of the potential case-mix, as these results were not adjusted for potential prognostic factors relating to mothers. For example, there could be some differences between treatment groups concerning the characteristics of women who gave birth to babies with low birthweight, very low birthweight, or babies who were small for their gestational age that may influence these outcomes. In addition, low birthweight, very low birthweight, and small for their gestational age are highly correlated.

Statistical analyses not performed and rationale

There were some statistical analyses that were not performed as planned in the SAP with justification. First, descriptive assessment of the differences in characteristics between completers and non-completers was not done since the number of non-completers (those who were lost to follow-up with missing live birth outcome) was very small. Second, the incidences of multiple births per single mother were very small and, as a result, related exploratory analysis accounting for multiple births was not performed [e.g. the analysis of the primary outcome (live birth) accounting for multiple births per single mother and the association between the number of eggs transferred and the number of live births].

Conclusions

The ES procedure was generally well tolerated and safe; however, we did not find sufficient evidence to support the clinical benefits of the ES procedure in increasing the chances of achieving pregnancy and a live birth significantly. In addition, the rates of all secondary clinical outcomes were similar between the ES and TAU treatment groups.

Chapter 8 Economic evaluation results

This chapter reports results relating to the cost-effectiveness analysis described in *Chapter 4*. Results are reported in line with Consolidated Health Economic Evaluation Reporting Standards (CHEERS) guidance.[67]

Resource use

Visits to accident and emergency

Between 3% and 6% of the participants accessed accident and emergency (A&E) during the study with slightly more accessing A&E at 3 months and 6 weeks post partum in the ES group (*Table 15*). Those who accessed A&E accessed it between one and five times. *Table 15* also shows the number of cases who went on to be admitted to hospital. *Appendix 20* provides a list of reasons for admission in addition to the unit costs, reasons for admission included infections, bleeding, pain and ovarian hyperstimulation syndrome (OHSS).

TABLE 15 Frequency of A&E visits and A&E visits followed by hospital admissions resource use

Resource	Baseline		3 months		6 weeks post partum	
	TAU group	ES group	TAU group	ES group	TAU group	ES group
A&E, *n* (%)	17 (3.2)	17 (3.3)	19 (3.6)	32 (6.1)	20 (3.8)	24 (4.6)
A&E and admitted to hospital	3	3	5		7	9
Planned cases, *n* (%)	168 (32.0)	196 (37.5)	65 (43.9)	87 (50.9)	116 (80.0)	130 (81.7)
Day cases	168	192	64	86	58	59
Overnight stay	0	4	1	1	58	71
GP, *n* (%)	184 (35.0)	205 (39.3)	78 (52.7)	104 (60.8)	107 (73.8)	122 (76.7)
Practice nurse, *n* (%)	78 (14.9)	105 (20.1)	29 (19.6)	34 (19.9)	44 (30.3)	44 (27.7)
Walk-in centre, *n* (%)	22 (4.2)	27 (5.2)	6 (4.1)	7 (4.1)	20 (13.8)	12 (7.5)
Dentist, *n* (%)	111 (21.1)	134 (25.7)	36 (24.3)	53 (31.0)	40 (27.6)	47 (29.6)
Pharmacist, *n* (%)	110 (21.0)	114 (21.8)	44 (29.7)	51 (29.8)	38 (26.2)	48 (30.2)
Physiotherapist, *n* (%)	18 (3.4)	29 (5.6)	2 (1.4)	7 (4.1)	14 (9.7)	18 (11.3)
Occupational therapist, *n* (%)	5 (1.0)	3 (0.6)	0 (0.0)	2 (1.2)	0 (0.0)	0 (0.0)
Dietitian, *n* (%)	1 (0.2)	3 (0.6)	1 (0.7)	0 (0.0)	2 (1.4)	0 (0.0)
Social worker, *n* (%)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.6)	0 (0.0)	1 (0.6)
Health visitor, *n* (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	122 (84.1)	136 (85.5)
Community psychiatrist/mental health nurse, *n* (%)	3 (0.6)	0 (0.0)	0 (0.0)	1 (0.6)	0 (0.0)	4 (2.5)
Midwife, *n* (%)	0 (0.0)	0 (0.0)	116 (78.4)	143 (83.6)	114 (78.6)	135 (84.9)
Planned services, *n* (%)	111 (21.1)	102 (19.5)	52 (9.9)	31 (6.0)	31 (21.4)	31 (19.5)
NHS ultrasound, *n* (%)	0 (0.0)	0 (0.0)	145 (98.0)	168 (98.2)	72 (49.7)	93 (58.5)
Private ultrasound, *n* (%)	0 (0.0)	0 (0.0)	50 (33.8)	50 (29.2)	11 (7.6)	14 (8.8)

The average cost of A&E visits was £26.63 in the TAU group and £34.21 in the ES group and the average cost of overnight stays after an A&E admission was also higher in the ES group (TAU mean £93.36; ES mean £130.94).

Ambulance services

Twenty-five participants used the ambulance service during the course of the study: 10 (1.9%) participants in the TAU arm and 15 (2.9%) participants in the ES arm. The cost of using the service was similar for both groups (TAU mean £7.89; ES mean £7.35).

Planned hospital visits

Planned hospital visits were higher post partum than at other time points and were slightly higher in the ES group than with TAU (see *Table 15*). Most planned visits were day cases, except at 6 weeks post partum when over half were overnight stays. *Appendices 21* and *22* provide a list of reasons why participants were listed for day-case visits and overnight stays, respectively, and the unit costs involved. Participants' most frequent reasons for day visits were IVF treatment, ultrasound scans, consultant appointments, blood tests and laparoscopy. Participants' most frequent reasons for overnight visits were for giving birth: labour, induction and caesarean sections.

The average cost of planned admissions was slightly higher in the TAU group (mean TAU = £334.71; mean ES = £295.74).

Services

Many participants accessed GPs, dentists, pharmacists and practice nurses during the study, with around 80% accessing midwives at 3 months and at 6 weeks post partum and > 80% accessing health visitors post partum (see *Table 15*). Those in the ES group were more likely to access services; this was true at baseline as well as at 3 months and 6 weeks post partum. The mean cost of services was higher for the ES group than for the TAU group (mean TAU = £375.88; mean ES = £422.84).

Other services

Around 20% of participants accessed other services at baseline and 6 weeks post partum with a lower percentage (between 6% and 10%) accessing services at 3 months (see *Table 15*). The types of services ranged from scans to breastfeeding clinics, assistive conception units and acupuncture. *Appendix 23* provides a list of reasons for admission in addition to the unit costs. The mean costs were slightly higher for the TAU group than for the ES group (TAU mean = £1178.62; ES mean = £930.83).

Ultrasound

Appendix 24 shows that the majority of participants had an ultrasound within 3 months of starting the study, which is as expected. A higher proportion in the ES group had scans later in their pregnancy. About one-third of participants had a scan privately. The average costs to the NHS were slightly higher for the ES group than for the TAU group (TAU mean = £597.95; ES mean = £621.61).

Treatment costs

The cost of IVF per cycle is £3218 and is incurred by all participants in the study undergoing IVF treatment; an additional study cost of £10 per participant is also included for the provision of ovulation kits. Those who underwent ES incurred an additional cost of £160.

Overall costs

The overall cost for TAU was on average £316 lower than for ES (TAU mean £6503; ES mean = £6819). *Table 16* presents a summary of costs per treatment group.

TABLE 16 Summary of mean treatment cost

| Treatment | Cost, mean (95% bias corrected bootstrapped CI) | |
	TAU group	ES group
A&E	£26.24 (£17.23 to £37.94)	£34.91 (£25.12 to £48.02)
A&E followed by hospital stay	£105.51 (£44.31 to £184.87)	£112.36 (£52.76 to £192.00)
Services	£447.20 (£374.06 to £522.50)	£565.45 (£479.75 to £661.47)
Other services	£436.25 (£183.59 to £835.79)	£265.24 (£104.40 to £538.70)
Ambulance	£10.28 (£3.67 to £22.03)	£8.92 (£3.43 to £16.27)
Ultrasound	£496.40 (£410.54 to £598.95)	£590.07 (£497.50 to £698.53)
Day cases	£155.45 (£84.64 to £236.13)	£114.05 (£63.10 to £201.54)
Overnight stay	£81.57 (£29.07 to £159.44)	£115.78 (£32.52 to £254.35)
IVF	£3027.98 (£2954.43 to £3083.15)	£3057.72 (£2989.90 to £3107.03)
Scratch	N/A	£138.59 (£133.33 to £142.84)
Ovulation kit	£10.00 (£10.00 to £10.00)	£10.00 (£10.00 to £10.00)
Delivery	£1706.14 (£1520.85 to £1881.09)	£1809.45 (£1621.93 to £1998.62)
Costs at baseline	£207.94 (£164.81 to £258.70)	£215.30 (£171.44 to £271.64)
Total costs	£6503.03 (£5989.60 to £7106.95)	£6819.38 (£6352.80 to £7353.40)

N/A, not applicable.

Cost-effectiveness analysis

Table 17 presents the main results for the cost-effectiveness analysis. The mean cost of fertility was £316 more in the ES group than in the TAU group (mean difference = £316; 95% CI –£468 to £1078). The mean incremental risk difference in live births was 16 (95% CI –44 to 77) per 1000 population and was identical for the worst-case, best-case and complete-cases analysis. Therefore, the ICERs were the same in each case at £11.90 (95% CI –£134.08 to £127.24) per successful live birth, meaning that, in order to achieve one extra live birth in 1000 women, £11,900 would need to be spent. *Figure 12* presents the cost-effectiveness plane for the worst-case scenario. The uncertainty in the ICER covers all four quadrants of the cost-effectiveness plane with the majority of points (64%) being in the north-east quadrant where ES is more effective but more costly compared with TAU. The cost-effectiveness acceptability curve (*Figure 13*) presents the probability of ES being cost-effective compared with TAU and different willingness-to-pay thresholds. The probability of ES being cost-effective at a willingness-to-pay threshold of £0 is 20%, and is 70% from a willingness-to-pay threshold of £1000.

TABLE 17 Incremental cost-effectiveness analysis with 95% bias corrected bootstrap CIs

Scenario	TAU group	ES group	Incremental difference (ES – TAU)	ICER
Total costs	£6503 (£5990 to £7107)	£6819.38 (£6352.80 to £7353.40)	£316.35 (–£468 to £1078)	
LB WC	373 (331 to 413)	388 (348 to 433)	15 (–42 to 75)	£11.90 (–£134.08 to £127.24
LB BC	373 (331 to 413)	388 (348 to 433)	15 (–42 to 75)	£11.90 (–£134.08 to £127.24
LB CC	373 (331 to 413)	388 (348 to 433)	15 (–42 to 75)	£11.90 (–£134.08 to £127.24

BC, best case; CC, complete case; LB, live birth; WC, worst case.
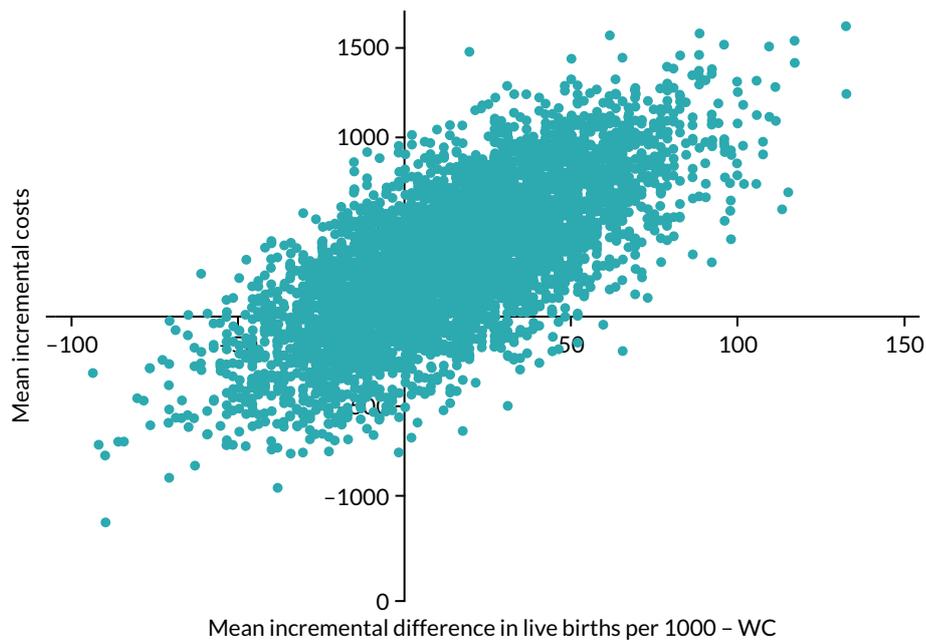
FIGURE 12 Incremental cost-effectiveness plane per incremental difference in successful live births between ES and TAU (5000 bootstrap replicates). WC, worst case.
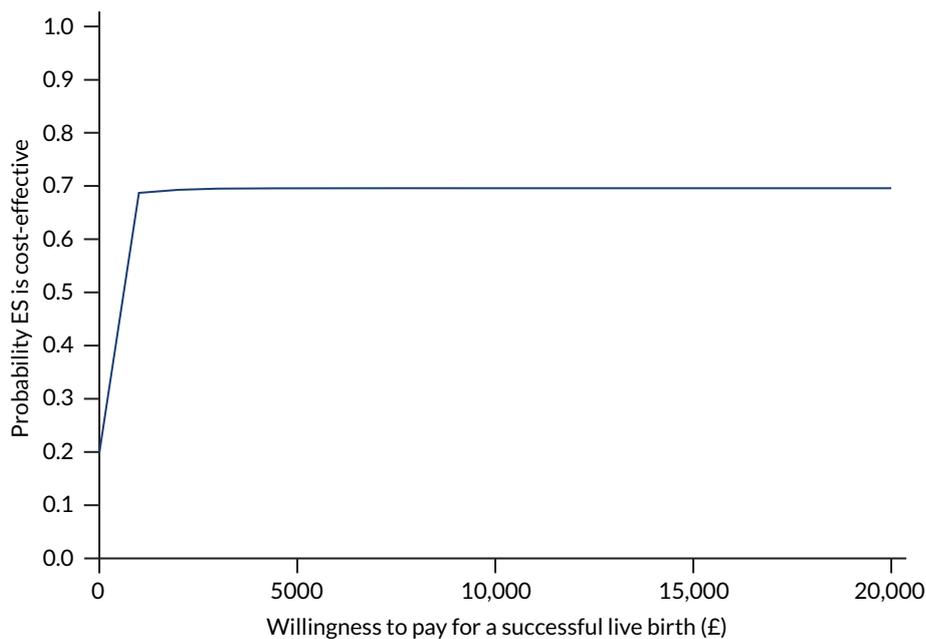
FIGURE 13 Probability that ES is cost-effective relative to TAU.

Secondary outcomes

There were increases in live births and clinical pregnancies and reductions in biochemical pregnancies, ectopic pregnancies, miscarriages, multiple births and preterm births in the ES group compared with the TAU group (*Table 18*). The ICER ranged from –£57.50 incremental cost per ectopic pregnancy avoided to £11.90 incremental cost per successful live birth with uncertainty in the cost-effectiveness for all outcomes.

TABLE 18 Summary of cost-effectiveness analysis with 95% bias corrected bootstrap CIs for secondary outcomes

Scenario	TAU group	ES group	Incremental difference (ES–TAU)	ICER
Total costs	£6503 (£5990 to £7107)	£6819.38 (£6352.80 to £7353.40)	£316.35 (–£468 to £1078)	
LB WC	373 (331 to 413)	388 (348 to 433)	15 (–42 to 75)	£11.90 (–£134.08 to £127.24)
Clinical pregnancies	406 (365 to 448)	426 (384 to 468)	21 (–38 to 81)	£11.29 (–£90.91 to £125.70)
Biochemical pregnancies	38 (23 to 55)	21 (10 to 35)	–17 (–38 to 4)	–£17.09 (–£173.98 to £105.55)
Ectopic pregnancies	4 (0 to 10)	2 (0 to 8)	–2 (–8 to 4)	–£57.50 (–£6466 to £5690)
Miscarriages	124 (96 to 153)	84 (62 to 109)	–39 (–77 to –3)	–£7.99 (–£54.60 to £23.57)
Stillbirths	2 (0 to 8)	2 (0 to 8)	0 (–6 to 6)	£6.45 (–£8458 to £7922)
Multiple births	21 (10 to 34)	11 (4 to 23)	–9 (–25 to 5)	–£22.78 (–£504.08 to £383.99)
Preterm births	38 (23 to 55)	27 (15 to 42)	–11 (–33 to 10)	–£15.60 (–£417.45 to £373.08)

LB, live birth; WC, worst case.

Subgroup analysis

Costs increase by day of embryo transfer in both groups and are less expensive in the ES group up to day 5. At day 6, they are more expensive for the ES group [*Table 19*; $F_{4,911} = 3.21$ ($p = 0.013$)]. There were no other significant differences in costs for fertilisation method, type of protocol, embryo transfer, the nature of the embryo, history of miscarriage or cycle programming (yes/no) between the two groups. Given there was a difference in costs by day of embryo transfer, this was adjusted for in addition to baseline costs in sensitivity analysis (see *Sensitivity analysis*).

TABLE 19 Summary of costs by subgroup with 95% bias corrected bootstrap CIs for TAU and ES

Subgroup	TAU group	ES group	Statistical test ANOVA comparing difference in costs between TAU and ES (*p*-value)
Day of embryo transfer			
Day 2	N = 27	N = 29	$F_{4,911} = 3.21$ ($p = 0.013$)
	£6041 (£4175 to £7908)	£5026 (£3715 to £6337)	
Day 3	N = 57	N = 68	
	£6330 (£4774 to £7886)	£5670 (£4551 to £6790)	
Day 4	N = 11	N = 6	
	£7043 (£4712 to £9373)	£6912 (£2282 to £11,542)	
Day 5	N = 367	N = 352	
	£7172 (£6479 to £7865)	£8036 (£7405 to £8667)	
Day 6	N = 2	N = 0	
	£3228 (N/A, N/A)	N/A	

TABLE 19 Summary of costs by subgroup with 95% bias corrected bootstrap CIs for TAU and ES (*continued*)

Subgroup	TAU group	ES group	Statistical test ANOVA comparing difference in costs between TAU and ES (*p*-value)
Fertility method			
IVF	264	241	$F_{2,935} = 0.49$ ($p = 0.612$)
	£6478 (£5916 to £7040)	£7616 (£6703 to £8529)	
ICSI	205	209	
	£7386 (£6258 to £8515)	£7245 (£6520 to £7970)	
IVF and ICSI	7	15	
	£7745 (£2398 to £13,091)	£5706 (£3858 to £7555)	
Type of protocol			
Long	197	291	$F_{2,985} = 0.21$ ($p = 0.808$)
	£6690 (£5890 to £7490)	£7100 (£6339 to £7860)	
Antagonistic	297	205	
	£6797 (£5997 to £7597)7	£7164 (£6480 to £7847)	
Embryo transfer			
Single	374	381	$t_{1,917} = 0.24$ ($p = 0.811$)
	£6856 (£6333 to £7379)	£7565 (£6961 to £8169)	
Double	90	74	
	£7508 (£5360 to £9656)	£7016 (£5798 to £8233)	
Nature of embryo			
Frozen	425	416	$t_{1,917} = -0.12$ ($p = 0.904$)
	£6968 (£6353 to £7583)	£7485 (£6917 to £8053)	
Fresh	39	39	
	£7143 (£5151 to £9136)	£7377 (£5913 to £8841)	
History of miscarriages			
0–2	519	517	$t_{1,1046} = -0.02$ ($p = 0.982$)
	£6486 (£5924 to £7048)	£6836 (£5970 to £7701)	
≥ 3	6	6	
	£7979 (£3567 to £12,391)	£5410 (£1449 to £9372)	
Cycle programme			
Yes	131	126	$t_{1,619} = -0.40$ ($p = 0.691$)
	£6097 (£4666 to £7527)	£7072 (£6170 to £7973)	
No	182	182	
	£6886 (£6010 to £7762)	£6758 (£5814 to £7702)	

ANOVA, analysis of variance.

Sensitivity analysis

Wider than NHS prospective (including travel costs and loss of earnings by participants and their companions)

Three-hundred and fourteen participants responded to the patient resource use questionnaire: 146 (27.8%) in the TAU group and 168 (32.1%) in the ES group. *Table 20* presents the mode of transport and the median and IQR for total travel time in minutes, to and from appointments and the cost for each mode of transport (note that participants could report more than one mode, e.g. train and taxi).

The majority of participants travelled by car to the appointment: 93% of the TAU and 90% of the ES participants. The median distance from the appointment for both groups was 10 miles. The median time taken to travel to and from appointments was slightly higher for the TAU group (median = 80; IQR 50–120 minutes) than for the ES group (median = 70; IQR 80–100). However, overall costs of travel were similar for both groups (TAU mean = £20.42, 95% CI £17.34 to £25.37; ES mean = £19.00, 95% CI £16.87 to £23.81). Two participants were reimbursed for time at appointments; both were in the TAU group.

The majority of participants were accompanied to their appointment and most participants and their companions were in paid work with a few who would have been doing household activities, been in full-time or part-time education, on sick leave or unemployed (*Table 21*). Between 16% and 30% of participants and their companions reported lost earnings, with companions more likely to lose earnings than participants.

Appendix 25 summarises the list of occupations for participants and their companions.

The median number of hours lost was 3 (IQR 0–9) hours in the TAU group and 2 (IQR 0–5) hours in the ES group. Time lost was estimated regardless of whether or not someone was in paid employment, as a person's time is valuable regardless of what they would have been occupied with. *Table 22* presents the cost of lost time for participants, their companions and the combined cost of both, in addition to

TABLE 20 Summary of modes of transport, cost and time spent travelling to fertility appointments

Mode	TAU group (N = 146)			ES group (N = 168)		
	n (%)	Time in minutes, median (IQR)	Cost, mean (95% bootstrap CI)	n (%)	Time in minutes, median (IQR)	Cost, mean (95% bootstrap CI)
Car	136 (93%)	80 (50–120)	£20.03 (£16.27 to £24.98)	152 (90%)	70 (4–1000)	£17.34 (£15.39 to 19.85)
Bus	2 (1.4%)	140 (N/A)	£9.10 (£8.60 to £9.60)	9 (5.4%)	60 (40–120)	£8.07 (£3.61 to £23.58)
Train	7 (4.8%)	60 (40–90)	£13.37 (£4.60 to £42.13)	8 (4.8%)	90 (4–110)	£15.08 (£4.01 to 26.38)
Taxi	5 (3.4%)	40 (30–60)	£10.76 (£4.00 to £17.93)	3 (1.8%)	150 (30–240)	£47.03 (£26 to £120)
Walk	3 (2.1%)	30 (30–60)	N/A	4 (2.4%)	15 (10–80)	N/A
Other	1	90 (N/A)	£4.50 (N/A)	0 (0.0%)	N/A	N/A
Overall	146	80 (50–120)	£20.42 (£17.34 to £25.37)	168	70 (40–100)	£19.00 (£16.87 to £23.81)
Distance from appointment (miles)	137 (94%)	10 (5–21)		164 (98%)	10 (6–20)	

N/A, not applicable.

TABLE 21 What you would have been doing if not attending appointment

	TAU group (N = 146), n (%)		ES group (N = 168), n (%)	
Activity	Participant	Companion	Participant	Companion
Accompanied (Yes)		138 (94.5%)		159 (94.6%)
Paid work	137 (93.8)	134 (97.1)	155 (92.3)	154 (96.9)
Household activities	2 (1.4)	1 (0.7)	6 (3.6)	2 (1.3)
Education (full time or part time)	0 (0.0)	0 (0.0)	1 (0.6)	2 (1.3)
Sick leave	2 (1.4)	0 (0.0)	2 (1.2)	1 (0.6)
Unemployed	0 (0.0)	0 (0.0)	2 (1.2)	0 (0.0)
Lost earnings (Yes)	25 (17.1)	41 (29.7)	27 (16.1)	33 (20.7)

TABLE 22 Mean cost of time lost for attending fertility appointment, cost of travel and overall costs incurred by participants and their companions

Cost	TAU group (N = 146)	ES group (N = 168)
Participant cost of lost time	£174.27 (£124.27 to £244.00)	£173.73 (£105.63 to £382.08)
Companion cost of lost time	£120.51 (£92.00 to £168.77)	£83.38 (£63.44 to £110.85)
Overall cost of lost time	£294.79 (£223.64 to £385.07)	£257.11 (£179.28 to £459.64)
Travel cost	£20.42 (£17.34 to £25.37)	£19.00 (£16.87 to £23.81)
Total non-NHS cost	£315.21 (£242.81 to £406.02)	£276.11 (£198.19 to £476.79)

travel costs and overall non-NHS costs for one appointment. On average, the TAU group incurred £38 more in lost time and £40 more overall than did the ES group.

The cost incurred by participants per visit was applied to seven IVF fertility visits for those receiving IVF treatment, plus one additional visit for delivery for those with a positive pregnancy outcome. If a participant indicated other hospital visits on the resource use questionnaire, the costs for these visits were also included. The mean cost incurred by participants and their companions over the study period was slightly higher for the TAU group than the ES group (mean costs to participants: TAU = £2422, 95% CI £2223 to £2729; ES = £2145, 95% CI £1924 to £2656) (*Table 23*). Including participant costs resulted in a mean difference in costs between the ES and TAU groups of £40, with an ICER per successful live birth of £2.67 (95% CI £120 to £142).

TABLE 23 Incremental cost-effectiveness analysis with 95% bootstrap CIs for main analysis and after including costs incurred by participants and their companions

	TAU group	ES group	Incremental difference (ES–TAU)	ICER
Total NHS costs	£6503 (£5990 to £7107)	£6819.38 (£6352.80 to £7353.40)	£316.35 (–£468 to £1078)	
LB WC	373 (331 to 413)	388 (348 to 433)	15 (–42 to 75)	£11.90 (–£134.08 to £127.24)
Total costs (including patient costs)	£8925 (£8354 to £9641)	£8965 (£8422 to £9616)	£40 (–£867 to £877)	£2.67 (–£120 to £142)

LB, live birth; WC, worst case.

Adjustment for baseline costs

Analysis of covariance was applied to adjust for baseline costs. There was no significant difference in costs between groups after adjusting for baseline resource use (mean difference £320.41 [95% CI −£492.46 to £1028.61: $z = 0.84$; $p = 0.402$] (*Table 24*). The ICER was lower than for the main analysis, although there is still uncertainty around the estimate (mean ICER = 5.38; 95% CI −£141 to £242).

After adjusting for baseline costs and day of embryo transfer, the difference in costs was slightly higher (mean £534.97, 95% CI −£325 to £1308: $z = 1.29$; $p = 0.196$), although it remained non-significant and the ICER per successful live birth was similar to that from the main analysis (mean ICER = 10.06; 95% CI −£206 to £332).

Triangulation of resource use

Table 25 shows the percentage of participants reporting IVF treatment and labour based on the resource use questionnaire. The table also reports those known to have IVF and their labour outcome collected via the CRF. Participants in the ES group were more likely to report resource use, but this resource use was not reported in 67–87% of cases. In the cost-effectiveness analysis, labour and IVF costs were taken based on the CRF; however, it is possible that other resource use, reported by participants, are also under-reported. Therefore, costs of other resources were inflated to reflect under-reporting based on rates reported for IVF (SA1) and rates reported for labour (SA2).

After inflating costs, the mean cost almost doubles under both analyses and the costs are less expensive for ES (SA1 TAU mean £15,260, ES mean £12,718; SA2 TAU mean £10,523, ES mean £9267), with uncertainty remaining in the ICER per successful live birth (*Table 26*).

TABLE 24 Incremental cost-effectiveness analysis with 95% bootstrap CIs after adjusting for baseline costs and day of embryo transfer

	TAU group	ES group	Incremental difference (ES–TAU)	ICER
Total costs	£6503 (£5990 to £7107)	£6819.38 (£6352.80 to £7353.40)	£316.35 (−£468 to £1078)	
LB WC	373 (331 to 413)	388 (348 to 433)	15 (−42 to 75)	£11.90 (−£134.08 to £127.24)
Adjusting for baseline costs			£320.41 (−£492 to £1029)	£5.38 (−£141 to £242)
Adjusting for baseline costs and embryo transfer day			£534.97 (−£325 to £1308)	£10.06 (−£206 to £332)

LB, live birth; WC, worst case.

TABLE 25 Frequency of reporting of IVF treatment and labour from the patient cost questionnaire and the CRF form

	TAU group			ES group		
Variable	Questionnaire, *n*	CRF, *n*	%	Questionnaire, *n*	CRF, *n*	%
IVF	62	494	12.6	85	497	17.1
Labour	51	213	23.9	74	223	33.2

TABLE 26 Incremental cost-effectiveness analysis with 95% bootstrap CIs after triangulation of IVF or labour costs

	TAU group	ES group	Incremental difference (ES–TAU)	ICER
Total NHS costs	£6503 (£5990 to £7107)	£6819.38 (£6352.80 to £7353.40)	£316.35 (–£468 to £1078)	
LB WC	373 (331 to 413)	388 (348 to 433)	15 (–42 to 75)	£11.90 (–£134.08 to £127.24)
Total costs (triangulation based on labour)	£15,260 (£12,890 to £19,957)	£12,718 (£11,090 to £15,186)	–£2542 (–£6664 to £1064)	–£169.47 (–£1343 to £1187)
Total costs (triangulation based on IVF)	£10,523 (£9201 to £13,039)	£9267 (£8327 to £10,576)	–£1255 (–£3517 to £750)	–£83.67 (–£690 to £619)

LB, live birth; WC, worst case.

Conclusions

The cost-effectiveness of ES demonstrates that it is more costly but more effective than TAU; however, there is uncertainty in the estimates. In the main analysis, 64% of the time ES would be more costly and more effective than TAU, with an ICER per successful live birth of £11.90 (95% CI –£134 to £127). Uncertainty in the ICER remained for all secondary outcome measures and there were no differences in the ICER after adjusting for baseline costs, day of embryo transfer and costs incurred by participants and their companions, and triangulating the reporting of resource usage. However, triangulation of resources suggested that ES could be cheaper than TAU. The triangulation assumed that the under-reporting of labour or IVF in the resource use questionnaire was the same for other items of resource use, whereas participants may have been aware that this was recorded routinely as part of the study and thus not reported this. Therefore, the amount of under-reporting adjusted for in the triangulation sensitivity analysis could be thought of as an extreme scenario and the under-reporting of other resources could be lower than this.

Chapter 9 Qualitative substudy results

In this chapter we present the results of the qualitative substudy, the methods for which are described in *Chapter 5*.

Interviews and characteristics of participants

Forty-four trial participants were approached to participate in the study, of which two actively declined to participate and 15 either did not respond or were not contactable. As a result, 27 trial participants took part.

Interviewed participants had an average duration of infertility of 3 years (range from 11 months to 8 years 6 months) and were interviewed, on average, 17.8 months after they were randomised to the trial (range from 8 to 33 months). Seven participants were interviewed who were randomised to receive ES but declined the procedure. The demographics of the trial participants who took part in this qualitative study are summarised in *Table 27*.

Seven site staff (one PI and six RNs) also took part in this qualitative study. Fifteen participants were invited (seven PIs and eight RNs); one participant actively declined because of lack of time, five participants did not reply to initial contact attempts, and attempts were made to schedule interviews for two individuals but a mutually agreed time could not be found. The characteristics of the site staff who participated in this study can be found in *Table 28*.

TABLE 27 Demographics of interviewed trial participants

ID	Time between randomisation and interview (months)	Age at randomisation	Duration of infertility	Live birth recorded?	Randomisation arm	ES received?
P1-SITE4	29	32	2 years 3 months	No	ES	Declined
P2-SITE4	16	35	3 years 6 months	Yes	ES	Yes
P3-SITE4	15	34	3 years	Yes	TAU	N/A
P4-SITE4	12	34	4 years	No	ES	Yes
P1-SITE1	20	28	4 years	Yes	TAU	N/A
P2-SITE7	21	35	3 years	Yes	ES	Yes
P3-SITE7	17	21	2 years	No	ES	Yes
P4-SITE7	15	33	1 year 2 months	Yes	TAU	N/A
P1-SITE1	29	37	1 year 6 months	No	ES	Declined
P2-SITE1	24	33	3 years 7 months	Yes	ES	Declined
P3-SITE1	22	32	2 years 5 months	No	ES	Yes
P4-SITE1	9	29	1 year 10 months	No	ES	Yes

continued

TABLE 27 Demographics of interviewed trial participants (*continued*)

ID	Time between randomisation and interview (months)	Age at randomisation	Duration of infertility	Live birth recorded?	Randomisation arm	ES received?
P5-SITE1	8	31	1 year 8 months	No	ES	Yes
P1-SITE3	19	33	3 years 6 months	No	ES	Declined
P1-SITE9	33	28	8 years 6 months	Yes	ES	Declined
P2-SITE9	14	30	1 year 5 months	Yes	ES	Declined
P1-SITE8	16	29	2 years	No	TAU	N/A
P2-SITE8	16	33	11 months	Yes	ES	Yes
P3-SITE8	11	28	3 years 5 months	No	ES	Yes
P4-SITE8	9	35	2 years 6 months	No	TAU	N/A
P1-SITE5	23	31	3 years 6 months	Yes	ES	Declined
P2-SITE5	19	32	3 years	Yes	TAU	N/A
P3-SITE5	16	24	3 years	No	ES	Yes
P4-SITE5	17	36	4 years	No	ES	Yes
P1-SITE6	25	30	6 years	Yes	TAU	N/A
P2-SITE6	15	29	2 years 2 months	Yes	ES	Yes
P3-SITE6	14	35	2 years 9 months	Yes	ES	Yes
P4-SITE6	15	32	3 years 7 months	No	TAU	N/A

ID, identification number; N/A, not applicable.

TABLE 28 Characteristics of interviewed site staff

ID	NHS/private	Site participating in E-Freeze?	Role	Sex
PI-SITE1	NHS	No	PI	Female
RN-SITE1			RN	Female
RN-SITE2	NHS	Yes	RN	Female
RN-SITE3	NHS	No	RN	Female
RN-SITE4	NHS	Yes	RN	Female
RN-SITE5	NHS	No	RN	Female
RN-SITE6	Private	Yes	RN	Female

ID, identification number.

Below, we summarise the results of the trial participant and site staff interviews under the prespecified research question set out in *Chapter 5*.

What were the reasons for participants withdrawing from the intervention?
Seven participants were interviewed who, having been randomly allocated to receive the ES, did not receive it.

Five participants declined the ES because of reasons that were out of the control of their fertility centre: illness, holidays and work commitments.

However, two participants stated reasons for declining ES that could have been prevented. For P1-SITE5, the site appeared to have forgotten to check the participant's records for a recent smear test result:

> *I, we were scheduled to start the IVF and then I got a call from, like the following cycle, I got a call from one of the nurses, not the research nurse one of the other nurses at the facility and she said 'Oh I've just checked your records and it turns out you're, you're overdue a smear test, we can't proceed with the IVF until you've had a smear test' . . .*
>
> *P1-SITE5*

For P1-SITE1, the site appeared to have forgotten to tell the participant to refrain from sexual intercourse in the menstrual cycle in which the ES was being performed. This participant recommended that discussions around abstinence should be more prominent in the recruitment process:

> *I think that, that discussion [abstinence from sexual intercourse] needs to be kind of one of the leading things.*
>
> *P1-SITE1*

One participant discussed not feeling supported when she told the fertility unit staff that she did not want to receive the procedure:

> *I just felt that the lady was quite abrupt with me, once we'd decided that I wasn't going to partake, I felt that the lady was a little bit abrupt. And didn't really understand why I'd made that decision.*
>
> *P1-SITE4*

However, other participants felt well supported:

> *They were lovely about it all, they were really nice.*
>
> *P2-SITE9*

Of those women who did receive the intervention, what were their thoughts of the procedure and the information they received in preparation for the procedure?

Preconceptions and preparedness
Participants were aware of the variability in the pain felt during the procedure, mainly from undertaking their own internet searches, and were apprehensive of this:

> *I'd done some googling and looked at some like Facebook [Facebook, Inc., Menlo Park, CA, USA] groups and stuff like that and seen what people's experiences were and that people said it was really painful so yeah I was a bit apprehensive about what it was actually going to be like.*
>
> *P5-SITE1*

Ten of the participants who received the ES felt prepared to receive it:

I was prepared as I could have been, I think. I had things explained to me.

P2-SITE4

Other participants felt that their preparation for the procedure could have been improved. One participant was unsure what to expect:

But I don't really remember anything [pause] specific and whether that's because I blocked it out, I'm not entirely sure, but I don't remember talking about it in detail with any of the staff members at all.

P2-SITE7

Another participant stated that she felt prepared for the procedure, but her experience could have been improved if more information was provided face to face:

I think I would prepare well enough, I don't think so there'd be anything more what I could prepare myself better for it . . . maybe giving just a little bit more information face to face and explain more because just asking to, do you have any questions is not enough because not knowing much about that procedure you don't have actually much question.

P3-SITE1

One participant suggested that general information about the scratch procedure itself was required:

Like they said there will be the scratch before the egg but I don't know what the scratch does. That's the one thing I didn't find out about; what will the scratch do.

P4-SITE8

A number of participants felt that, having received the procedure, information relating to the pain and side effects of the procedure were lacking in the PIS:

Perhaps if there was any so like when I mentioned about the, the time involved like is there recovery time involved. So, would it mean taking time off work for example?

P1-SITE8

Staff perception of conversations with participant regarding the endometrial scratch procedure

Site staff discussed introducing the ES in terms of relating the procedure to other procedures the participant may have been aware of:

It is rather like having a smear test or if couples have had an IUI, saying it's very similar to that and then a small scratch of the lining of the womb.

RN-SITE2

One nurse would show the participants the pipelle (the instrument used to undertake ES) to reassure them:

I had a little pipelle so I could show them what it was like and so they could like feel the end of it their self, it wasn't like a scary thing you know. I found that helped, was quite helpful and most people were like, 'oh it will be fine'. They were quite happy.

RN-SITE3

Participants were made aware by the sites that they might feel pain during the ES procedure:

We are aware that it will cause some discomfort.

RN-SITE4

However, one site discussed not wanting to give patients a preconceived idea of the pain involved in receiving the procedure, owing to the variability in pain that patients feel:

> *I didn't want to give them a preconceived idea, because obviously you wanted them to do a pain score . . . some people can find it you know uncomfortable, rather than say it's going to be painful.*
>
> *RN-SITE2*

The preconception that the ES may be painful may have been due to the way in which nurses described the procedure, relating it to a painful one:

> *It's like a smear test, we don't know, it's a bit painful, etc. etc. Quite a few women went 'oh no you know I don't want to do that, I don't want to do that it sounds horrible'.*
>
> *RN-SITE6*

Experience of pain during and after the procedure

The pain experienced during and after the procedure varied considerably across the interviewed participants. For five participants, the procedure was more painful than they had prepared for. For these participants, the information they received prior to receiving the procedure did not reflect the actual pain experienced:

> *I can say from my experience it was quite a painful process. It wasn't that straightforward and it took you know all day to recover afterwards. So, it wasn't that easy and like soft, like as [the] nurse described [to] me at that time.*
>
> *P3-SITE1*

Three participants' pain lasted after the procedure itself and affected the activities they could undertake in the following days:

> *. . . and that [the endometrial scratch] then put me out of action really for a few days. I just didn't want to, you know – I just wanted to get home, curl up and the only way I can think about it now is probably like the worst period pain I've ever sort of had.*
>
> *P2-SITE7*

Other participants stated that the procedure was not painful ($n = 1$), not very painful ($n = 2$) and painful but tolerable ($n = 4$).

Site staff were aware of the variability in pain levels:

> *Let's be fair you know some people didn't feel a thing, others found it fairly uncomfortable, but that's you know, because people do.*
>
> *RN-SITE2*

What were recruited trial participants' and research site staff's perceptions of the recruitment process and how did research sites deal with recruiting to two large randomised trials; how can the recruitment process be improved in future trials?

Process of recruitment

The majority of participants were happy with the way in which they were approached to participate in the trial. Many could not think of ways in which recruitment to the trial could be improved:

> *I think it was all done really well. It was just, it was really well explained and I don't, I don't think, if it was happening again, I think it, I wouldn't change anything about it.*
>
> *P2-SITE8*

Exceptions to this were three participants who felt that the approach to take part in the trial was too informal, with either the participants having to ask the site staff about the trial or the trial being mentioned to the participant at the last minute when there were previous opportunities to raise this:

> I don't know if it should be more formal really, because it was just kind of, if they remembered on that day that I would be a, I'd be a possible candidate so it wasn't, do you know what I mean, maybe you should get it in a pack or something at the beginning, or told at the very beginning.
>
> *P2-SITE1*

One particular important factor was the relationship with the RN; there was a particular emphasis on discussing the trial and asking questions:

> There was plenty of opportunity to ask questions, and to sort of hear about the rationale as to why the trial was taking part, and what they were hoping to prove and disprove by it, and the reasons as to why the research was being conducted.
>
> *P4-SITE7*

Site staff were aware of the importance of this relationship. One staff member thought that her visibility in the fertility unit aided her first approach to participants:

> I found [approaching participants] quite easy really, because we're quite visible on the unit and people know us. I think it wasn't like there was somebody from a different unit asking them if they were interested in participating in the study.
>
> *RN-SITE1*

Site staff were also happy with the recruitment process and did not recommend any improvements.

Reflections on recruitment materials

Four participants could not recall the recruitment materials, owing to the time that had elapsed since they were recruited. Five participants stated that the materials were clear and informative:

> I thought the information sheet was really useful and informative.
>
> *P4-SITE7*

Some participants stated that the trial materials provided were not too onerous to read:

> Yeah, I think it was maybe like a double-sided page it wasn't kind of too much information.
>
> *P2-SITE9*

Others reflected on the amount of information provided more negatively, with two participants discussing the amount of information they received as 'overwhelming':

> . . . it can be very overwhelming, having lots of trial information given to you at the same time as all your IVF treatment. It just means that your pack of information to read through, I think I had about 20 different documents or leaflets to read through.
>
> *P4-SITE7*

Interviewees were asked if they required any additional information not included in the PIS. One participant required additional information regarding whether or not she could receive ES outside the trial:

> If they can provide people to say look if you want do it with the scratch this is all information and if you wanna do it yourself that's fine. So people will have more support and more kind of, what do you call, they can go and do it themselves to just . . .
>
> *P4-SITE8*

Another participant suggested that the PIS should contain more information about related studies:

> *After reading about other studies in the newspapers like The Guardian [Guardian Media Group, London, UK] and things like that, it might have been also interesting to hear a little bit about other studies that were going on, if there are any things like that. Because, I think when I did read about this study in The Guardian . . . and it suggested that it wasn't helpful. There was part of me that I think became a little bit angry that maybe I'd put myself through it at the time.*
>
> *P2-SITE7*

Two participants discussed that they would like more information about the possible effect of the ES on IVF success rates:

> *For me it was about the prospect of how the chances of the research would improve chances of a successful IVF treatment . . .*
>
> *P1-SITE1*

Trial staff perceived the information provided on the PIS as adequate:

> *I thought the recruitment was fine, the letter and the information leaflet, yeh. Very clear.*
>
> *RN-SITE1*

Staff stated that there was no need to supplement the PIS with additional information:

> Interviewer: *Did you need to supplement any of them, with additional information?*
>
> Participant: *No, I found it fine and I had very few enquiries from patients, they understood it really well.*
>
> *RN-SITE4*

Preconceptions

Many participants had preconceptions of the ES procedure that influenced their experience of the recruitment process.

Eighteen participants appeared to have a positive preconception of the effect that the ES would have on the outcome of the IVF. These participants stated that a major reason to participate was to increase their chances of a positive outcome from their IVF cycle:

> *. . . obviously if they'd have said to us the Scratch trial probably won't have improved our chances of getting pregnant, I probably wouldn't have, wouldn't have taken part [hmm]. So, we, you know we did it on the basis that we wanted to be on the clinical sample so that we could improve our chances of getting pregnant, so yes.*
>
> *P1-SITE1*

For many, these positive preconceptions originated from the internet. Some of these positive preconceptions from the internet came from participants reading that other women attending privately run units had received the procedure:

> *I think I had positive preconceptions, I think, because I'd been reading forums and read about private clinics doing scratches before.*
>
> *P2-SITE8*

There were subtle differences in the way participants described the information provided by research staff about the possible effects of the ES on their chances of a successful IVF outcome. For some

participants, the information provided by the clinic staff ensured that they were aware that there was insufficient evidence to suggest that the procedure would improve their chances of success:

> *The consultant suggested that there wasn't much evidence about whether the scratch was helpful or not when it comes to success of IVF so . . . So the purpose of the trial was to get some evidence behind that and so that decision was completely up to me as to whether I wanted to take or not.*
>
> *P1-SITE8*

Other participants seemed aware of the possible negative effects on their chances of success, but seemed to believe that it was more likely to be beneficial:

> *The positive effects were obviously, that you know, that if you did have the scratch, it could mean that you could fall pregnant naturally in the future . . . the lady had explained to us, and it came across very positive, if you like, the way that she put it across to us, but she did also make us fully understand that it wasn't 100% guaranteed, it's a bit like IVF, you might fall pregnant, or you might not.*
>
> *P1-SITE9*

Some participants discussed being told by staff at their fertility unit that the ES would benefit them, despite the lack of evidence to suggest this:

> *. . . obviously they do believe scratches increase chances of success.*
>
> *P3-SITE7*

One participant discussed that she thought the success rate for the scratch was high; it was unclear where this preconception originated from:

> *Just the success rate for the people who were trying to do IVF, who were trying to have babies. The success rate after the scratch was high.*
>
> *P2-SITE5*

Five participants discussed that they believed the ES procedure could not 'harm' them, but did not discuss being specifically told this by the research staff:

> *There wasn't anything that could potentially harm my chances of success. It wasn't, you know, it wasn't like trying out a new medicine or anything like that . . .*
>
> *P2-SITE8*

One participant felt influenced by clinic staff in a subtle way:

> *But unless, well you don't know whether it helps or not but there seems to be this, this influence from some within the clinic that almost nudge, wink, wink, you know, 'you're gonna be fine because this really does help' . . .*
>
> *P2-SITE7*

Other participants (*n* = 4) had not heard of the scratch prior to their involvement in the trial and stated that they did not have any preconceptions.

Site staff were aware of positive preconceptions:

> *They would generally say 'oh I've heard that, about the procedure and it can benefit the chances of you getting pregnant'.*
>
> *RN-SITE4*

Other staff discussed negative preconceptions:

> *Some patients did have concerns whether it would negatively affect [OK] their treatment outcome.*
>
> *PI-SITE1*

It was important for the sites to understand the participants' preconceptions, as sites discussed not wanting to recruit individuals who were not in equipoise:

> *If they are going, to really want a scratch but sign up anyway and then sneak off and get one privately and not let me know [yes] then that's not good for the research. So just very open and said if you feel very strongly, then you know I wouldn't want to recruit you for the trial . . .*
>
> *RN-SITE2*

Setting participants' expectations was seen as important in order to prevent resentful demoralisation of being randomised to the control arm:

> *We don't really know if it works or not, so I had to try and be a bit err I don't know pragmatic maybe, if that is the word. So that they didn't get, so they didn't feel sort of, too upset if they weren't randomised to have scratch done.*
>
> *RN-SITE3*

Dispelling positive preconceptions

Staff described the conversations they had with participants regarding taking part in the trial. Staff highlighted the importance of providing 'honest' and accurate information to participants:

> *. . . so to just be honest but say I had to say well they don't know if there is evidence for this, so they could get on it and it might be beneficial or we don't know if it is beneficial or not so and they are still going to get the same care.*
>
> *RN-SITE3*

Staff discussed adjusting participants' expectations regarding the possible effects of the ES procedure. This was mainly undertaken by explaining the lack of an evidence base for the scratch:

> *. . . we did explain to them comprehensively the evidence, around the endometrial scratch and the requirement for a big trial like this to be conducted for us to be able to know the answers and to know whether they should be offered to all patients going through forward for IVF treatment in the future.*
>
> *PI-SITE1*

One member of staff would send participants the abstract of academic manuscripts if they required more information:

> *I just used to talk about things like that but no more, no nothing else and sometimes people wanted more about that. So I would say if they really, really pushed it, I would send them abstracts of some of the research that I knew, that sort of thing.*
>
> *RN-SITE6*

However, site staff discussed that trial participants found it difficult to accept that if one study suggested benefit, they could be randomised to the control arm:

> *It wasn't that difficult to explain it to them although sometimes patients found it difficult to accept that if there has been some study that indicate potential benefit, then patients found it difficult to accept that they could be randomised into the control arm.*
>
> *PI-SITE1*

Randomisation to the control arm

Disappointment at being randomised to the control arm influenced the experience of recruitment for five out of the eight participants randomised to this arm.

Two of these participants discussed slight disappointment, but quickly moved on from this. These participants put their disappointment into the context of usual care:

> *I think there was a slight bit of disappointment, which is ridiculous really because it's randomised. I wasn't sort of like devastated or anything like that . . . At the end of the day I was having the treatment. It was, that's what current practice is, to not have the scratch for a first time if under 35 so I was just like well we'll go with that and see what happens. So I didn't dwell on it or anything like that.*
>
> *P4-SITE7*

For one of these participants, randomisation to the control group triggered an emotional response:

> *I was really quite disappointed and quite emotional when I, didn't get on the, scratch trial itself because I was worried that because I didn't have that, that my chances of getting pregnant were minimised.*
>
> *P1-SITE1*

One participant felt staff approached randomisation to the control group too sensitively:

> *I remember thinking they seemed like I was going to be upset about it 'cos they did it in a really sensitive way when they told you. But I was a bit like yeah it's fine don't worry it's not the end of the world like I'm fine with it. I remember thinking oh they were like really sensitive about it as if I should be upset about it, should I be upset about it . . .*
>
> *P1-SITE6*

A small number of participants did not seem to understand the role of the control group in the research, with one participant suggesting that more patients should have the opportunity to receive the ES procedure:

> *I don't think so, I know when we were asked to do it she did have to put our details onto a system and we would be randomly chosen and it would either be yes you're participating or no you're not, if you wanted greater numbers, then you could automatically ask everybody.*
>
> *P1-SITE3*

A misunderstanding of the role of the control group was also evident when participants discussed randomisation to the intervention arm of the trial in terms of being 'picked' or 'selected' for the intervention:

> *So, this, it was just an option given by my consultant at the time that if I wanted to take part in the trial and, that would mean that I either was or wasn't selected to have the scratch.*
>
> *P1-SITE8*

Site staff perceptions of control group disappointment

Many site staff were aware of some participants' disappointment about being randomised to the control arm:

> *Some people did express great disappointment and in particularly that way round actually it was nearly always the ones that had wanted it obviously 'cos it's the intervention.*
>
> *RN-SITE6*

Once learning of their randomisation outcome, some participants enquired about receiving the ES procedure outside the trial:

> *We had a couple, I think we had and I can't remember exactly but we had at least one person who did, who was in the control group and then they paid for a private scratch.*
>
> *RN-SITE6*

However, this seemed to be the minority of patients, with most randomised to the control arm accepting the outcome of randomisation:

> *The absolute vast majority of them did, they just accepted that, you know we didn't have many just kind of going 'oh well in that case I'm not going to do it' or 'I'm going to go and do it somewhere else'.*
>
> *RN-SITE6*

The negative reaction from participants was such that one member of staff described feeling worried about informing participants of their randomisation outcome:

> *I did have some trepidation, phoning them to tell them [they] couldn't do it.*
>
> *RN-SITE3*

Another individual reflected on the importance of wording the outcome of randomisation to the patient:

> *Certainly I was very careful not to say you know, I'm sorry you haven't been in. I would just say, you've been randomised to the control arm.*
>
> *RN-SITE2*

Staff described mechanisms they had developed for dealing with this disappointment; this included explaining the importance of the control group, describing the lack of evidence to support ES and focusing on the increased contact with participants that research allows:

> *So I think it's explaining the importance of being in that control group to them. And how we can, you know, we don't know if there was any benefit until we've actually, that's the whole purpose of doing the trial, to answer a question.*
>
> *RN-SITE1*

> *Try and sort of explain that, there was no evidence whether it was a benefit anyway, and that normally if they were going through their treatment, I understand they wouldn't be having a scratch, so it wasn't as if they missing out on anything.*
>
> *RN-SITE6*

> *If it's the control you know it's a bit of extra contact with people and I would always say 'look I'm here in the unit, you're not having the scratch, but if I can help you in any other way, feel free to ring the research mobile'. You know to try and always make it I suppose to soften the blow really, and for those who were disappointed.*
>
> *RN-SITE2*

Delay to in vitro fertilisation

Discussing the potential delay to IVF that participating in the trial may cause was discussed by site staff as a barrier to recruitment:

> *Well they had to delay, delay starting their treatment, so I had to explain to them that they had the option to withdraw from the study, if they felt that they found their treatment was being compromised because they were taking part in the study.*
>
> *RN-SITE4*

Two site staff stated that it was important to be clear with participants from the outset that if they were randomised to ES they would need to delay their IVF cycle:

> *We made it very clear at the beginning looking at, we did look [at] their menstrual diary and everything and we did say that if they were randomised for the intervention arm what the wait would be.*
>
> *PI-SITE1*

It was important to ensure that it was the participant's decision to delay their cycle:

> *I think I would always let them try and make the decisions, whether they really wanted to take part in the study or not [mm]. I never tried to like say oh well, not to, yes I'd like, I always used to say it's, it was their decision, whether they really, really wanted to take part in the study or not and if they felt that they didn't want to leave their IVF for another month then that was, that was their decision to do that.*
>
> *RN-SITE5*

Delaying IVF due to participants not being available to receive the ES was discussed by one participant as the biggest challenge during the trial:

> *I suppose the big challenge was the time issue of, if patients weren't around to have the scratch, then it meant they had delay in their treatment.*
>
> *RN-SITE4*

Recruitment to multiple trials

Nine of the trial participants were receiving treatment from sites that were also recruiting to the E-Freeze trial,[42] and therefore these participants may have been eligible for both trials. Eight of these participants did not recall having been approached for the E-Freeze trial.[42]

One participant recalled being provided the recruitment materials for both trials, and felt positive regarding the choice she was given:

> *I think that it was nice because, if we didn't qualify for you know, one particular trial, there was another trial that we could, we could participate in . . . So it was kind of, it wasn't difficult but it was nice that we had options to choose between and what works for myself and my husband.*
>
> *P2-SITE4*

One of these participants reflected on how the process of approaching participants for multiple studies could have been improved by not including information about research by default within information packs:

> *They probably should have, maybe picking one trial at a time to sort of drop in, so they kind of given the options that are relevant to you at different stages might be helpful. Or perhaps even, thinking about it, when you were going for your initial appointment, they could ask you when you're first seeing somebody in the clinic, and there are various research trials that you could potentially be involved with, would you like any information on any of them? Rather than just including it in everyone's pack.*
>
> *P4-SITE7*

Three of the interviewed site staff had been involved in recruiting to both this trial and the E-Freeze trial.[42]

These individuals described different processes used when approaching for the first time participants who were potentially eligible for both trials. Two of the staff members stated that they would give information about both trials to participants who were eligible for both:

> *I didn't really select, a patient could only be approached for one trial, and so if a patient was seen, that they could, you know be approached for more than one trial, I would give them the information, on both trials.*
>
> *RN-SITE4*

The other interviewee stated that they would approach participants eligible for both trials at two different time points:

> *Well to be honest we saw them at different stages, which was good, so actually the Endo Scratch, anyone potentially for Endo Scratch we would approach earlier. And with the E-Freeze, we would actually approach at the time they were starting their treatment. So completely different clinics and different situations really.*
>
> RN-SITE2

This individual cited two reasons for approaching participants at different time points. The first was concerns around providing participants with too much written information:

> *Yes, and it's a lot to take on board anyway. I mean when they are coming for an appointment it's pretty hard core, they are getting an awful lot of information. And to try and give them information about more than one trial as well, I don't, I think that would be difficult. I'd find that difficult [laughing].*
>
> RN-SITE2

The second reason seemed to be logistical, in that the Endometrial Scratch Trial required recruitment at an earlier time point in the patient pathway than usual:

> *Having worked here for a sort of 2 or 3 years now we, if they're ones for IVF we tend to approach them at the point they are coming through for a down-regulation scan. But obviously that just wouldn't have worked with the endometrial scratch, because you need to do the procedure beforehand.*
>
> RN-SITE2

What were recruited participants' perceptions of the data collection tools used during the trial, specifically text messages to collect pain and electronic questionnaires?

Perceptions of the trial questionnaires

The majority of participants stated that the questionnaires they had been asked to complete in order to collect data for the health economic component of the trial were straightforward and simple:

> *. . . they were quite short, sharp and to the point which is, which is how it should have been.*
>
> P2-SITE4

A minority of participants felt that the questionnaire was too repetitive ($n = 2$) and challenging to complete due to requiring retrospective data ($n = 2$):

> *It was because it was asking me you know how many times in the last week you know or month I was seeing midwives, doctors, consultants and things like that. And obviously because I had to see quite a few people it was you know, oh how many, you know you were there thinking oh how many times I had to write it down – several times.*
>
> P1-SITE1

Because of this, participants recommended sending the questionnaire more regularly, rather than requesting more information less frequently:

> *I would just be more likely to remember what's going on. If I had a long one after 6 months, it'd be fine remembering the more recent times but if it was looking back to 6 months prior, I think it's difficult to try and remember what was happening and how you felt about things, to accurately respond to it.*
>
> P4-SITE7

Regarding whether they would prefer to receive electronic or paper questionnaires in the future, five participants would prefer electronic, three paper and two did not have a preference.

Perceptions of the text messages to collect pain scores

The majority of participants asked about the text messages felt that it was a good way to collect their pain post procedure as it was quick and simple:

> *I found it obviously quite simplistic which is good, because – you receive that text message at any point of the day, so if you're busy, you can quickly reply straight away.*
>
> *P3-SITE7*

One participant reflected on the difficulty of measuring pain using a Likert scale:

> *. . . there isn't much nuance to it if you, just, you know, it is just a number I think but I can see that, you know, you need to have that as a kind of, as statistical information really, I can see that it's important to have, but because it probably does vary so much it's perhaps more useful to just have a range of scores than maybe to have people going well it was alright, or it wasn't that bad, or it was quite bad, how do you kind of measure that.*
>
> *P4-SITE4*

But others thought that this was a suitable way to collect pain scores:

> *I believe it captured it quite well. I know everyone's perspectives on 10 is different, but I think it's the best way to get the overall view on the matter.*
>
> *P3-SITE7*

Conclusions

- Participants recruited to the Endometrial Scratch Trial were, on the whole, happy with the recruitment practices utilised in the trial.
- A small number of participants felt that the invitation to participate in the trial was too informal, a minority described the amount of written material they received as 'overwhelming', and the potential for participation in the trial to delay the IVF cycle caused issues.
- Many participants did not recall being approached for both this trial and the E-Freeze trial,[42] despite being recruited at sites that were running both trials.
- Interviewees had largely positive preconceptions regarding the possible effect of the ES procedure on the likelihood of a successful outcome to their IVF cycle. Most of these preconceptions originated from participants' own internet research. Site staff were aware of participants' positive preconceptions and, in order to dispel these, described the evidence base to participants. Some participants described being subtly informed by site staff that the procedure would be beneficial for them, and appeared to not understand that receiving the ES could potentially reduce their chances of a successful IVF outcome. This, coupled with the positive preconceptions of the ES, led to dissatisfaction among some participants randomised to the control arm.
- Seven participants were interviewed who decided not to undergo the ES procedure, despite being randomised to do so. For two of these participants, 'dropout' from the intervention may have been avoided with improved organisation at the fertility unit.
- A small number of participants described the ES procedure as being more painful than then expected. For three participants, the pain post procedure affected the activities they could undertake in the days following ES.
- Participants were generally happy with the data collection processes used during the trial (questionnaires and text messages).

Chapter 10　Systematic review and meta-analysis results

This chapter reports the results of the systematic review and meta-analyses that incorporated results from this trial (presented in *Chapter 7*).

Study selection

The searches identified a total of 1462 records (*Figure 14*). After removal of duplicates, 926 records were assessed for inclusion in the review at the title and abstract stage. A total of 845 records were excluded at this stage, for reasons that included not being a RCT, not being the correct population of interest (i.e. not in those undergoing IVF) and not being in the English language.
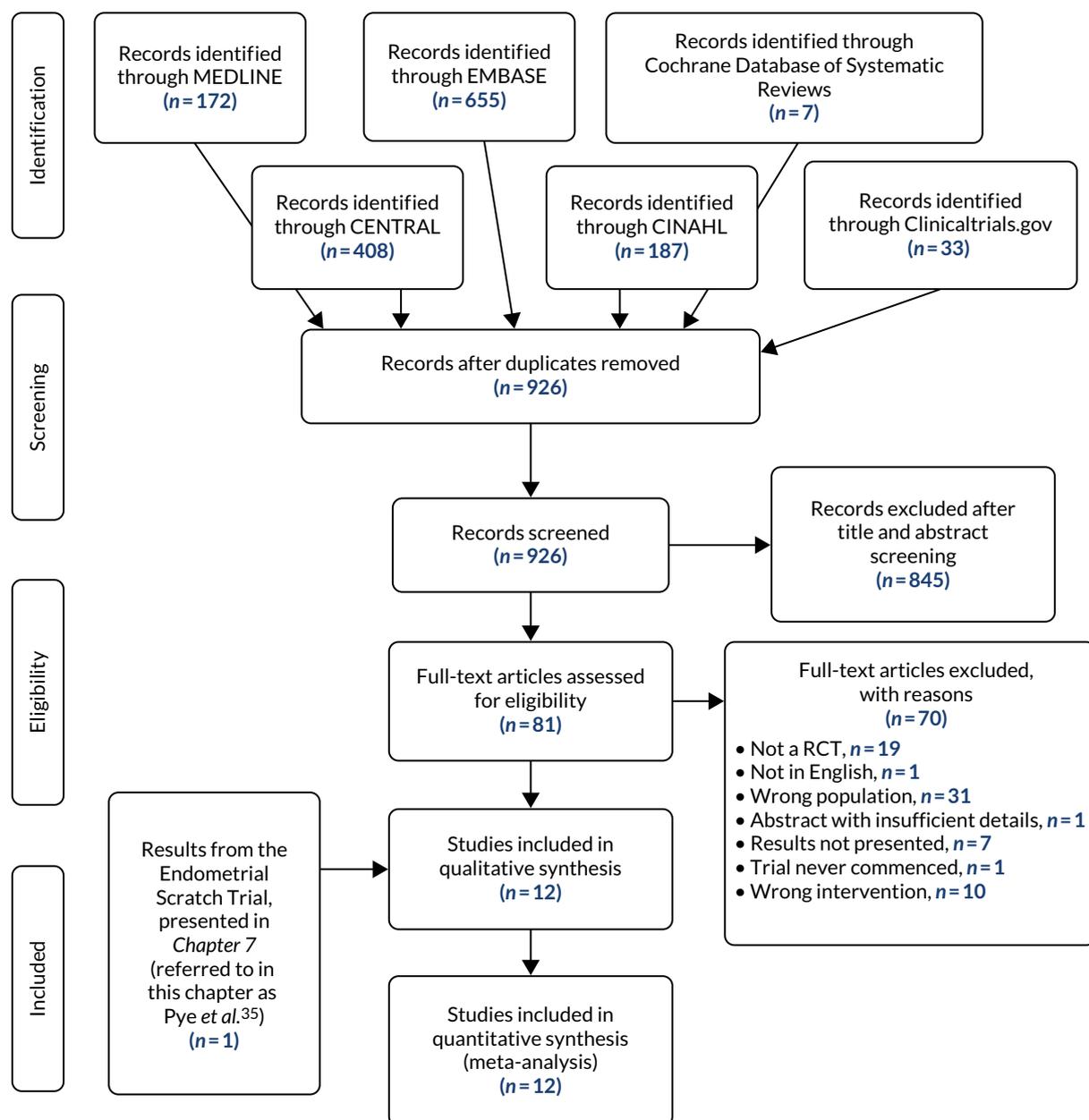
FIGURE 14　The PRISMA flow chart.

Eighty-one studies were assessed at the full-text stage, of which 70 were excluded for reasons presented in *Appendix 26*. We attempted to contact the authors of studies where results were not presented and the timelines of the trial suggested that the trial results may be available ($n = 4$),[68–71] where the abstract included insufficient detail and a full-text article was not available ($n = 1$),[72] and studies undertaken in an unselected population, where it appeared that patients undergoing their first IVF cycle were included but their outcomes were not presented separately ($n = 14$).[73–86] Of the 14 authors that were contacted, eight did not respond. One author confirmed that the trial was not eligible for inclusion as participant recruitment never commenced.[70] Separate outcomes data for those participants undergoing their first cycle of IVF were obtained from the authors of four trials undertaken in unselected IVF populations and were therefore included in the review.[14,22,28,87]

Polanski *et al.*,[88] whose study included women undergoing unselected IVF cycles, could not be contacted to obtain data for first-cycle participants only; however, unpublished data presented in a recent systematic review were used.[16] A risk-of-bias assessment could not be conducted for this study owing to a lack of methodological details described in the abstract.

After screening, 12 RCTs were eligible for inclusion in the review.

Characteristics of included trials

Table 29 summarises the characteristics of the 12 included RCTs comprising 3234 participants undergoing their first IVF/ICSI cycle.

Nature of trials and geographical coverage

Only three of these studies were multicentre RCTs.[14,20,35] All 12 studies were individually randomised and were conducted across 10 countries: Iran, Hong Kong, Egypt (two studies), China, USA, Belgium, Spain, Turkey, Brazil and the UK. One RCT was undertaken multinationally across five countries.[14] The number of participants included in each trial undergoing their first IVF cycle ranged from 18 to 1048.

Eleven studies were two-arm RCTs[14,17,19–22,28,35,87–89] and one was a four-arm RCT.[18] Nine of the two-arm RCTs compared the ES procedure with usual care and two trials included a comparator involving a sham procedure.[21,87] The four-arm trial compared ES at two different time points with a sham procedure undertaken at the same two different time points in the menstrual cycle: proliferative and luteal.[18]

Recruitment to three of the trials was prematurely ended owing to an unplanned futility analysis showing no differences in clinical pregnancy rates between intervention and control groups in one trial,[21] a planned interim analysis identifying higher miscarriage rates in the intervention group in another trial,[28] and identifying a significant benefit of the intervention during a planned interim analysis in the final trial.[87]

Endometrial scratch procedure and timing

The method of undertaking ES was largely similar across studies. Most used a pipelle to invoke injury, except one trial that used an embryo transfer catheter,[22] one that used a Novak Curette[17] and another that used either a pipelle or Wallace endometrial sampler.[88] However, the timing of ES differed across trials. Two trials undertook ES during the IVF cycle, either on the day of egg collection[17] or during ovarian stimulation.[28] Ten trials undertook ES in the menstrual cycle prior to IVF, with seven within the luteal phase[19–22,35,88,89] and one during the early or mid-luteal phase.[87] Lensen *et al.*[14] undertook ES at any point between day 3 of the menstrual cycle prior to ES and day 3 of the cycle in which IVF was being undertaken. However, this trial reported that the median time (IQR) between ES and embryo transfer was 35 (22–39) days, and therefore it is likely that the majority of participants received ES in the menstrual cycle prior to IVF. Liu *et al.*[18] undertook ES either during the proliferative or luteal

TABLE 29 Characteristics of included trials

Author, publication year and trial registration number	Single/multi-centre, country and period of study	Number of arms, intervention and control conditions	Key inclusion criteria: age, BMI, ovarian markers/ ovarian response, donor eggs	IVF treatment received	Intervention: instrument, procedure	Timing of ES: prior to IVF/during IVF, detail	Controls	Outcomes
Karimzade 2010[17] (published data only) (NCT00846183)	Single centre Iran June 2008 to January 2009	Two arms: ES, usual care	< 38 years > 19 and < 30 kg/m² Day 3 FSH < 12 mIU/ml N/A	IVF with or without ICSI	Novak Curette Two small biopsies obtained from anterior and posterior walls of uterus	During IVF Day of oocyte retrieval	Usual care	CPR, OPR, implantation rate, AE
Yeung 2014[89] (published data only) (NCT01977976)	Single centre China March 2011 to October 2013	Two arms: ES, usual care	N/A N/A N/A Excluded	IVF with or without ICSI	Pipelle Endometrial tissue taken between the fundus and internal os	Prior to IVF 7 days post LH surge/day 21 of cycle preceding IVF	Usual care	LBR, OPR, CPR, implantation rate, MR, MPR, QP, AE
Mahran 2016[20] (published data only) (ISRCTN61316186)	Multicentre (two centres) Egypt June 2012 to September 2014	Two arms: ES, usual care	20–40 years N/A FSH ≤ 12 mIU/ml and ≥ 2 good-quality embryos replaced N/A	IVF with or without ICSI	Pipelle The posterior wall of the uterine cavity scratched three times	Prior to IVF Day 21–24 of cycle preceding IVF	Usual care	LBR, OPR, CPR, implantation rate, MR, MPR, QP, AE
Maged 2018[19] (published data only) (NCT02660125)	Single centre Egypt January 2016 to March 2017	Two arms: ES, usual care	< 40 years N/A FSH < 10 mIU/ml N/A	ICSI	Pipelle Endometrial tissue removed from between the fundus and inner os	Prior to IVF	Usual care	CPR, implantation rate, MR, MPR, AE

TABLE 29 Characteristics of included trials (*continued*)

Author, publication year and trial registration number	Single/multi-centre, country and period of study	Number of arms, intervention and control conditions	Key inclusion criteria: age, BMI, ovarian markers/ ovarian response, donor eggs	IVF treatment received	Intervention: instrument, procedure	Timing of ES: prior to IVF/during IVF, detail	Controls	Outcomes
Liu 2017[18] (published data only) (ChiCTR-IOR-17011506)	Single centre China February 2012 to November 2014	Four arms: ES (proliferative phase), ES (luteal phase), sham procedure (proliferative phase), sham procedure (luteal phase)	≤ 40 years None FSH < 12 mIU/ml None	IVF with or without ICSI	Pipelle The piston was drawn back to the end of the sheath to create a negative pressure. The sheath was rotated and moved back and forth within the uterine cavity to ensure adequate endometrial tissue has been obtained	Prior to IVF Proliferative phase (day 10–12) or luteal phase (days 7–9) of preceding cycle	Sham procedure	LBR, CPR, BPR, implantation rate, MR, EPR, MPR, QP, AE
Eskew 2019[21] (published data only)	Single centre USA September 2013 to July 2017	Two arms: ES, usual care	18–43 years None None None	IVF	Pipelle The sheath was rotated and the pipelle moved up and down three or four times ensuring tissue was obtained	Prior to IVF 7–13 days post LH surge in cycle preceding IVF	Sham procedure	LBR, CPR, MR
Lensen 2019[14] (published and unpublished data) (ACTRN12614000626662)	Multicentre (13 centres) New Zealand, Belgium, Sweden and UK June 2014 to June 2017	Two arms: ES, usual care	> 18 years None None None	IVF with or without ICSI	Pipelle The catheter is advanced to the fundus and then moved back and forth four times, rotating 90° each time, to collect the biopsy	Prior to IVF Day 3 of the preceding cycle to day 3 of the IVF cycle	Usual care	LBR, OPR, CPR, MR, EPR, MPR, SBR, NP, AE

Author, publication year and trial registration number	Single/multi-centre, country and period of study	Number of arms, intervention and control conditions	Key inclusion criteria: age, BMI, ovarian markers/ ovarian response, donor eggs	IVF treatment received	Intervention: instrument, procedure	Timing of ES: prior to IVF/during IVF, detail	Controls	Outcomes
Mackens 2020[28] (published and unpublished data) (NCT02061228)	Single-centre Belgium April 2014 to October 2017	Two arms: ES, usual care	≥ 18 and < 40 years ≤ 35 or ≥ 18 kg/m^2 None Excluded donor eggs	IVF with or without ICSI	Pipelle The piston was then withdrawn and the device rotated while moving up and down four times	During IVF Days 6–7 of ovarian stimulation	Usual care	LBR, CPR, MR, EPR, QP, AE
Izquierdo Rodriguez 2020[22] (published and unpublished data) (NCT03108157)	Single centre Spain January 2017 to October 2018	Two arms: ES, usual care	None None None Only donor eggs	ICSI	Endometrial biopsy catheter Once in the cavity, the catheter piston was partially removed to create a suction effect and the catheter was then moved back and forth and rotated 360° to scratch the four walls	Prior to IVF Luteal phase: 5–10 days before the start of the period	Usual care	LBR, CPR, OPR, implantation rate, MR, MPR
Nastri 2013[87] (published and unpublished data) (NCT01132144)	Single centre Brazil June 2010 to March 2012	Two arms: ES, sham procedure followed by usual care	< 38 years None None None	IVF with or without ICSI	Pipelle The device was introduced through the cervix up to the uterine fundus when the piston was drawn back to generate pressure. The device was moved back and forth while being rotated over a period of 20 seconds	Prior to IVF 7–14 days prior to planned start of controlled ovarian stimulation	Sham procedure	LBR, CPR, MR

TABLE 29 Characteristics of included trials (*continued*)

Author, publication year and trial registration number	Single/multi-centre, country and period of study	Number of arms, intervention and control conditions	Key inclusion criteria: age, BMI, ovarian markers/ovarian response, donor eggs	IVF treatment received	Intervention: instrument, procedure	Timing of ES: prior to IVF/during IVF, detail	Controls	Outcomes
Pye 2018[35] (data published within this report) (ISRCTN23800982)	Multicentre (16 sites) UK July 2016 to October 2018	Two arms, ES, usual care	18–37 years ≥ 35 kg/m² FSH< 10 mIU/ml None	IVF with or without ICSI	Pipelle The plunger of the pipelle was withdrawn to generate negative pressure before being rotated and withdrawn several times	Prior to IVF Mid-luteal phase defined as 5–7 days before the expected next period, or 7–9 days after a positive ovulation test	Usual care	LBR, CPR, implantation rate, SBR, NP, AE, PTR
Polanski 2014[88] (data obtained from published systematic review) (NCT01882842)	Single centre UK February 2013 to June 2015	Two arms: ES, usual care	< 49 years None None None	IVF with or without ICSI	Pipelle or Wallace endometrial sampler Sampler will be introduced up to the uterine fundus, and attempts will be made to obtain samples from all four walls of the endometrial cavity. The endometrial sampling repeated a maximum of four times. If following the first two attempts the endometrial sample is sufficient, procedure will be deemed complete	Prior to IVF 7–9 days post LH surge in the cycle preceding IVF	Usual care	LBR, MPR, CPR, MR, EPR

BPR, biochemical pregnancy rate; CPR, clinical pregnancy rate; EPR, ectopic pregnancy rate; IU: international units; MBR, multiple birth rate; MPR, multiple pregnancy rate; MR, miscarriage rate; NP, numerical pain score; OPR, ongoing pregnancy rate; PTR, preterm delivery rate; SBR, stillbirth rate; QP, qualitative pain score.

phases of the menstrual cycle; therefore, in this review, the two time points of delivery of ES or the sham procedure were combined, so that for each outcome there was one rate for the ES arm (both proliferative and luteal) and another for the sham arm (both proliferative and luteal).

We undertook an additional analysis in order to further assess the impact of the timing of ES procedure; we split trials into those that performed ES in the menstrual cycle prior to that in which IVF was being undertaken[17,28] and those that undertook ES in the menstrual cycle during which IVF was also delivered.[14,18–22,35,87–89] Two trials provided hysteroscopy to all trial participants, prior to IVF.[20,89]

Participant eligibility

Trials utilised different participant eligibility criteria. Ten trials had age restrictions with an upper limit of between 35 and 49 years.[14,17–21,28,35,87,88] Three trials restricted the BMI to an upper limit ranging from 30 to 35 kg/m[2].[17,19,35] Five trials selected participants who were deemed to have a good ovarian reserve, by allowing only those with a certain level of FSH to participate, with a maximum level of 10 IU/ml in two studies[19,35] and 12 IU/ml in three studies.[17,18,20] At the point of assessing eligibility, two studies set requirements for the number of oocyctes collected or embryos replaced: ≥ 2 oocytes replaced in one study[20] and 4–14 oocyctes collected in another.[17] Only one trial stipulated that the embryos replaced had to be of a certain quality, with Mahran *et al.*[20] stipulating that the two or more embryos replaced had to be of 'good' quality. However, the exact method of defining a good-quality embryo was not reported. Two trials excluded women receiving donor eggs[28,89] and one study only included those receiving donor eggs.[22]

Trial outcomes, definitions and timing of end points

Ten trials reported LBRs.[14,18,20–22,28,35,87–89] Clinical pregnancy rates were reported in all trials. However, this was defined inconsistently, with variation in the time point at which this outcome was assessed: at 4 weeks post embryo transfer in two trials,[19,20] 5 weeks in one trial,[17] 6 weeks in four trials[14,18,22,89] and 7[28] and 8 weeks[35] in one trial each (and unknown in three trials[21,87,88]). Ongoing pregnancy rates were reported in four trials, the timing of which varied between 12[14,17,22] and 20[89] weeks post embryo transfer.

Implantation rates were reported in seven studies.[17,18,20,22,35,89,90] Six studies defined this similarly as the number of gestational sacs, divided by the number of embryos transferred, whereas in Pye *et al.*,[35] this was defined as the number of gestational sacs divided by the number of participants randomised to each arm (under ITT principles). Miscarriage rates per clinical pregnancy were reported in 11 trials,[14,18–22,28,35,87–89] with the time point of data collection differing between 12 and 24 weeks of gestation, but were unclear in two trials.[14,28] Seven trials reported multiple pregnancy rates[14,18–20,22,88,89] and multiple birth rates were reported in one study.[35] Two trials reported stillbirth rate,[14,35] the definition of which was reported in only one trial,[35] with one also reporting preterm delivery rates.[35] A subjective assessment of pain of the ES procedure on a numerical rating scale was reported in three trials,[14,35,87] with four studies providing qualitative reports of pain.[18,20,28,89] Eight trials reported AEs and/or complications in the participating women,[14,17–20,28,35,89] and four trials reported such events in the baby or neonate.[14,22,35,87]

Risk-of-bias assessment

It was not possible to conduct a risk-of-bias analysis for Polanski *et al.*[88] The only information available was from a recent review, which used a previous version of the risk-of-bias tool and therefore the authors' assessments could not be considered in this review.[16]

The risk-of-bias summary can be seen in *Figure 15*. This relates to the review authors' judgement for each category in the risk-of-bias tool. Domains 1–3 were consistent across all outcomes per study. Across all studies and outcomes, there was none that was considered to be at a high risk of bias.

Domain 1 assesses whether or not the allocation sequence was random, if the allocation sequence was concealed and if there were baseline differences between the intervention groups. Nine trials used a computerised system to undertake randomisation,[14,17,19,21,22,28,35,87,89] one trial used sealed envelopes,[20]

Study	Domain 1. Randomisation process	Domain 2. Deviations from intended intentions	Domain 3. Missing outcome data	Domain 4. Measurement of the outcome	LBR Domain 5. Selection of the reported result	LBR Domain 6. Overall bias	Clinical pregnancy rate Domain 5. Selection of the reported result	Clinical pregnancy rate Domain 6. Overall bias	Ongoing pregnancy rate Domain 5. Selection of the reported result	Ongoing pregnancy rate Domain 6. Overall bias	Miscarriage rate Domain 5. Selection of the reported result	Miscarriage rate Domain 6. Overall bias	Multiple pregnancy rate Domain 5. Selection of the reported result	Multiple pregnancy rate Domain 6. Overall bias	Ectopic pregnancy rate Domain 5. Selection of the reported result	Ectopic pregnancy rate Domain 6. Overall bias
Eskew[21]	+	+	+	+	?	?	?	?	NA	NA	?	?	NA	NA	NA	NA
Izquierdo Rodriguez[22]	?	?	+	+	+	?	+	?	NA	NA	+	?	?	?	?	?
Karimzade[17]	?	?	+	+	NA	NA	?	?	?	?	NA	NA	NA	NA	NA	NA
Lensen[14]	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Liu[18]	?	+	+	+	?	?	?	?	NA	NA	?	?	?	?	?	?
Mackens[18]	+	?	+	+	+	?	+	?	NA	NA	?	?	NA	NA	?	?
Maged[19]	?	+	+	+	NA	NA	?	?	NA	NA	?	?	?	?	NA	NA
Mahran[20]	+	+	+	+	?	?	?	?	NA	NA	?	?	NA	NA	?	?
Nastri[87]	+	+	+	+	+	+	+	+	NA	NA	+	+	NA	NA	NA	NA
Polanski[88]	AN	AN	AN	AN	AN	AN	AN	AN	NA	NA	NA	NA	NA	NA	NA	NA
Pye[35]	+	+	+	+	+	+	+	+	NA	NA	+	+	NA	NA	+	+
Yeung[89]	+	+	+	+	?	?	?	?	+	+	+	+	?	?	NA	NA

Legend: + Low risk; ? Some concerns; – High risk; NA Not assessed; AN Assessment not possible.

FIGURE 15 Risk-of-bias assessments of the included trials.

and another used a table of random numbers.[18] Four trials were assessed as 'some concerns' in domain 1.[17–19,22] Studies were assessed as having 'some concerns' due to there being no information on how the allocation sequence was concealed. In all trials, there were no substantial differences in baseline characteristics across treatment arms.

Domain 2 considered the risk of bias due to deviations from the intended interventions, which was deemed to be of 'some concern' in three trials.[17,22,28] The differing dropout rate between the intervention group (8.5%) and the control group (2.2%) in Izquierdo Rodriguez et al.[22] resulted in this assessment. Similarly, in Karimzade et al.[17] there were four patients excluded from the analysis in the intervention arm and none in the control arm.

Domain 3 considered the risk of bias due to missing outcome data. The rate of missing data was low across all trials, therefore all were considered to be at low risk of bias for this domain.

All studies were judged to be at low risk of bias for domain 4, 'measurement of the outcome'. Given the nature of the outcomes being assessed, patient knowledge of the intervention was unlikely to have affected the analysis. Therefore, even though only 3 of the 12 included studies involved some form of blinding,[18,21,87] this is not considered to affect the patient outcomes. There was also no blinding in most of the included trials; however, it is unlikely that participants' or clinicians' knowledge of the intervention could have biased outcome assessment owing to the included trials using objective outcome measures unlikely to be influenced by the placebo effect.

Domain 5 assessed the risk of bias in the selection of the reported result. For all outcomes, studies were deemed to be of 'some concern' if the outcome was not specified prior to the start of the trial

(no protocol, trials registry or a retrospectively added trials registry), the timing of the outcome was not specified or the outcome specification in the paper did not match the protocol/registry.

Only Pye *et al.*,[35] Lensen *et al.*[14] and Nastri *et al.*[87] were considered to have a low risk of bias across all assessments.

Meta-analysis results

Primary synthesis: live birth outcome

The overall effect from the 10 trials that reported live birth outcome did not provide evidence to support a significant effect of ES on increasing the LBR (*Figure 16*): OR 1.17 (95% CI 0.76 to 1.79).[14,18,20–22,28,35,87–89] However, this analysis should be interpreted with caution, as there was evidence of significant heterogeneity between trials. Over three-quarters of the between-trial variability was not attributable to chance alone ($I^2 = 83\%$).

Exclusion of Mahran *et al.*[20] from the analysis minimised heterogeneity ($I^2 = 0\%$), but did not alter the conclusions of the analysis, which was consistent when both fixed-effects (see *Appendix 27*; OR 1.03, 95% CI 0.87 to 1.21) and random-effects (*Figure 17*; OR 1.03, 95% CI 0.87 to 1.22) models were used.

Study or subgroup	ES Events	ES Total	Control Events	Control Total	Weight (%)	OR M–H, random, 95% CI	OR M–H, random, 95% CI
Eskew[21]	19	39	19	27	7.6	0.40 (0.14 to 1.13)	
Izquierdo Rodriguez[22]	43	77	36	78	10.5	1.48 (0.78 to 2.78)	
Lensen[14]	81	298	74	278	12.4	1.03 (0.71 to 1.49)	
Liu[18]	25	70	29	72	10.2	0.82 (0.42 to 1.62)	
Mackens[28]	23	70	29	78	10.2	0.83 (0.42 to 1.63)	
Mahran[20]	134	209	56	209	12.1	4.88 (3.22 to 7.40)	
Nastri[87]	2	8	4	10	3.3	0.50 (0.07 to 3.85)	
Polanski[88]	25	51	23	60	9.6	1.55 (0.73 to 3.30)	
Pye[35]	202	523	195	525	13.1	1.06 (0.83 to 1.37)	
Yeung[89]	32	105	32	104	10.9	0.99 (0.55 to 1.78)	
Total (95% CI)		1450		1441	100.0	1.17 (0.76 to 1.79)	
Total events	586		497				

Heterogeneity: $\tau^2 = 0.35$; $\chi^2 = 53.15$, df = 9 ($p < 0.00001$); $I^2 = 83\%$
Test for overall effect: $Z = 0.70$ ($p = 0.48$)

0.1 0.2 0.5 1.0 2.0 5.0 10.0
Favours control   Favours ES

FIGURE 16 Forest plot of all trials reporting LBRs. df, degrees of freedom; M–H, Mantel–Haenszel test.

Study or subgroup	ES Events	ES Total	Control Events	Control Total	Weight (%)	OR M–H, random, 95% CI	OR M–H, random, 95% CI
Eskew[21]	19	39	19	27	2.6	0.40 (0.14 to 1.13)	
Izquierdo Rodriguez[22]	43	77	36	78	6.9	1.48 (0.78 to 2.78)	
Lensen[14]	81	298	74	278	20.4	1.03 (0.71 to 1.49)	
Liu[18]	25	70	29	72	6.0	0.82 (0.42 to 1.62)	
Mackens[28]	23	70	29	78	6.0	0.83 (0.42 to 1.63)	
Nastri[87]	2	8	4	10	0.7	0.50 (0.07 to 3.85)	
Polanski[88]	25	51	23	60	4.8	1.55 (0.73 to 3.30)	
Pye[35]	202	523	195	525	44.5	1.06 (0.83 to 1.37)	
Yeung[89]	32	105	32	104	8.0	0.99 (0.55 to 1.78)	
Total (95% CI)		1241		1232	100.0	1.03 (0.87 to 1.22)	
Total events	452		441				

Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 6.93$, df = 8 ($p < 0.54$); $I^2 = 0\%$
Test for overall effect: $Z = 0.35$ ($p = 0.73$)

0.1 0.2 0.5 1.0 2.0 5.0 10.0
Favours control   Favours ES

FIGURE 17 Sensitivity analysis of LBRs excluding Mahran *et al.*[20] df, degrees of freedom; M–H, Mantel–Haenszel test.

Mahran *et al.*[20] is a clear outlier, being the only trial to report an extremely significant effect of undertaking ES on LBR. There are several potential explanations for this heterogeneity, however, uncertainty remains in the actual cause. First, this is the only trial that required participants to have two 'good' quality embryos transferred to be eligible to participate in the trial; including such a stipulation prior to randomisation (i.e. prior to egg collection) may have resulted in participants being randomised to the trial and excluded because of poor embryo quality, biasing the results of this trial in a positive direction. Second, all participants received hysteroscopy prior to IVF, a procedure that can have similar effects to ES. Finally, an average of three embryos were transferred, whereas all other trials included in this review averaged one or two embryos transferred.

A subgroup analysis was undertaken to explore the effect of the timing of ES (see *Appendix 28*); one trial reported LBR and undertook ES during the IVF cycle,[28] and nine trials undertook ES in the menstrual cycle prior to ES.[14,18,20–22,35,87–89] The analysis identified a non-significant interaction between subgroups ($p = 0.37$), but with the point estimates (ORs) of the two subgroups suggesting that ES influenced LBR in different directions; the effect when ES was undertaken prior to IVF was 1.21 (95% CI 0.76 to 1.92) and during the IVF cycle it was 0.83 (95% CI 0.42 to 1.63). However, with only one trial in the 'during IVF' subgroup, there are insufficient data to draw robust conclusions, and the interaction test has insufficient power to detect meaningful differences between these subgroups even if such a difference exists.

Heterogeneity was substantial both in the 'prior to IVF' subgroup ($I^2 = 84\%$) and across subgroups ($I^2 = 83\%$). Evidence of both levels of heterogeneity was reduced to a minimal level ($I^2 = 0$) by the removal of Mahran *et al.*[20] This did not alter the conclusions of the analysis, which were consistent regardless of whether random- (see *Appendix 29*) or fixed-effects models were used (see *Appendix 30*).

Synthesis of secondary outcomes

Clinical pregnancy

Meta-analysis of all 12 trials that reported clinical pregnancy did not show evidence of a significant effect of ES in increasing clinical pregnancy rate when analysed using a random-effects model (see *Figure 18*; OR 1.18, 95% CI 0.82 to 1.72).[14,17–19,21,22,28,35,87–89] Owing to substantial evidence of heterogeneity ($I^2 = 82\%$), sensitivity analysis using a fixed-effects model was not undertaken.

Study or subgroup	ES Events	ES Total	Control Events	Control Total	Weight (%)	OR M–H, random, 95% CI	OR M–H, random, 95% CI
Eskew[21]	21	39	19	27	6.1	0.49 (0.17 to 1.39)	
Izquierdo Rodriguez[22]	48	77	44	78	8.6	1.28 (0.67 to 2.43)	
Karimzade[17]	9	77	26	79	7.3	0.27 (0.12 to 0.62)	
Lensen[14]	94	298	87	278	10.5	1.01 (0.71 to 1.44)	
Liu[18]	29	70	30	72	8.4	0.99 (0.51 to 1.93)	
Mackens[28]	35	70	33	78	8.6	1.36 (0.71 to 2.61)	
Maged[19]	61	150	41	150	9.7	1.82 (1.12 to 2.96)	
Mahran[20]	147	209	71	209	10.1	4.61 (3.05 to 6.96)	
Nastri[87]	3	8	4	10	2.9	0.90 (0.13 to 6.08)	
Polanski[88]	33	51	34	60	7.7	1.40 (0.65 to 3.02)	
Pye[35]	223	523	213	525	11.0	1.09 (0.85 to 1.39)	
Yeung[89]	42	105	38	104	9.2	1.16 (0.66 to 2.02)	
Total (95% CI)		1677		1670	100.0	1.18 (0.82 to 1.72)	
Total events	745		640				

Heterogeneity: $\tau^2 = 0.31$; $\chi^2 = 59.87$, df = 11 ($p < 0.00001$); $I^2 = 82\%$
Test for overall effect: $Z = 0.88$ ($p = 0.38$)

0.2   0.5   1.0   2.0   5.0
Favours control   Favours ES

FIGURE 18 Forest plot including all trials reporting clinical pregnancy rates. df, degrees of freedom; M–H, Mantel–Haenszel test.

Removal of Mahran et al.[20] reduced heterogeneity to a moderate level ($I^2 = 46\%$), but the conclusion drawn from the analysis did not change, which was consistent across random-effects models (Figure 19; OR 1.06, 95% CI 0.84 to 1.35) and fixed-effects models (see Appendix 31; OR 1.08, 95% CI 0.93 to 1.26).

We undertook a subgroup analysis to assess the effect of the timing of ES, including two trials that undertook ES during the IVF cycle,[17,28] and 10 trials undertaking ES prior to IVF (see Appendix 32).[14,18–22,35,87–89] The analysis identified a non-significant interaction between subgroups ($p = 0.36$), but with the point estimates (ORs) of the two subgroups suggesting that ES influenced LBR in different directions; the effect prior to IVF was 1.33 (95% CI 0.91 to 1.95) and during IVF it was 0.62 (95% CI 0.13 to 3.04). However, with only one trial in the 'during IVF' subgroup, there are insufficient data to draw robust conclusions, and the interaction test has insufficient power to detect meaningful differences between these subgroups even if such a difference exists. There was evidence of substantial heterogeneity within the subgroups themselves (prior to IVF: $I^2 = 89\%$; during IVF: $I^2 = 80\%$) and between trials ($I^2 = 82\%$).

To reduce heterogeneity in the 'prior to IVF' subgroup, Mahran et al.[20] was excluded from this analysis, resulting in no evidence of heterogeneity within the 'prior to IVF' subgroup ($I^2 = 0\%$), but substantial evidence of heterogeneity remained in the 'during IVF' subgroup ($I^2 = 89\%$) and moderate evidence of heterogeneity between trials ($I^2 = 46\%$) (see Appendix 33). This did not affect the conclusions of the analysis, identifying a non-significant interaction between subgroups ($p = 0.46$), but with the point estimates still suggesting that ES influenced LBR in two different directions, with an effect when ES was undertaken prior to IVF of 1.13 (95% CI 0.97 to 1.33), and the point estimate and CIs of the 'during IVF' subgroup remaining the same (OR 0.62, 95% CI 0.13 to 3.04). The conclusions were the same when a fixed-effects model was used (see Appendix 34).

Implantation per embryo transferred

One trial where the implantation rate was only reported as a percentage without the absolute numbers and authors did not respond to attempts to collect these data was excluded from this analysis.[20] Consequently, five trials were included in this review that reported this outcome.[18,19,22,35,89] However, issues with using implantation rate as an outcome have been documented.[91] Therefore, because of the sensitivity of implantation rate to the number of embryos transferred, and artificial inflation of a sample size when more than one embryo is transferred per woman, a meta-analysis of this outcome was not undertaken.

Study or subgroup	ES Events	ES Total	Control Events	Control Total	Weight (%)	OR M–H, random, 95% CI	OR M–H, random, 95% CI
Eskew[21]	21	39	19	27	4.3	0.49 (0.17 to 1.39)	
Izquierdo Rodriguez[22]	48	77	44	78	8.7	1.28 (0.67 to 2.43)	
Karimzade[17]	9	77	26	79	6.0	0.27 (0.12 to 0.62)	
Lensen[14]	94	298	87	278	15.3	1.01 (0.71 to 1.44)	
Liu[18]	29	70	30	72	8.3	0.99 (0.51 to 1.93)	
Mackens[28]	35	70	33	78	8.6	1.36 (0.71 to 2.61)	
Maged[19]	61	150	41	150	11.8	1.82 (1.12 to 2.96)	
Mahran[20]	223	523	213	525	18.5	1.09 (0.85 to 1.39)	
Nastri[87]	3	8	4	10	1.5	0.90 (0.13 to 6.08)	
Polanski[88]	33	51	34	60	6.8	1.40 (0.65 to 3.02)	
Yeung[89]	42	105	38	104	10.2	1.16 (0.66 to 2.02)	
Total (95% CI)		1468		1461	100.0	1.06 (0.84 to 1.35)	
Total events	598		569				

Heterogeneity: $\tau^2 = 0.07$; $\chi^2 = 18.67$, df = 10 ($p = 0.04$); $I^2 = 46\%$
Test for overall effect: $Z = 0.51$ ($p = 0.61$)

0.1 0.2 0.5 1.0 2.0 5.0 10.0
Favours control Favours ES

FIGURE 19 Forest plot of clinical pregnancy rates, with Mahran et al.[20] excluded. df, degrees of freedom; M–H, Mantel–Haenszel test.

Ongoing pregnancy rate

Ongoing pregnancy rate was reported in four studies,[14,17,22,89] with a meta-analysis of all four trials showing no evidence of a significant effect on undertaking ES on improving this outcome when a random-effects model was used (see *Appendix 35*; OR 0.86, 95% CI 0.49 to 1.48). There was evidence of substantial heterogeneity ($I^2 = 71\%$) and, therefore, sensitivity analysis using a fixed-effects model was not undertaken.

A subgroup analysis was undertaken to assess the impact of the timing of the ES on ongoing pregnancy rates (see *Appendix 36*), Karmizade *et al.*[17] was the only trial that reported ongoing pregnancy rates that undertook ES during the IVF cycle (OR 0.24, 95% CI 0.10 to 0.61), the three other trials reporting this outcome undertook the procedure in the menstrual cycle prior to ES (OR 1.08, 95% CI 0.82 to 1.42).[14,22,89] There was evidence of a significant difference in ongoing pregnancy rates between the subgroups ($p = 0.002$) and evidence of substantial heterogeneity between trials ($I^2 = 89\%$), but no evidence of heterogeneity within the 'prior to IVF' subgroup ($I^2 = 0\%$). However, as there was only one trial in the 'during IVF' subgroup, this conclusion should be interpreted with caution.

Miscarriage

Ten trials were included in this meta-analysis,[14,18–22,28,35,87,89] all undertaking ES in the menstrual cycle prior to IVF, which showed no evidence of a significant reduction when undertaking ES on this outcome when using a random-effects model (*Figure 20*; OR 0.96, 95% CI 0.57 to 1.63). There was evidence of moderate between-trial heterogeneity ($I^2 = 46\%$). This conclusion was consistent with sensitivity analysis results from a fixed-effects model (see *Appendix 37*). Removal of Mahran *et al.*[20] from this analysis did not substantially reduce the evidence of heterogeneity ($I^2 = 40\%$; see *Appendix 38*) and also did not alter the conclusions (OR 1.12, 95% CI 0.64 to 1.96).

However, this result should be interpreted with caution, as the method of calculating miscarriage rate varied substantially between studies. One trial collected this measure at 12 weeks,[19] six trials between 20 and 24 weeks,[14,20,28,35,87,89] and one trial collected both 'early' (prior to 12 weeks) and late (12–24 weeks) outcomes.[22] The timing of the collection of this outcome was unknown in two trials.[18,21] We therefore, undertook a subgroup analysis (*Figure 21*), separating those trials that collected miscarriages at an 'early' time point, and those that reported miscarriages at a 'late' time point. Izquierdo Rodriguez *et al.*[22] was included in both analyses by including both measures of miscarriage in the 'late' subgroup analysis (however, no late miscarriages were recorded). There was no evidence of a significant difference in miscarriage rates between subgroups when a random-effects model was used ($p = 0.50$), with no evidence of heterogeneity between pooled subgroup effects ($I^2 = 0\%$). This result was consistent when a fixed-effects model was used (see *Appendix 39*).

Study or subgroup	ES Events	ES Total	Control Events	Control Total	Weight (%)	OR M–H, random, 95% CI
Eskew[21]	4	21	0	19	2.8	10.03 (0.50 to 199.86)
Izquierdo Rodriguez[22]	5	48	8	44	10.9	0.52 (0.16 to 1.74)
Lensen[14]	4	85	5	80	9.5	0.74 (0.19 to 2.86)
Liu[18]	1	29	1	30	3.1	1.04 (0.06 to 17.38)
Mackens[28]	11	35	3	33	9.2	4.58 (1.15 to 18.31)
Maged[19]	11	61	9	41	13.4	0.78 (0.29 to 2.10)
Mahran[20]	13	147	14	71	15.8	0.39 (0.17 to 0.89)
Nastri[87]	1	3	0	4	2.0	5.40 (0.15 to 188.83)
Pye[35]	32	223	43	213	20.6	0.66 (0.40 to 1.09)
Yeung[89]	13	42	7	38	12.6	1.99 (0.70 to 5.67)
Total (95% CI)		694		573	100.0	0.96 (0.57 to 1.63)
Total events	95		90			

Heterogeneity: $\tau^2 = 0.28$; $\chi^2 = 16.93$, df = 9 ($p = 0.05$); $I^2 = 47\%$
Test for overall effect: $Z = 0.14$ ($p = 0.89$)

FIGURE 20 Forest plots of all trials reporting miscarriage rates. df, degrees of freedom; M–H, Mantel–Haenszel test.

Study or subgroup	ES Events	Total	Control Events	Total	Weight (%)	OR M–H, random, 95% CI	OR M–H, random, 95% CI
Early miscarriage							
Izquierdo Rodriguez[22]	5	48	8	44	10.1	0.52 (0.16 to 1.74)	
Maged[19]	11	61	9	41	12.7	0.78 (0.29 to 2.10)	
Subtotal (95% CI)		109		85	22.8	0.67 (0.31 to 1.43)	
Total events	16		17				
Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 0.26$, df = 1 ($p = 0.61$); $I^2 = 0\%$							
Test for overall effect: $Z = 1.05$ ($p = 0.29$)							
Late miscarriage							
Izquierdo Rodriguez[22]	5	48	8	44	10.1	0.52 (0.16 to 1.74)	
Lensen[14]	4	85	5	80	8.6	0.74 (0.19 to 2.86)	
Mackens[28]	11	35	3	33	8.3	4.58 (1.15 to 18.31)	
Mahran[20]	13	147	14	71	15.3	0.39 (0.17 to 0.89)	
Nastri[87]	1	3	0	4	1.7	5.40 (0.15 to 188.83)	
Pye[35]	32	223	43	213	21.2	0.66 (0.40 to 1.09)	
Yeung[89]	13	42	7	38	11.9	1.99 (0.70 to 5.67)	
Subtotal (95% CI)		583		483	77.2	0.94 (0.50 to 1.75)	
Total events	79		80				
Heterogeneity: $\tau^2 = 0.36$; $\chi^2 = 14.03$, df = 6 ($p = 0.03$); $I^2 = 57\%$							
Test for overall effect: $Z = 0.21$ ($p = 0.83$)							
Total (95% CI)		692		568	100.0	0.83 (0.52 to 1.35)	
Total events	95		97				
Heterogeneity: $\tau^2 = 0.21$; $\chi^2 = 14.41$, df = 8 ($p = 0.07$); $I^2 = 44\%$							
Test for overall effect: $Z = 0.74$ ($p = 0.46$)							
Test for subgroup differences: $\chi^2 = 0.46$, df = 1 ($p = 0.50$); $I^2 = 0\%$							

0.01 0.10 1.00 10.00 100.00
Favours ES   Favours control

FIGURE 21 Subgroup analysis of miscarriages, by early and late miscarriage definitions. df, degrees of freedom; M–H, Mantel–Haenszel test.

Multiple pregnancy and multiple birth rate

Evidence synthesis of multiple birth rates was not possible since only one trial reported multiple birth rates.[35] Meta-analysis of multiple pregnancy rates included five trials that reported data,[14,18,19,22,89] and results did not show a significant improvement of undertaking ES on this outcome (see *Appendix 40*; OR 1.09, 95% CI 0.68 to 1.75), with no evidence of heterogeneity between trials ($I^2 = 0\%$). This result was consistent with that from a fixed-effects model (see *Appendix 41*).

Ectopic pregnancy

Five studies reported ectopic pregnancy per clinical pregnancy;[14,18,20,28,35] however, only one trial defined this outcome.[34] Meta-analysis provided no evidence of a significant effect of ES on this outcome (see *Appendix 42*; OR 0.66, 95% CI 0.17 to 2.51), with no evidence of between-trial heterogeneity ($I^2 = 0\%$). However, the number of events reported was low, causing high uncertainty in the estimate of the overall effect. As a result, robust conclusions cannot be drawn regarding the effect of ES in reducing ectopic pregnancies. This result was consistent when a fixed-effects model was used (see *Appendix 43*). Interestingly, the smaller trials included in this analysis appeared to report more events than the larger trials. The cause of this is unknown, but may be due to different definitions of this outcome across trials.

Stillbirth

Owing to only two trials reporting this outcome and rare events reported [ES: *n* = 0, TAU: *n* = 1 (0.36%) in Lensen *et al.*,[14] and see *Chapter 7, The effects of endometrial scratch on the secondary outcomes* for Pye *et al.*[35]], meta-analysis was deemed to be inappropriate.

Pain

It was not possible to perform a meta-analysis on the effect of ES procedure on the pain experienced by participants as only three trials reported quantitative outcomes used to assess pain.[14,35,87] Two of the trials reported similar mean (SD) pain levels (on a rating scale of 0–10) directly post procedure [4.1 (2.4)][21] and during the procedure [4.2 (2.5)], whereas the third reported slightly higher pain levels

[6.42 (2.35)] directly after the procedure.[87] Pye *et al.*[35] further reported mean pain scores, which are presented in *Appendix 13*. Other trials reported pain qualitatively, with three trials reporting a low proportion of participants reporting severe pain [0/150 in Yeung *et al.*,[89] 3/209 (1.4%) in Mahran *et al.*,[20] 0/70 in Liu *et al.*[18] and 2/70 (2.9%) in Mackens *et al.*[28]].

Adverse events

In the trial participant/mother

Four trials did not report AEs.[21,22,87,88] Of the eight trials that reported AEs, seven appeared to have collected AEs in the intervention (ES) arm only, although information was often not presented in the reports. Four trials reported that there were 'no complications' recorded,[17–19,89] and one trial reported 'minimal' spotting for a few days post ES (number of participants not reported).[20] Trials that reported rare complications comprised one trial that reported that 1.3% ($n = 4$) of participants experienced fainting, 0.67% ($n = 2$) reported excessive pain and 0.33% ($n = 1$) excessive bleeding,[14] and another trial that reported that 5% ($n = 3$) of participants experienced bleeding.[28] Pye *et al.*[35] was the only study to report and compare safety events in both trial arms, finding comparable/similar incidences of AEs between groups (see *Chapter 7, Effect of endometrial scratch on safety outcomes*).

In the baby/neonate

Four trials reported AEs in the baby or neonate.[14,22,35,87] Separate data for participants undergoing their first IVF cycle were not requested for three of these trials that involved participants who had undergone an unselected number of previous IVF cycles.[14,22,87] However, the incidences of AEs in participants undergoing the first or subsequent cycle of IVF in these trials was low, with no 'major' fetal malformations reported in one trial;[87] rare congenital abnormalities (ES arm: 1.6%, $n = 3$; TAU arm: $n = 0$), fetal group restriction (ES: 5.2%, $n = 10$; TAU: 4.3%, $n = 8$) and neonatal death (ES: $n = 0$; TAU: 0.5%, $n = 1$) in another trial;[14] and rare placentation abnormalities (ES: 1.0%, $n = 1$; TAU: $n = 0$) and intrauterine growth restriction (ES: $n = 0$; TAU: $n = 1$, 1%) in the final trial.[22] Pye *et al.*[35] was the only trial that provided accessible data separately presented for those babies born to participants undergoing their first IVF cycle; they reported no neonatal deaths in both arms, no severe congenital abnormalities in the ES arm and low levels of congenital abnormalities in both trial arms (TAU 1.1%, $n = 3$; ES: $n = 0$) (see *Chapter 7* and *Table 21*).

Subgroup analyses

Analyses to assess the effects of the prespecified subgroups (age, duration of infertility and IVF protocol) could not be undertaken, as the included trials did not present outcome data individually for these subgroups.

Conclusions

Combining data across trials did not provide evidence to support the beneficial effect of ES on LBR, clinical pregnancy rate, ongoing pregnancy rate, miscarriage rate or multiple pregnancy rate. One trial contributed substantial heterogeneity to these analyses, which may have been as a result of this trial requiring participants to have two 'good' embryos transferred and providing hysteroscopy to all participants. The former may have resulted in participants being excluded from the trial post randomisation, as investigators will have had to predict which participants were going to produce good-quality embryos prior to IVF starting. When this trial was removed from the LBR and clinical pregnancy analyses, heterogeneity was reduced, and results were consistent regardless of whether fixed- or random-effects models were used. Five trials reported implantation rate, but a meta-analysis was not undertaken because of inherent biases in the measurement of this outcome. Only five trials reported ectopic pregnancy rates and, for those trials that did, there were very few events reported and, surprisingly,

smaller trials reported more events than larger trials. This is difficult to explain, but may be attributable to trials using different definitions for this outcome. Only two of the trials reported stillbirth rates; owing to a very small number of events, a meta-analysis was not appropriate. One trial reported preterm delivery rates. Few of the trials reported pain post ES ($n = 6$); however, four trials reported a low proportion of participants experiencing 'severe' pain post ES, and three trials reported similar, moderate post-ES pain ratings. The procedure seemed to be safe, with studies reporting very rare complications post ES. None of the included trials was deemed to be at a high risk of bias.

Chapter 11 Discussion

We undertook a pragmatic RCT of the ES procedure followed by 'usual' IVF treatment compared with 'usual' IVF treatment with the objectives of assessing its clinical effectiveness, safety and cost-effectiveness. In addition, we undertook a qualitative substudy to assess the experiences of trial participants and staff, and a systematic review, including a meta-analysis, to incorporate the results of this trial with others undertaken in the same population.

Statement of main findings

Clinical effectiveness

We did not find evidence that performing the ES procedure before IVF increases the chance of achieving pregnancy and a live birth in women undergoing IVF for the first time. The observed LBRs were similar between treatment groups with 37.1% in the TAU group and 38.6% in the ES group; a very small increase in LBR of 1.5% (95% CI −4.4% to 7.4%; $p = 0.621$) was attributed to the ES procedure, which is not statistically significant and not of clinical importance. Sensitivity analyses under different assumptions concerning missing data and adjusting for prespecified covariates produced similar results. In the per-protocol analysis, ES increased LBR by only 3.6% (95% CI −3.3% to 10.5%; $p = 0.312$) and there was only a 2.9% probability that ES increases the LBR by at least the 10% targeted AD, given the observed data; therefore, this observed effect was not statistically significant and extremely unlikely to be of clinical importance.

Across secondary outcomes, we did not find evidence to support the beneficial effects of the ES procedure; rates were similar between groups. These secondary outcomes were implantation, clinical pregnancy, multiple birth, preterm delivery, miscarriage, ectopic pregnancy and stillbirth.

In general, the effect of ES on pregnancy and live birth appears consistent across prespecified subgroups, including protocol followed (long or antagonist), fertilisation method (IVF, ICSI or split IVF and ICSI), nature of embryo transfer (single or double embryo transfer), embryo transferred (fresh or frozen) and history of miscarriage (0–2 or ≥ 3). However, we could not rule out potential benefits when embryos are transferred on day 5, or in women who had cycle programming performed using oral contraception, progestogens or oral oestrogen. Readers should interpret these exploratory results with caution owing to the small numbers of participants and events within each subgroup.

Safety and tolerability

The ES procedure was found to be well tolerated, with high levels of reported tolerability (99.8%) and low pain levels, with a median pain score (IQR) of 4.0 (2.0–6.0) within 30 minutes post procedure, 1.0 (0.0–3.0) at 1 day and 0.0 (0.0–0.0) at 7 days post procedure. ES was found to be safe; rates of clinical outcomes that also relate to safety (ectopic pregnancy, miscarriages, preterm delivery, multiple births and stillbirth) and the incidences of AEs and SAEs were comparable between groups. There were slightly lower rates of babies born with low birthweight, very low birthweight, or who were small for their gestational age in the ES group than in the TAU group. However, these results should be interpreted with caution, owing to these events being highly correlated, based on small numbers and not adjusted for other potential confounding factors.

No deaths or severe congenital abnormalities were recorded. Only 2.4% (11/458) of women experienced at least one unexpected AE that occurred between ES and IVF (in the ES group).

Health economics

The cost-effectiveness analysis found that ES was, on average, £316 more expensive (£6819.38) than TAU (£6503). The ICER per successful live birth was calculated at £11.90 per woman (95% CI −£134

to £127), meaning that in order to achieve one extra live birth in 1000 women £11,900 would need to be spent. The ICER was similar across secondary outcomes, prespecified subgroups and sensitivity analyses.

Meta-analyses

To corroborate evidence of clinical effectiveness, we undertook a systematic review combining the results of this trial with other trials assessing the effectiveness of ES prior to the first IVF cycle, using meta-analysis. One trial contributed substantial heterogeneity to the analyses, possibly owing to an eligibility criteria stipulating that participants had to have two good-quality embryos transferred.[20] When this trial was excluded from the analyses, heterogeneity was reduced to minimal or moderate levels, finding no evidence that ES had a significant effect on LBR (nine trials: OR 1.03, 95% CI 0.87 to 1.22) and pregnancy rate (11 trials: OR 1.06, 95% CI 0.84 to 1.35), using random-effects models. We found no evidence of ES having a significant effect on ongoing pregnancy rate (four trials: OR 0.86, 95% CI 0.49 to 1.48), miscarriage rate (10 trials: OR 0.96, 95% CI 0.57 to 1.63) or multiple pregnancy rate (five trials: OR 1.09, 95% CI 0.68 to 1.74), also using random-effects models. Owing to a small number of trials ($n = 5$) reporting ectopic pregnancy rates, and the small number of events these trials reported, the results of a meta-analysis of ectopic pregnancy rates was highly uncertain (OR 0.66, 95% CI 0.17 to 2.51). Results were consistent across all outcomes when fixed-effects models were used. We did not undertake a meta-analysis of implantation rates because of concerns around this outcome being biased by the number of embryos transferred, which has been highlighted elsewhere.[91] Only two trials reported stillbirth rate, and one reported preterm delivery rate; therefore, a meta-analyses was not undertaken for these outcomes.

Only six of the trials reported pain post procedure;[14,18,20,28,35,89] in those that did, few participants reported severe pain, with three studies reporting a numerical measure of pain, all three of which reported moderate pain post procedure.[14,35,87] Eight of the trials reported AEs;[14,17–20,28,35,89] however, only our trial recorded such events in both ES and TAU arms of the trial, limiting the conclusions that can be made. Four trials reported AEs in the fetus or neonate;[14,22,35,87] however, data were not obtained for three of these,[14,22,87] also limiting the conclusions that can be drawn.

Qualitative substudy

To inform the design and conduct of future, related trials, we undertook a qualitative substudy to explore trial participants' and site staff members' experiences of being involved in the trial. Of the small proportion of trial participants ($n = 27$) and site staff ($n = 7$) who took part in the qualitative interviews, both participants and staff were generally happy with the recruitment process, and participants were well prepared to receive the ES procedure. Three participants discussed recruitment being too informal. Some participants and site staff felt that more information was required regarding the evidence base for or against ES. Eighteen of the interviewed participants discussed having positive preconceptions regarding the effect of the ES on the outcome of their IVF cycle, with site staff, in some instances, appearing to have contributed to these positive preconceptions. Some participants seemed to be unaware of the possibility that the ES could potentially reduce their chances of a live birth, instead choosing to focus on the potential positive impact that the intervention may have.

Participants' positive preconceptions meant that the recruitment process was challenging for both staff and participants owing to some participants experiencing demoralisation when they were randomised to the control arm. Site staff developed mechanisms to assist participants in coming to terms with this. The potential for participation in the trial to delay the IVF cycle was also challenging. Five participants described the procedure as being more painful than they expected. Participants therefore recommended that more information regarding the potential painfulness of the ES is provided in trial recruitment materials.

Those participants who declined the ES procedure did so, on the whole, because of personal reasons that could not have been prevented (e.g. holidays and illness). However, the withdrawal of two participants could

have been prevented with improved organisation by the fertility unit; a smear test was missed in one participant, and the other was not told to abstain from sexual intercourse prior to ES.

Strengths and weaknesses of the research

Strengths

This well-powered, pragmatic RCT recruited the targeted number of women undergoing IVF across 16 UK fertility units in both the private and the public sectors. To the best of our knowledge,[16,92] this is the largest definitive trial evaluating the effect of ES procedure in women undergoing first-time IVF, thus bridging an important evidence gap on the benefits and safety of IVF 'add-ons'.[2,8,93,94] Furthermore, the observed levels of clinical effectiveness and safety of the ES procedure are likely to be reflected in clinical practice because of the pragmatic nature of the trial,[95] which mimicked the real world by recruiting participants from both NHS and private facilities across the UK, including both long and antagonist protocols, and by allowing pragmatic delivery of the ES.

Another major strength of this study is the specific focus on one particular population in order to minimise heterogeneity. The potential confounding effect of embryo quality was addressed by including only good responders who were likely to have a single embryo replacement. Indeed, our results show that the majority of participants responded well to stimulation and approximately 80% received a SET on day 5. We excluded patients with factors that could potentially influence endometrial quality (anovulation, BMI $\geq$ 35kg/m^2 and severe endometriosis).

There was a very high uptake (86.6%) of the ES procedure as per protocol in the ES group. The proportion of women who received IVF in both treatment groups was very high and similar (95.0% in the ES group and 94.1% in the TAU group). The contamination in the TAU group was negligible with only 1% (5/525) of women receiving ES procedure outside the trial, indicating high compliance and negligible contamination of the TAU group.

Last, we undertook a meta-analysis of all the available evidence within RCTs, including the results from our trial, which highlighted that, to our knowledge, this is the only trial that has comprehensively reported AE data both in participating women and in born babies.

Weaknesses

The study did not include a sham procedure in the control group. However, we do not believe that this would have influenced our results, as the objective fertility outcomes used in this study are unlikely to be influenced by a placebo effect.

Nearly 10% (9.2%; 48/523) of women who were randomised to the ES group did not receive the ES procedure, but they subsequently received IVF. These women technically received TAU although they were randomised to receive the ES procedure. As a result, this could have potentially diluted the effect of the ES procedure under the ITT principle. However, post hoc analysis that excluded women who did not receive the allocated treatment (ES plus IVF in the ES group or IVF in the TAU group) or those who were allocated to TAU but received the ES procedure outside the trial produced similar results to the main analysis, showing a non-significant effect of ES on live birth.

For ethical reasons, and under the advice from the PPI group, women who did not become pregnant or experienced a negative outcome, such as a loss of pregnancy, were not followed up further.[35] As a result, safety outcomes after these events were not collected in these participants, which may have resulted in some important events being missed, and the impact of interventions underestimated. Readers should bear this in mind when interpreting safety results.

A small proportion (0.9%; 10/1048) of women were found to be ineligible following randomisation, although these individuals remained in the trial and were followed up. Reasons for ineligibility included the identification of medical/gynaecological conditions (e.g. fibroids), slightly high BMI or age. However, this pragmatic trial mimicked routine practice where such issues occur and these numbers were very negligible relative to those randomised. Furthermore, sensitivity analyses that included per-protocol analysis were performed that excluded these women, and this produced consistent results.

The qualitative substudy included only participants who had participated in the trial, and, therefore, participants' reflections on their experience of the recruitment process and delivery of the ES procedure may have been biased by the fact that they had agreed to participate (also known as survivorship bias). Those that withdrew from the trial were not approached for interview, meaning that the experience of those who could not tolerate the trial protocol, or who experienced particular demoralisation when randomised to the TAU arm, may have been under-represented. Interviews with fertility unit staff were undertaken by members of the central research team. The staff would have been aware of the interviewers' involvement in the trial, which may have also biased how they reported their experience.

Owing to time constraints, we were unable to contact all authors within the systematic review to collect information to aid the assessment of the risk of bias. For one trial included in the review,[88] we were unable to obtain the full-text article or correspond with the author in order to obtain data and, therefore, data for this trial were extracted from a recent systematic review.[16] As a result, a risk-of-bias assessment for this trial could not be undertaken.

In order to compare the quality of embryos between fertility units, we developed a method of converting between the three embryo grading techniques used across the trial. However, this was developed by one principal embryologist, and therefore consensus was not sought between fertility units. This method of comparing embryo quality was not validated and therefore may be open to bias.

Evidence in the context of other similar studies

This study provides definitive evidence that the ES does not significantly improve the chances of a successful outcome of IVF in women undergoing their first cycle. Previous studies have been undertaken in 'unselected' populations with significant sample heterogeneity, or have been small studies with other major methodological flaws.[16]

Studies undertaken in unselected populations have not been specifically powered to detect a realistic and clinically worthwhile difference in the first-time IVF group.[14,21,87,89] A recent study by Lensen et al.[14] included a cohort of participants undergoing their first IVF cycle, but was not powered specifically to detect a successful outcome in this group. Furthermore, significant heterogeneity was introduced by including populations with different prognostic potential and a high proportion (26%) of FETs. Unpublished data obtained from the authors for those patients undergoing their first cycle of IVF showed no significant difference between the ES and TAU arms (RR 0.99, 95% CI 0.90 to 1.10).

Other studies have been undertaken in women undergoing their first IVF cycle but have had relatively small sample sizes and have suffered from other major methodological flaws.[17,18,96] Two of the largest RCTs to date in patients receiving first-time IVF included 418 and 300 participants, respectively.[19,20] Mahran et al.[20] identified significant increases in live birth, implantation and pregnancy rates, and significant decreases in miscarriage rates in those that received ES, while Maged et al.[19] identified significant increases in pregnancy and implantation rates in those that received ES. However, a recently conducted systematic review[16] identified potential sources of bias in these studies, with either an unclear or high risk of reporting bias, detection bias and other biases in Mahran et al.,[20] or unclear risk of bias in allocation concealment, and other biases, in Maged et al.[19,20] In addition, Maged et al.[19] did not follow up participants until delivery.

Mahran *et al.*[20] reported very high LBRs in the ES arm (67% in the ES arm and 28% in the control arm), and included only patients who had two or more 'good quality' embryos replaced. Although a subgroup analysis for embryo quality was not prespecified or undertaken within our trial, the general quality of embryos transferred was good, with at least 70.3% (646/919) women receiving at least one embryo transfer of excellent, very good or good quality. Therefore, it is extremely unlikely that the lack of positive effect of the ES procedure that we observed in this trial can be attributed to poor embryo quality.

To our knowledge, no study has qualitatively explored the experience of patients and staff taking part in and delivering an interventional fertility trial. We found that some participants felt discontent at being randomised to the control arm, which has been discussed in numerous previous studies in other populations.[97–99] However, only a small proportion of participants actively withdrew from the ES arm of the trial ($n = 8$), and only five participants in the TAU arm received ES 'outside' the trial, leading us to conclude that, although participants may have felt demoralised when randomised to TAU, this did not result in a high proportion of participants withdrawing from the trial or seeking ES elsewhere. Participants required more information, at recruitment, regarding the evidence base for and against ES. However, including such information would be at odds with some participants in the qualitative study reporting that the amount of information received prior to recruitment was 'overwhelming', and other research that suggests that participants prefer shorter versions of the PIS and that some participants only minimally use the PIS prior to consent.[100,101] Other studies have also reported that participants overestimate the benefits, and underestimate the risks, of participating in research;[102,103] however, to our knowledge, this is the first time that this has been described in those participating in a fertility trial. In this trial, underestimation of the potential harms may have been exacerbated by the time-sensitive nature of fertility treatment; both clinicians and patients have been reported to feel the need to improve the outcomes of treatment immediately following a failed treatment cycle.[104]

In addition, no study has looked at the cost-effectiveness of ES, or any other IVF 'add-on'. Exploring cost-effectiveness is particularly important in the UK in order to understand whether or not new effective treatments can be afforded by the NHS. A study in the Netherlands looked at the cost-effectiveness of immediate compared with delayed IVF over a 2-year time horizon using a societal perspective.[105] They reported mean costs (converted to Great British pounds from Euros, and inflated to 2018/19 prices using inflation rates reported by Curtis and Burns[106]) of IVF of £6485 for unexplained fertility and £6480 for endometriosis when IVF was postponed for 1 year compared with costs of £7046 for unexplained fertility and £6669 for endometriosis when IVF was not delayed for those aged 30 years.[105] This study reported mean results over a longer time horizon (2 years) and a different population, so is not comparable with our results.

Discussion of results

The ES procedure performed in the mid-luteal phase of the menstrual cycle preceding IVF was found to be well tolerated and safe in women undergoing their first IVF cycle. These results are consistent with other similar trials that also reported very rare AEs.[14,19,20,28,89] However, we did not find evidence to support the beneficial use of ES procedure in improving the chances of achieving a live birth and other pregnancy outcomes in this population. These results were similar across the sensitivity analyses performed. This lack of evidence to support beneficial effects of ES across pregnancy outcomes was consistent with our meta-analysis, which pooled together results from similar RCTs,[14,17–22,28,35,87,89] including this trial. It should be noted that there was substantial study-to-study heterogeneity mainly driven by one trial,[20] potentially due to the selection of participants who were predicted to have two 'good' quality embryos transferred. However, overall results were consistent when this trial was excluded from the meta-analyses for LBRs and clinical pregnancy rates.

The timing of the ES within the IVF cycle is an important consideration. A small proportion of previous trials have undertaken the ES at time points in the menstrual cycle other than the mid-luteal phase: on the day of egg collection,[17] day 6 of stimulation,[96] and IVF cycle day 10–12.[18] However, in this trial we undertook ES in the mid-luteal phase because, when we sought funding, evidence was available that suggested that there were potential benefits.[12] Although those trials that have undertaken ES at time points other than the mid-luteal phase do not show a benefit of ES at this time point,[17,18,96] we cannot ascertain whether or not undertaking the ES procedure at the mid-luteal phase of the menstrual cycle preceding IVF is one of the contributing factors to the results observed in this trial.

There was no evidence to show that ES procedure was more costly than usual care, although on average ES was £316 more costly than TAU, with the 95% CI ranging from –£468 to £1078. However, the ICER per successful live birth of £11.90 (95% CI –£134 to £127) was relatively low, although unlike QALYs, which have a well-defined threshold, the interpretation of the ICER per successful live birth is complicated by the lack of a set threshold for this measure. Based on the health economic evidence alone, a decision-maker may decide to fund ES in this population; however, the clinical effectiveness results, which showed that ES did not result in a significant increase in LBRs and benefits on other pregnancy outcomes compared with TAU, should be considered alongside this finding. Therefore, it is unlikely that ES would benefit patients undergoing their first IVF cycle. In a qualitative substudy, five participants reported significant levels of pain post ES that limited the activities they could undertake in the days following the procedure. These findings are not contradictory, as even though trial participants reported, on a pain rating scale of 0 (no pain) to 10 (worst pain imaginable), a mean (SD) pain score of 4.1 (2.4) within 30 minutes of the procedure and 1.8 (2.5) at 1 day and 0.8 (2.1) at 7 days post ES, some participants did report high post-procedure pain scores at all three time points (see *Appendix 13*). It may be that, by chance, those participants who rated high pain levels were interviewed in the qualitative study. Nevertheless, we can conclude that the experience of receiving the ES is variable, with severe pain being felt by some participants for days after the procedure, but the majority experiencing milder, short-term pain.

Participants' experience of pain post ES feeds into our interpretation of the health economic evaluation when it focuses on a wider than NHS perspective to include costs incurred by women and their companions. Within the trial, we collected data regarding participants' health resource use and the costs they incurred in receiving IVF treatment, but we did not collect information regarding lost earnings or lost productivity other than those for attending fertility appointments. If the loss of productivity post ES due to the pain of the procedure or other side effects was incorporated into the analysis, then including a wider than NHS perspective may increase the difference in costs between ES and TAU and increase the ICER, although the ICER may remain relatively low.

Implications for health care

Despite a lack of evidence to support its use, some in the medical community have adopted ES, including for those undergoing their first IVF cycle.[27] Evidence from this trial concludes that the ES procedure does not significantly improve pregnancy outcomes for women undergoing their first IVF cycle. Furthermore, an updated meta-analysis of trials in this population, including unpublished data obtained from study authors, demonstrates a consistent overall effect, across all included trials and outcomes, of no benefit.

Implications for future research studies

Recruiting staff should be aware that participants may base their decision to participate on internet research, which may result in positive preconceptions of the potential benefits and a downplaying of the potential harms of receiving the novel intervention. Recruiting staff should try to introduce inclusion in the control group with similar emphasis to inclusion in the intervention group; however, once randomised to the control group discontent may still be felt, in which case we suggest three ways to lessen this, namely, explaining the importance of the control group, the lack of evidence to support the intervention, and the increased contact with patients that participating in research allows.

Other studies have provided recommendations to guide 'recruitment conversations' in surgical trials; these guidelines are also applicable to studies in fertility.[107–109]

The quantity of information provided to participants during the recruitment stages of fertility studies may need to be revised. Other methods of providing such information to potential participants should be explored, possibly by providing participants with a choice of levels of information, as suggested in previous studies.[101,110]

Qualitative and quantitative evaluations of the experience of receiving ES highlighted the variability in pain experienced by patients, with some participants experiencing more severe pain post procedure. This information should be taken into account when undertaking ES in other populations.

Implications for future research needs

Our trial provides definitive evidence that ES does not significantly improve pregnancy outcomes in the first IVF cycle when the procedure is performed in the mid-luteal phase. However, three pertinent questions remain.

First, as described above, a small number of trials have undertaken ES at time points other than the mid-luteal phase.[17,18,96] It could be argued that further investigation is warranted within a high-quality RCT. However, these trials concluded that ES had no benefit on IVF outcomes, with one trial identifying a reduction in pregnancy rates when the ES was undertaken on the day of egg collection.[17] Given that smaller, less well conducted trials are thought to overestimate treatment effects,[111] we argue that further trials for this reason are unwarranted.

Second, we found that babies born to women randomised to the ES group had a slightly lower chance of being born small for their gestational age (AD –4.4%, 95% CI –8.1% to –0.7%) or with a low birthweight (AD –8.3%, 95% CI –13.5% to –3.2%) or very low birthweight (AD –3.9%, 95% CI –6.9% to –0.8%). Given the low chance of these events occurring and also that they are correlated, this result should be interpreted as an exploratory finding, and we do not recommend future research in this area owing to the low chance of significant findings in a larger study. We also found that a subgroup analysis identified that a benefit could not be ruled out in women who had day 5 embryo transfers or those who underwent cycle programming. Again, we suggest that further research is unwarranted. The fact that the majority of participants had a day 5 embryo transferred is merely proof that we selected a homogeneous population of good responders. This means that the embryo effect is minimal and the focus is on the endometrium on which the scratch should act if it had an effect; however, our secondary outcomes found no effect on implantation or pregnancy rates.

A meta-analysis of individual participant data (IPD), where individual patient records from studies in this population are aggregated and analysed, is planned by Farquhar and colleagues (Professor Cindy Farquhar, The University of Auckland, 2021, personal communication). We recommend that further research is not undertaken until this IPD has been completed, as doing so may waste resources in evaluating an intervention that two recent, large RCTs (this study, in first-cycle IVF, and the PiP study,[14] in unselected IVF cycles) have concluded as ineffective.[14] The IPD should aim to elucidate the two pertinent remaining questions on this topic, described above.

Our systematic review has highlighted that studies undertaken in this population have defined key outcomes differently, with trials missing key outcomes and the time point of assessment differing. This issue has previously been highlighted by Wilkinson et al.[112] The development of a core data set for studies in women undergoing infertility treatments is in progress, but the outcome has not yet been published.[113] We recommend that this core data set is utilised in order to reduce inefficiencies and enable useful comparisons to be made in future meta-analyses.

Future research should focus on how the potential harms of taking part in a research study can be relayed to participants, in order to deal with overly positive preconceptions and the impact of these preconceptions on feeling discontent when randomised to the control arm of the trial. The PIS for the Endometrial Scratch Trial did not explicitly state the evidence for or against the ES procedure, nor did it explicitly state that the ES may potentially harm the chances of a successful IVF outcome. Including this information in the PIS is at odds with participants in this study stating that the information provided to them was 'overwhelming', and other research that suggests that participants prefer shorter information sheets[101] and that participants only minimally use the PIS.[100] However, future research should seek trial participants' views on how evidence regarding the potential harms of a procedure, or at least the unknowns, should be imparted, including the medium (i.e. PIS or face to face) and how the information should be structured in order to not unnecessarily perturb potential participants about participating in research. In addition, patient-acceptable alternatives to the patient information sheet should be investigated, in order to reduce the burden on patients at the start of their fertility treatment.

Finally, we do not recommend any future research is undertaken regarding the cost-effectiveness of ES, as, given the ES was found to not significantly impact on LBRs, further research in this area is unwarranted.

Conclusions

In conclusion, this study provides evidence that, although safe and tolerable, ES does not significantly improve pregnancy outcomes in women undergoing their first IVF cycle. A meta-analysis that combined our trial with other randomised trials in this population also showed no benefit of undertaking ES in the first cycle of IVF, when assessed using the same outcomes. The health economic evaluation, although showing that ES has the potential to be cost-effective at the population level, should be considered in context alongside there being no clinical benefits to improve clinical pregnancy.

Acknowledgements

Contributions of authors

Lamiya Mohiyiddeen (https://orcid.org/0000-0002-4616-5893) (Consultant Gynaecologist and Subspecialist in Reproductive Medicine and Surgery; reproductive medicine) contributed to the design and oversight of the trial, was a member of the TMG, contributed to the interpretation of the results, provided advice on clinical aspects of the trial, was a site investigator and reviewed the report for important intellectual content.

Tim Chater (https://orcid.org/0000-0002-1138-0147) (Data and Information Systems Manager; data management) contributed to the design and management of the trial, oversaw data management, was a member of the TMG and reviewed the report for important intellectual content.

Kirsty Pemberton (https://orcid.org/0000-0002-3666-2386) (Data Manager; data management) contributed to the design and management of the trial, undertook data management, was a member of the TMG and reviewed the report for important intellectual content.

Chris Turtle (https://orcid.org/0000-0003-3109-425X) (Data Specialist; data management) undertook data management activities, was a member of the TMG and reviewed the report for important intellectual content.

Jamie Hall (https://orcid.org/0000-0003-4042-5591) (Research Assistant; trial management) contributed to the systematic review and reviewed the report for important intellectual content.

Liz Taylor (https://orcid.org/0000-0002-4769-3970) (Research Nurse; assisted conception) contributed to the design and management of the trial and interpretation of the results, was a member of the TMG, was a site RN who contributed to recruitment to the trial, and reviewed the report for important intellectual content.

Kate Brian (https://orcid.org/0000-0002-1583-0416) (PPI representative: PPI) contributed to the design and management of the trial and interpretation of the results, provided PPI input and reviewed the report for important intellectual content.

Anya Sizer (https://orcid.org/0000-0002-6631-1085) (PPI representative; PPI) contributed to the design and management of the trial and interpretation of the results, was a member of the TMG provided PPI input and reviewed the report for important intellectual content.

Helen Hunter (https://orcid.org/0000-0002-1506-2410) (Principal Embryologist; embryology) contributed to embryo grading and the comparison of embryo grading between grading systems, and reviewed the report for important intellectual content.

Publications

Pye C, Chatters R, Cohen J, Brian K, Cheong YC, Laird S, *et al.* Induced endometrial trauma (endometrial scratch) in the mid-luteal menstrual cycle phase preceding first cycle IVF versus usual IVF therapy: study protocol for a randomised controlled trial. *BMJ Open* 2018;**8**:e020755. https://doi.org/10.1136/bmjopen-2017-020755

Chatters R, White D, Pye C, Petrovic A, Sizer A, Kumar P, Metwally M. Experiences of trial participants and site staff of participating in and running a large randomised trial within fertility (the endometrial scratch trial): a qualitative interview study. *BMJ Open* 2021;**11**:e051698. https://doi.org/10.1136/bmjopen-2021-051698

Metwally M, Chatters R, Dimairo M, Walters S, Pye C, White D, *et al.* A randomised controlled trial to assess the clinical effectiveness and safety of the endometrial scratch procedure prior to first-time IVF, with or without ICSI. *Hum Reprod* 2021;**36**:1841–53. https://doi.org/10.1093/humrep/deab041

Metwally M, Chatters R, White D, Hall J, Walters S. Endometrial scratch in women undergoing first time IVF treatment: a systematic review and meta-analysis of randomised controlled trials. *Reprod Biomed Online* 2022; in press.

Data-sharing statement

Requests for patient-level data and statistical code should be made to the corresponding author and will be considered on a case-by-case basis following the principles for sharing patient-level data as described by Smith *et al.*[114] The presented data do not contain any direct identifiers; we will minimise indirect identifiers and remove free-text data to minimise the risk of identification.

Patient data

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease, develop new treatment, monitor safety, and plan NHS services. Patient data should be kept safe and secure to protect everyone's privacy, and it's important there are safeguards to make sure that it is stored and used responsibly. Everyone should be able to find out how patient data are used. #datasaveslives. You can find out more about the background to this citation here: https://understandingpatientdata. org.uk/data-citation.

References

1. Human Fertilisation and Embryology Authority (HFEA). *Fertility Treatment 2018: Trends and Figures*. 2020. URL: www.hfea.gov.uk/about-us/publications/research-and-data/fertility-treatment-2018-trends-and-figures/ (accessed 12 March 2021).

2. Macklon NS, Ahuja KK, Fauser B. Building an evidence base for IVF 'add-ons'. *Reprod Biomed Online* 2019;**38**:853–6. https://doi.org/10.1016/j.rbmo.2019.04.005

3. Bouwmans CA, Lintsen BA, Al M, Verhaak CM, Eijkemans RJ, Habbema JD, *et al.* Absence from work and emotional stress in women undergoing IVF or ICSI: an analysis of IVF-related absence from work in women and the contribution of general and emotional factors. *Acta Obstet Gynecol Scand* 2008;**87**:1169–75. https://doi.org/10.1080/00016340802460305

4. Cousineau TM, Domar AD. Psychological impact of infertility. *Best Pract Res Clin Obstet Gynaecol* 2007;**21**:293–308. https://doi.org/10.1016/j.bpobgyn.2006.12.003

5. Verhaak CM, Smeenk JM, Evers AW, Kremer JA, Kraaimaat FW, Braat DD. Women's emotional adjustment to IVF: a systematic review of 25 years of research. *Hum Reprod Update* 2007;**13**:27–36. https://doi.org/10.1093/humupd/dml040

6. Wade JJ, MacLachlan V, Kovacs G. The success rate of IVF has significantly improved over the last decade. *Aust N Z J Obstet Gynaecol* 2015;**55**:473–6. https://doi.org/10.1111/ajo.12356

7. Wise J. Show patients evidence for treatment 'add-ons', fertility clinics are told. *BMJ* 2019;**364**:l226. https://doi.org/10.1136/bmj.l226

8. Heneghan C, Spencer EA, Bobrovitz N, Collins DR, Nunan D, Plüddemann A, *et al.* Lack of evidence for interventions offered in UK fertility centres. *BMJ* 2016;**355**:i6295. https://doi.org/10.1136/bmj.i6295

9. Barash A, Dekel N, Fieldust S, Segal I, Schechtman E, Granot I. Local injury to the endometrium doubles the incidence of successful pregnancies in patients undergoing in vitro fertilization. *Fertil Steril* 2003;**79**:1317–22. https://doi.org/10.1016/S0015-0282(03)00345-5

10. Potdar N, Gelbaya T, Nardo LG. Endometrial injury to overcome recurrent embryo implantation failure: a systematic review and meta-analysis. *Reprod Biomed Online* 2012;**25**:561–71. https://doi.org/10.1016/j.rbmo.2012.08.005

11. Lensen S, Venetis C, Ng EHY, Young SL, Vitagliano A, Macklon NS, *et al.* Should we stop offering endometrial scratching prior to in vitro fertilization? *Fertil Steril* 2019;**111**:1094–101. https://doi.org/10.1016/j.fertnstert.2019.04.017

12. Nastri CO, Gibreel A, Raine-Fenning N, Maheshwari A, Ferriani RA, Bhattacharya S, Martins WP. Endometrial injury in women undergoing assisted reproductive techniques. *Cochrane Database Syst Rev* 2012;**7**:CD009517. https://doi.org/10.1002/14651858.CD009517.pub2

13. van Hoogenhuijze NE, Kasius JC, Broekmans FJM, Bosteels J, Torrance HL. Endometrial scratching prior to IVF; does it help and for whom? A systematic review and meta-analysis. *Hum Reprod Open* 2019;**2019**:hoy025. https://doi.org/10.1093/hropen/hoy025

14. Lensen S, Osavlyuk D, Armstrong S, Stadelmann C, Hennes A, Napier E, *et al.* A randomized trial of endometrial scratching before in vitro fertilization. *N Engl J Med* 2019;**380**:325–34. https://doi.org/10.1056/NEJMoa1808737

15. McLernon DJ, Steyerberg EW, Te Velde ER, Lee AJ, Bhattacharya S. Predicting the chances of a live birth after one or more complete cycles of in vitro fertilisation: population based study of linked cycle data from 113 873 women. *BMJ* 2016;**355**:i5735. https://doi.org/10.1136/bmj.i5735

16. Vitagliano A, Andrisani A, Alviggi C, Vitale SG, Valenti G, Sapia F, *et al.* Endometrial scratching for infertile women undergoing a first embryo transfer: a systematic review and meta-analysis of published and unpublished data from randomized controlled trials. *Fertil Steril* 2019;**111**:734–46. https://doi.org/10.1016/j.fertnstert.2018.12.008

17. Karimzade MA, Oskouian H, Ahmadi S, Oskouian L. Local injury to the endometrium on the day of oocyte retrieval has a negative impact on implantation in assisted reproductive cycles: a randomized controlled trial. *Arch Gynecol Obstet* 2010;**281**:499–503. https://doi.org/10.1007/s00404-009-1166-1

18. Liu W, Tal R, Chao H, Liu M, Liu Y. Effect of local endometrial injury in proliferative vs. luteal phase on IVF outcomes in unselected subfertile women undergoing in vitro fertilization. *Reprod Biol Endocrinol* 2017;**15**:75. https://doi.org/10.1186/s12958-017-0296-8

19. Maged AM, Rashwan H, AbdelAziz S, Ramadan W, Mostafa WAI, Metwally AA, Katta M. Randomized controlled trial of the effect of endometrial injury on implantation and clinical pregnancy rates during the first ICSI cycle. *Int J Gynaecol Obstet* 2018;**140**:211–16. https://doi.org/10.1002/ijgo.12355

20. Mahran A, Ibrahim M, Bahaa H. The effect of endometrial injury on first cycle IVF/ICSI outcome: a randomized controlled trial. *Int J Reprod Biomed* 2016;**14**:193–8. https://doi.org/10.29252/ijrm.14.3.193

21. Eskew AM, Reschke LD, Woolfolk C, Schulte MB, Boots CE, Broughton DE, *et al.* Effect of endometrial mechanical stimulation in an unselected population undergoing in vitro fertilization: futility analysis of a double-blind randomized controlled trial. *J Assist Reprod Genet* 2019;**36**:299–305. https://doi.org/10.1007/s10815-018-1356-5

22. Izquierdo Rodriguez A, de la Fuente Bitaine L, Spies K, Lora D, Galindo A. Endometrial scratching effect on clinical pregnancy rates in patients undergoing egg donor in vitro fertilization cycles: the ENDOSCRATCH randomized clinical trial (NCT03108157). *Reprod Sci* 2020;**27**:1863–72. https://doi.org/10.1007/s43032-020-00204-8

23. Howard S. The hidden costs of infertility treatment. *BMJ* 2018;**361**:k2204. https://doi.org/10.1136/bmj.k2204

24. Tuckerman E, Mariee N, Prakash A, Li TC, Laird S. Uterine natural killer cells in peri-implantation endometrium from women with repeated implantation failure after IVF. *J Reprod Immunol* 2010;**87**:60–6. https://doi.org/10.1016/j.jri.2010.07.001

25. Gnainsky Y, Granot I, Aldo PB, Barash A, Or Y, Schechtman E, *et al.* Local injury of the endometrium induces an inflammatory response that promotes successful implantation. *Fertil Steril* 2010;**94**:2030–6. https://doi.org/10.1016/j.fertnstert.2010.02.022

26. Kalma Y, Granot I, Gnainsky Y, Or Y, Czernobilsky B, Dekel N, Barash A. Endometrial biopsy-induced gene modulation: first evidence for the expression of bladder-transmembranal uroplakin Ib in human endometrium. *Fertil Steril* 2009;**91**:1042–49.E9. https://doi.org/10.1016/j.fertnstert.2008.01.043

27. Lensen S, Sadler L, Farquhar C. Endometrial scratching for subfertility: everyone's doing it. *Hum Reprod* 2016;**31**:1241–4. https://doi.org/10.1093/humrep/dew053

28. Mackens S, Racca A, Van de Velde H, Drakopoulos P, Tournaye H, Stoop D, *et al.* Follicular-phase endometrial scratching: a truncated randomized controlled trial. *Hum Reprod* 2020;**35**:1090–8. https://doi.org/10.1093/humrep/deaa018

29. Human Fertilisation and Embryology Authority (HFEA). *Improving Outcomes for Fertility Patients: Multiple Births. A Statistical Report.* London: HFEA; 2015. URL: https://ifqlive.blob.core.windows.net/umbraco-website/1169/multiple_births_report_2015.pdf (accessed 12 March 2021).

30. Pinborg A. IVF/ICSI twin pregnancies: risks and prevention. *Hum Reprod Update* 2005;**11**:575–93. https://doi.org/10.1093/humupd/dmi027

31. Martikainen H, Tiitinen A, Tomás C, Tapanainen J, Orava M, Tuomivaara L, *et al.* One versus two embryo transfer after IVF and ICSI: a randomized study. *Hum Reprod* 2001;**16**:1900–3. https://doi.org/10.1093/humrep/16.9.1900

32. Ledger WL, Anumba D, Marlow N, Thomas CM, Wilson EC, Cost of Multiple Births Study Group (COMBS Group). The costs to the NHS of multiple births after IVF treatment in the UK. *BJOG* 2006;**113**:21–5. https://doi.org/10.1111/j.1471-0528.2005.00790.x

33. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ* 2015;**350**:h2147. https://doi.org/10.1136/bmj.h2147

34. Metwally M, Chatters R, Pye C, Dimairo M, Walters S. *A Multicentre Randomised Controlled Trial of Induced Endometrial Scratch in Women Undergoing First Time In Vitro Fertilisation (IVF) – Protocol.* Sheffield: University of Sheffield; 2020. https://doi.org/10.15131/shef.data.12886997.v1

35. Pye C, Chatters R, Cohen J, Brian K, Cheong YC, Laird S, *et al.* Induced endometrial trauma (endometrial scratch) in the mid-luteal menstrual cycle phase preceding first cycle IVF/ICSI versus usual IVF/ICSI therapy: study protocol for a randomised controlled trial. *BMJ Open* 2018;**8**:e020755. https://doi.org/10.1136/bmjopen-2017-020755

36. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, *et al.* Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;**348**:g1687. https://doi.org/10.1136/bmj.g1687

37. European Surveillance of Congenital Anomalies (EUROCAT). *EUROCAT Guide 1.4, Section 3.2: Minor Anomalies for Exclusion.* Brussels: EUROCAT; 2014. URL: https://eu-rd-platform.jrc.ec. europa.eu/sites/default/files/Section 3.2- 27_Oct2016.pdf (accessed 12 March 2021).

38. Cutting R, Morroll D, Roberts SA, Pickering S, Rutherford A, BFS and ACE. Elective single embryo transfer: guidelines for practice British Fertility Society and Association of Clinical Embryologists. *Hum Fertil* 2008;**11**:131–46. https://doi.org/10.1080/14647270802302629

39. Scientists in Reproductive Medicine, ESHRE Special Interest Group of Embryology. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. *Hum Reprod* 2011;**26**:1270–83. https://doi.org/10.1093/humrep/der037

40. Harbottle S, Hughes C, Cutting R, Roberts S, Brison D, Association Of Clinical Embryologists (ACE), The British Fertility Society (BFS). Elective single embryo transfer: an update to UK best practice guidelines. *Hum Fertil* 2015;**18**:165–83. https://doi.org/10.3109/14647273.2015. 1083144

41. Dimairo M, Walters S, Metwally M, Watkins A. *Endometrial Scratch Trial Statistical Analysis Plan (Final Version 2.1).* Sheffield: University of Sheffield; 2020.

42. Maheshwari A, Bhattacharya S, Bowler U, Brison D, Child T, Cole C, *et al.* Study protocol: e-freeze – freezing of embryos in assisted conception: a randomised controlled trial evaluating the clinical and cost effectiveness of a policy of freezing embryos followed by thawed frozen embryo transfer compared with a policy of fresh emb. *Reprod Health* 2019;**16**:81. https://doi.org/10.1186/s12978-019-0737-2

43. Miller D, Pavitt S, Sharma V, Forbes G, Hooper R, Bhattacharya S, *et al.* Physiological, hyaluronan-selected intracytoplasmic sperm injection for infertility treatment (HABSelect): a parallel, two-group, randomised trial. *Lancet* 2019;**393**:416–22. https://doi.org/10.1016/ S0140-6736(18)32989-1

44. European Medicines Agency. *Guideline for Good Clinical Practice E6(R2)*. URL: https://s3.eu-west-2.amazonaws.com/www.hra.nhs.uk/media/documents/ema-gcp-guidance.pdf (accessed 27 October 2021).

45. Senn S. Testing for baseline balance in clinical trials. *Stat Med* 1994;**13**:1715–26. https://doi.org/10.1002/sim.4780131703

46. Knol MJ, Groenwold RH, Grobbee DE. P-values in baseline tables of randomised controlled trials are inappropriate but still common in high impact journals. *Eur J Prev Cardiol* 2012;**19**:231–2. https://doi.org/10.1177/1741826711421688

47. de Boer MR, Waterlander WE, Kuijper LD, Steenhuis IH, Twisk JW. Testing for baseline differences in randomized controlled trials: an unhealthy research behavior that is hard to eradicate. *Int J Behav Nutr Phys Act* 2015;**12**:4. https://doi.org/10.1186/s12966-015-0162-z

48. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, *et al.* CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;**340**:c869. https://doi.org/10.1136/bmj.c869

49. Norton EC, Miller MM, Kleinman LC. Computing adjusted risk ratios and risk differences in Stata. *Stata J* 2013;**13**:492–509. https://doi.org/10.1177/1536867X1301300304

50. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med* 1998;**17**:857–72. https://doi.org/10.1002/(SICI)1097-0258(19980430)17:8<857::AID-SIM777>3.0.CO;2-E

51. Villar J, Cheikh Ismail L, Victora CG, Ohuma EO, Bertino E, Altman DG, *et al.* International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. *Lancet* 2014;**384**:857–68. https://doi.org/10.1016/S0140-6736(14)60932-6

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

53. Department of Health. *NHS Reference Costs 2012/13*. London: Department of Health; 2013. URL: www.gov.uk/government/publications/nhs-reference-costs-2012-to-2013 (accessed 12 March 2021).

54. Curtis L. *Unit Costs of Health and Social Care 2013*. Canterbury: PSSRU, University of Kent; 2013. URL: www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2013/ (accessed 12 March 2021).

55. Office for National Statistics. *Annual Survey for Hours and Earnings (ASHE)*. 2012. URL: www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/bulletins/annualsurveyofhoursandearnings/2012-11-22 (accessed 12 March 2021).

56. Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. New York, NY: Chapman and Hall/CRC Press; 1994. https://doi.org/10.1201/9780429246593

57. UK Government. *Expenses and Benefits: Business Travel Mileage for Employees' Own Vehicles*. URL: www.gov.uk/expenses-and-benefits-business-travel-mileage/rules-for-tax (accessed 20 August 2020).

58. Office for National Statistics. (2019) *Earnings and Hours Worked, All Employees: AHSE Table 1*. URL: www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/datasets/allemployeesashetable1 (accessed 12 March 2021).

59. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Heal Care* 2007;**19**:349–57. https://doi.org/10.1093/intqhc/mzm042

60. Braun V, Clark V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;**3**:77–101. https://doi.org/10.1191/1478088706qp063oa

61. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, *et al*. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLOS Med* 2009;**6**:e1000100. https://doi.org/10.1371/journal.pmed.1000100

62. The Cochrane Collaboration. *RoB 2: A Revised Cochrane Risk-of-bias Tool for Randomized Trials.* URL: https://methods.cochrane.org/bias/resources/rob-2-revised-cochrane-risk-bias-tool-randomized-trials (accessed 26 October 2021).

63. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester: John Wiley & Sons Ltd; 2019. https://doi.org/10.1002/9781119536604

64. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;**340**:c332. https://doi.org/10.1136/bmj.c332

65. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, *et al*. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008;**337**:a2390. https://doi.org/10.1136/bmj.a2390

66. Ioannidis JP, Evans SJ, Gøtzsche PC, O'Neill RT, Altman DG, Schulz K, *et al*. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004;**141**:781–8. https://doi.org/10.7326/0003-4819-141-10-200411160-00009

67. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, *et al*. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) – explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. *Value Health* 2013;**16**:231–50. https://doi.org/10.1016/j.jval.2013.02.002

68. Geslevich J. *Endometrial Curettage Before Embryo Transfer*. 2006. URL: https://clinicaltrials.gov/show/NCT00367367 (accessed 12 March 2021).

69. Marsh E. *Endometrial Injury and In Vitro Fertilization Outcomes*. 2013. URL: https://clinicaltrials.gov/show/NCT02153814 (accessed 12 March 2021).

70. Checa MA. *Endometrial Scratching in In Vitro Fertilization Cycles with Oocyte Donation*. 2013. URL: https://clinicaltrials.gov/show/NCT01842178 (accessed 12 March 2021).

71. Shehata A, Mohamed A. *Endometrial Scratching by Pipelle on Pregnancy Rate in Unexplained Infertility*. 2014. URL: https://clinicaltrials.gov/show/NCT02084914 (accessed 12 March 2021).

72. Shawki HE, Youssef AM, Hassan MMM. Routine endometrial scratching in women with first ICSI cycle; does it matter? *Int J Gynecol Obstet* 2018;**143**:811.

73. Aflatoonian A, Baradaran Bagheri R, Hosseinisadat R. The effect of endometrial injury on pregnancy rate in frozen-thawed embryo transfer: a randomized control trial. *Int J Reprod Biomed* 2016;**14**:453–158. https://doi.org/10.29252/ijrm.14.7.3

74. Frantz S, Parinaud J, Kret M, Rocher-Escriva G, Papaxanthos-Roche A, Creux H, *et al*. Decrease in pregnancy rate after endometrial scratch in women undergoing a first or second in vitro fertilization. A multicenter randomized controlled trial. *Hum Reprod* 2019;**34**:92–9. https://doi.org/10.1093/humrep/dey334

75. Zhao SY, Liu Y, Yu LP, Chao H, Liu MH, Zhao YY. Endometrial injury performed during the cycle preceding ovarian stimulation increases the biochemical pregnancy rate in unselected infertile women undergoing in vitro fertilization: a randomized placebo controlled trial. *Hum Reprod* 2015;**30**:i279.

76. Sherbiny Y. *Effect of Endometrial Injury before Frozen Embryo Transfer on Pregnancy Rate*. 2017. URL: https://clinicaltrials.gov/ct2/show/NCT03220503?term=sherbiny&draw=2&rank=1 (accessed 12 March 2021).

77. Shriver EK. *Endometrial Biopsy in Infertile Patients*. 1999. URL: https://clinicaltrials.gov/ct2/show/NCT00064935 (accessed 12 March 2021).

78. Hurst B. *Comparing Two Types of Endometrial Activation Prior to Embryo Transfer*. 2014. URL: https://clinicaltrials.gov/ct2/show/NCT04363879 (accessed 12 March 2021).

79. Guven S, Kart C, Unsal MA, Yildirim O, Odaci E, Yulug E. Endometrial injury may increase the clinical pregnancy rate in normoresponders undergoing long agonist protocol ICSI cycles with single embryo transfer. *Eur J Obstet Gynecol Reprod Biol* 2014;**173**:58–62. https://doi.org/10.1016/j.ejogrb.2013.11.005

80. Hebeisha SA, Moiety FS, Samir M, Hussein M. Effect of endometrial injury on implantation and pregnancy rates: a randomised controlled trial. *Clin Exp Obstet Gynecol* 2018;**45**:105–8.

81. Hilton J, Liu KE, Laskin CA, Havelock J. Effect of endometrial injury on in vitro fertilization pregnancy rates: a randomized, multicentre study. *Arch Gynecol Obstet* 2019;**299**:1159–64. https://doi.org/10.1007/s00404-019-05044-9

82. Mehrafza M, Asgharnia M, Abdollahiyan P, Oudi M, Nikpuri Z, Mohammad Tabar Z, *et al*. The effect of local injury to endometrium on pregnancy rate in ICSI: a randomized controlled trial. *Iran J Reprod Med* 2011;**9**:44–5.

83. Merriam K, Rozario N, Marshburn P, Matthews M, Usadi RS, McCall J, *et al*. Comparing two types of endometrial activation before embryo transfer: a pilot study. *Fertil Steril* 2017;**108**:e83. https://doi.org/10.1016/j.fertnstert.2017.07.258

84. Sherif A, Abou-Talib Y, Ibrahim M, Arafat R. The effect of day 6 endometrial injury of the ICSI cycle on pregnancy rate: a randomized controlled trial. *Middle East Fertil Soc J* 2018;**23**:292–6. https://doi.org/10.1016/j.mefs.2018.02.001

85. Spandorfer SD, Liu HC, Xu KP, Barmat LI, Davis OK. The day of the luteal phase endometrial biopsy (EBx) is an important predictor of success when autologous endometrial coculture (AECC) is utilized in IVF-ET. *Fertil Steril* 1998;**30**:S117.

86. Vidal C, Giles J, Labarta E, Castillon G, Martinez-Salazar J, Fernandez L, *et al*. Intentional endometrial injury trying to improve clinical outcomes of an oocyte donation program in patients without RIF. Interim analysis of a randomized controlled trial. *Fertil Steril* 2019;**112**:e164–5. https://doi.org/10.1016/j.fertnstert.2019.07.546

87. Nastri CO, Ferriani RA, Raine-Fenning N, Martins WP. Endometrial scratching performed in the non-transfer cycle and outcome of assisted reproduction: a randomized controlled trial. *Ultrasound Obstet Gynecol* 2013;**42**:375–82. https://doi.org/10.1002/uog.12539

88. Polanski L, Raine-Fenning N. *A Study to Assess if Scratching the Lining of the Womb Prior to IVF Treatment Increases the Chances of Pregnancy*. 2014. URL: https://clinicaltrials.gov/ct2/show/NCT01882842 (accessed 12 March 2021).

89. Yeung TW, Chai J, Li RH, Lee VC, Ho PC, Ng EH. The effect of endometrial injury on ongoing pregnancy rate in unselected subfertile women undergoing in vitro fertilization: a randomized controlled trial. *Hum Reprod* 2014;**29**:2474–81. https://doi.org/10.1093/humrep/deu213

90. Maged AM, Al-Inany H, Salama KM, Souidan II, Abo Ragab HM, Elnassery N. Endometrial scratch injury induces higher pregnancy rate for women with unexplained infertility undergoing IUI with ovarian stimulation: a randomized controlled trial. *Reprod Sci* 2016;**23**:239–43. https://doi.org/10.1177/1933719115602776

91. Griesinger G. Beware of the 'implantation rate'! Why the outcome parameter 'implantation rate' should be abandoned from infertility research. *Hum Reprod* 2016;**31**:249–51. https://doi.org/10.1093/humrep/dev322

92. El-Toukhy T, Sunkara S, Khalaf Y. Local endometrial injury and IVF outcome: a systematic review and meta-analysis. *Reprod Biomed Online* 2012;**25**:345–54. https://doi.org/10.1016/j.rbmo.2012.06.012

93. Zech N, Holl P. IVF add-ons: fact, fiction, fake or fortune? *J Fertil In Vitro IVF Worldw Reprod Med Genet Stem Cell Biol* 2020;**8**:1–7.

94. Datta AK, Campbell S, Deval B, Nargund G. Add-ons in IVF programme – hype or hope? *Facts Views Vis Obgyn* 2015;**7**:241–50.

95. Zuidgeest MGP, Goetz I, Groenwold RHH, Irving E, van Thiel GJMW, Grobbee DE. Series: pragmatic trials and real world evidence: paper 1. Introduction. *J Clin Epidemiol* 2017;**88**:7–13. https://doi.org/10.1016/j.jclinepi.2016.12.023

96. Yoldemir T, Erenus M. Does the timing of mock embryo transfer trial improve implantation in intracytoplasmic sperm injection cycles? *Gynecol Endocrinol* 2011;**27**:396–400. https://doi.org/10.3109/09513590.2010.491927

97. Skingley A, Bungay H, Clift S, Warden J. Experiences of being a control group: lessons from a UK-based randomized controlled trial of group singing as a health promotion initiative for older people. *Health Promot Int* 2014;**29**:751–8. https://doi.org/10.1093/heapro/dat026

98. Meinich Petersen S, Zoffmann V, Kjærgaard J, Graff Stensballe L, Greisen G. Disappointment and adherence among parents of newborns allocated to the control group: a qualitative study of a randomized clinical trial. *Trials* 2014;**15**:126. https://doi.org/10.1186/1745-6215-15-126

99. Lindström D, Sundberg-Petersson I, Adami J, Tönnesen H. Disappointment and drop-out rate after being allocated to control group in a smoking cessation trial. *Contemp Clin Trials* 2010;**31**:22–6. https://doi.org/10.1016/j.cct.2009.09.003

100. Jenkins SP, Calvert MJ, Draper H. Potential research participants' use of information during the consent process: a qualitative pilot study of patients enrolled in a clinical trial. *PLOS ONE* 2020;**15**:e0234388. https://doi.org/10.1371/journal.pone.0234388

101. Antoniou EE, Draper H, Reed K, Burls A, Southwood TR, Zeegers MP. An empirical study on the preferred size of the participant information sheet in research. *J Med Ethics* 2011;**37**:557–62. https://doi.org/10.1136/jme.2010.041871

102. Lidz CW, Appelbaum PS, Grisso T, Renaud M. Therapeutic misconception and the appreciation of risks in clinical trials. *Soc Sci Med* 2004;**58**:1689–97. https://doi.org/10.1016/S0277-9536(03)00338-1

103. Canvin K, Jacoby A. Duty, desire or indifference? A qualitative study of patient decisions about recruitment to an epilepsy treatment trial. *Trials* 2006;**7**:32. https://doi.org/10.1186/1745-6215-7-32

104. Perrotta M, Geampana A. The trouble with IVF and randomised control trials: professional legitimation narratives on time-lapse imaging and evidence-informed care. *Soc Sci Med* 2020;**258**:113115. https://doi.org/10.1016/j.socscimed.2020.113115

105. Eijkemans MJC, Kersten FAM, Lintsen AME, Hunault CC, Bouwmans CAM, Roijen LH, *et al.* Cost-effectiveness of 'immediate IVF' versus 'delayed IVF': a prospective study. *Hum Reprod* 2017;**32**:999–1008. https://doi.org/10.1093/humrep/dex018

106. Curtis L, Burns A. *Unit Costs of Health and Social Care 2019*. Canterbury: PSSRU, University of Kent; 2019.

107. Paramasivan S, Rogers CA, Welbourn R, Byrne JP, Salter N, Mahon D, *et al.* Enabling recruitment success in bariatric surgical trials: pilot phase of the By-Band-Sleeve study. *Int J Obes* 2017;**41**:1654–61. https://doi.org/10.1038/ijo.2017.153

108. Elliott D, Hamdy FC, Leslie TA, Rosario D, Dudderidge T, Hindley R, *et al.* Overcoming difficulties with equipoise to enable recruitment to a randomised controlled trial of partial ablation vs radical prostatectomy for unilateral localised prostate cancer. *BJU Int* 2018;**122**:970–7. https://doi.org/10.1111/bju.14432

109. Blencowe NS, Cook JA, Pinkney T, Rogers C, Reeves BC, Blazeby JM. Delivering successful randomized controlled trials in surgery: methods to optimize collaboration and study design. *Clin Trials* 2017;**14**:211–18. https://doi.org/10.1177/1740774516687272

110. Kirkby HM, Calvert M, Draper H, Keeley T, Wilson S. What potential research participants want to know about research: a systematic review. *BMJ Open* 2012;**2**:e000509. https://doi.org/10.1136/bmjopen-2011-000509

111. Dechartres A, Trinquart L, Boutron I, Ravaud P. Influence of trial sample size on treatment effect estimates: meta-epidemiological study. *BMJ* 2013;**346**:f2304. https://doi.org/10.1136/bmj.f2304

112. Wilkinson J, Roberts SA, Showell M, Brison DR, Vail A. No common denominator: a review of outcome measures in IVF RCTs. *Hum Reprod* 2016;**31**:2714–22. https://doi.org/10.1093/humrep/dew227

113. Duffy JM, van 't Hooft J, Gale C, Brown M, Grobman W, Fitzpatrick R, *et al.* A protocol for developing, disseminating, and implementing a core outcome set for pre-eclampsia. *Pregnancy Hypertens* 2016;**6**:274–8. https://doi.org/10.1016/j.preghy.2016.04.008

114. Smith CT, Hopkins C, Sydes M, Woolfall K, Clarke M, Murray G, *et al.* Good practice principles for sharing individual participant data from publicly funded clinical trials. *Trials* 2015;**16**(Suppl. 2):O1. https://doi.org/10.1186/1745-6215-16-S2-O1

115. Department of Health and Social Care. *National Cost Collection: National Schedule of Costs 2018-19*. URL: https://improvement.nhs.uk/resources/national-cost-collection/ (accessed 12 March 2021).

116. NHS Employers. *NHS Terms and Conditions (AfC) Pay Scales 2018/19*. URL: www.nhsemployers.org/sites/default/files/media/NHS-terms-and-conditions-pay-poster-201819_0.pdf (accessed 12 March 2021).

117. Bell H. *Natural at Heart Price List*. URL: www.naturalatheart.co.uk/price-list.html (accessed 12 March 2021).

118. National Institute for Health and Care Excellence (NICE). *Resource Impact Report: Flu Vaccination: Increasing Uptake (NG103)*. London: NICE; 2018. URL: www.nice.org.uk/guidance/ng103/resources/resource-impact-report-pdf-6532121197 (accessed 12 March 2021).

119. McKinnon G. *Hypnotherapy Session Prices*. URL: www.hypnosiskent.co.uk/session-prices (accessed 24 December 2021).

120. Young J. *Janine Young Reflexology Price List*. URL: www.janineyoungreflexology.com/phdi/p1.nsf/supppages/1860?opendocument&part=4 (accessed 12 March 2021).

121. Real World Reiki. *Real World Reiki Price List*. URL: http://realworldreiki.co.uk/how-much-does-a-session-cost/4593171091 (accessed 12 March 2021).

122. Shahrokh-Tehraninejad E, Dashti M, Hossein-Rashidi B, Azimi-Nekoo E, Haghollahi F, Kalantari V. A randomized trial to evaluate the effect of local endometrial injury on the clinical pregnancy rate of frozen embryo transfer cycles in patients with repeated implantation failure. *J Family Reprod Health* 2016;**10**:108–14.

123. Abdelmagied AM, Abdalmageed OS, Abbas A, Farghaly TA, Ali MK, Nassr AA, *et al.* Hysteroscopic surgery as a surrogate marker for local endometrial injury in fresh IVF cycles. *Fertil Steril* 2016;**106**:e220. https://doi.org/10.1016/j.fertnstert.2016.07.635

124. Chang E, Check JH, Liss JR, Choe J, Cohen R. An endometrial scratch can improve pregnancy rates in natural cycles of women with unexplained infertility given luteal phase support. *Fertil Steril* 2017;**108**:e369. https://doi.org/10.1016/j.fertnstert.2017.07.1078

125. Chawla S, Purandare N, Mocanu E, Hughes C, Deignan K, Naasan M, *et al.* Does the endometrial scratch improve implantation rates? *Ir Med J* 2015;**108**:1.

126. Du YB, Gao MZ, Shi Y, Sun ZG, Wang J. Endocrine and inflammatory factors and endometriosis-associated infertility in assisted reproduction techniques. *Arch Gynecol Obstet* 2013;**287**:123–30. https://doi.org/10.1007/s00404-012-2567-0

127. Lensen S, Wilkinson J, Sadler L. A randomized trial of endometrial scratching before in vitro fertilization. Reply. *N Engl J Med* 2019;**380**:1777–8. https://doi.org/10.1056/NEJMc1902642

128. Liang Y, Han J, Jia C, Ma Y, Lan Y, Li Y, Wang S. Effect of endometrial injury on secretion of endometrial cytokines and IVF outcomes in women with unexplained subfertility. *Mediators Inflamm* 2015;**2015**:757184. https://doi.org/10.1155/2015/757184

129. McKenzie L. Does endometrial injury improve ART outcomes? *Contemp Ob Gyn* 2012;**57**:11.

130. Mitic D, Basic M, Petric A, Stamenovic S. The effects of local endometrial injury made by hysteroscopy on in vitro fertilization outcome. *Vojnosanit Pregl* 2018;**75**:570–5. https://doi.org/10.2298/VSP160526233M

131. Moghissi KS, Taylor HS. Controversies in OB/GYN. Should endometrial biopsy be part of an infertility evaluation? *Contemp Ob Gyn* 2005;**50**:71–6.

132. Mol BW, Barnhart KT. Scratching the endometrium in in vitro fertilization – time to stop. *N Engl J Med* 2019;**380**:391–2. https://doi.org/10.1056/NEJMe1815042

133. Nastri CO, Teixeira DM, Martins WP. Endometrial injury in the menstrual cycle prior to assisted reproduction techniques to improve reproductive outcomes. *Gynecol Endocrinol* 2013;**29**:401–2. https://doi.org/10.3109/09513590.2012.760193

134. Nayar KD, Jain G, Chulet S, Mohan A, Kant G. Scheduling hysteroscopic endometrial biopsy in endometrial preparation cycle just preceding embryo transfer (ET) v/s any time in the last one year: does it change the reproductive outcome? *Fertil Steril* 2015;**104**:e169. https://doi.org/10.1016/j.fertnstert.2015.07.525

135. Shokeir T. Impact of luteal phase hysteroscopy and concurrent endometrial biopsy on subsequent IVF cycle outcome. *Arch Gynecol Obstet* 2015;**291**:477. https://doi.org/10.1007/s00404-014-3579-8

136. Siristatidis C, Rigos I, Pergialiotis V, Karageorgiou V, Christoforidis N, Daskalakis G, *et al.* Endometrial injury for patients with endometriosis and polycystic ovary syndrome undergoing medically assisted reproduction: current data and a protocol. *Horm Mol Biol Clin Investig* 2018;**35**:20180040. https://doi.org/10.1515/hmbci-2018-0040

137. Vanvitelli L. *Office Hysteroscopy in Infertility Work-out: The Role of Endometrial Hyperplasia in Southern Italian Women*. 2019. URL: www.clinicaltrials.gov/ct2/show/NCT03965585 (accessed 12 March 2021).

138. Endometrial scratching improves IVF results. *BMJ* 2013;**347**:4. https://doi.org/10.1136/bmj.f6007

139. Wise J. Endometrial scratching improves IVF pregnancy rate. *BMJ* 2013;**347**:f6007. https://doi.org/10.1136/bmj.f6007

140. Zhou L, Li R, Wang R, Huang HX, Zhong K. Local injury to the endometrium in controlled ovarian hyperstimulation cycles improves implantation rates. *Fertil Steril* 2008;**89**:1166–76. https://doi.org/10.1016/j.fertnstert.2007.05.064

141. Genod A. *Endometrial Local Injury before First IVF: Evaluation of Pregnancy Rate*. 2015. URL: https://clinicaltrials.gov/show/NCT02522806 (accessed 12 March 2021).

142. Lee VCY. *A RCT on the Effect of Endometrial Injury on Improving Ongoing Pregnancy Rate in Subfertile Women Undergoing FET Cycles*. URL: https://clinicaltrials.gov/ct2/show/NCT02197832 (accessed 4 November 2021).

143. Martins WP. *Effect of Endometrial Scratching on Assisted Reproduction Outcomes: A Randomized Controlled Trial*. 2014. URL: https://clinicaltrials.gov/show/NCT02180256 (accessed 12 March 2021).

144. Martins WP. *Endometrial Scratching for Cryopreserved Embryo Transfer*. 2014. URL: https://clinicaltrials.gov/show/NCT02093442 (accessed 12 March 2021).

145. Bahaa A. *Does Endometrial Injury Improve Intrauterine Insemination Outcome?* 2015. URL: https://clinicaltrials.gov/show/NCT02542280 (accessed 12 March 2021).

146. Checa MA. *Endometrial Priming for Embryo Transfer*. 2010. URL: https://clinicaltrials.gov/ct2/show/NCT01430650 (accessed 12 March 2021).

147. Gibreel A, Badawy A, El-Refai W, El-Adawi N. Endometrial scratching to improve pregnancy rate in couples with unexplained subfertility: a randomized controlled trial. *J Obstet Gynaecol Res* 2013;**39**:680–4. https://doi.org/10.1111/j.1447-0756.2012.02016.x

148. Gibreel A, Ali R, Hemida R, Sherif L, El-Adawi N. Endometrial scratch for infertile polycystic ovary syndrome (PCOS) women undergoing laparoscopic ovarian drilling: a randomized controlled trial. *Middle East Fertil Soc J* 2020;**24**:1. https://doi.org/10.1186/s43043-019-0001-2

149. Movahedi S. *Local Injury to the Endometrium Does Not Improve the Implantation Rate in Good Responder Patients Undergoing In-vitro Fertilization*. 2011. URL: https://trialsearch.who.int/?TrialID=IRCT201008154572N1 (accessed 27 October 2021).

150. Shokeir T, Ebrahim M, El-Mogy H. Hysteroscopic-guided local endometrial injury does not improve natural cycle pregnancy rate in women with unexplained infertility: randomized controlled trial. *J Obstet Gynaecol Res* 2016;**42**:1553–7. https://doi.org/10.1111/jog.13077

151. Soliman BS, Harira M. Local endometrial scratching under ultrasound-guidance after failed intrauterine insemination and cycle outcome: a randomized controlled trial. *Middle East Fertil Soc J* 2017;**22**:60–6. https://doi.org/10.1016/j.mefs.2016.06.006

152. Vilela M, Bello A, Sueldo C, Quintana R, Diradourian M, Marconi G. Abstract no: O262. Endometrial lesions caused by catheters used for embryo transfer. *Hum Reprod* 2003;**18**:90.

153. Xu B, Zhang Q, Hao J, Xu D, Li Y. Two protocols to treat thin endometrium with granulocyte colony-stimulating factor during frozen embryo transfer cycles. *Reprod Biomed Online* 2015;**30**:349–58. https://doi.org/10.1016/j.rbmo.2014.12.006

154. Alalfy M. *Physical Endometrial Manipulation and its Effect on ICSI*. 2017. URL: https://clinicaltrials.gov/ct2/show/NCT03345251 (accessed 12 March 2021).

155. Baum M. *Endometrial Sampling (Pipelle) in IVF Patients*. 2006. URL: https://clinicaltrials.gov/show/NCT00411021 (accessed 12 March 2021).

156. Chi CI. *Effect of Endometrial Injury on MMPs and its Inhibitor and Pregnancy Outcome before Frozen-thawed Embryo Transfer*. 2017. URL: www.chictr.org.cn/hvshowproject.aspx?id=12400 (accessed 12 March 2021).

157. Farrag MM. *Autologous Intrauterine Platelet-Rich Plasma Instillation and Endometrial Scratching for Thinned Endometrium*. 2020. URL: https://clinicaltrials.gov/ct2/show/NCT04240860 (accessed 12 March 2021).

158. Gomaa MF. *Value of Routine Hysteroscopy Prior to IVF/ICSI Cycles*. 2014. URL: https://clinicaltrials.gov/ct2/show/NCT02245750 (accessed 12 March 2021).

159. Hare KJ. Office hysteroscopy before IVF treatment-is fertility improved? *Gynecol Surg* 2013;**10**:S71.

160. Helmy MEE, Maher MA, Elkhouly NI, Ramzy M. A randomized trial of local endometrial injury during ovulation induction cycles. *Int J Gynaecol Obstet* 2017;**138**:47–52. https://doi.org/10.1002/ijgo.12178

161. Kandavel V, Quenby S. *Patient Acceptability and Perception of Randomisation in a Trial of Scratch in Recurrent Miscarriage (SiM trial)*. 34th Annual Meeting of the European Society of Human Reproduction and Embryology (ESHRE), 1–4 July 2018, Barcelona, Spain, abstract no. 33 (Suppl 1). https://doi.org/10.26226/morressier.5af300b1738ab10027aa9745

162. Levin D. *Endometrial Biopsy Protocol for In Vitro Fertilization (IVF)*. 2011. URL: https://clinicaltrials.gov/show/NCT01278706 (accessed 12 March 2021).

163. Maged A. *Effect of Endometrial Injury in Couples with Unexplained Infertility*. 2018. URL: https://clinicaltrials.gov/ct2/show/NCT03398993 (accessed 12 March 2021).

164. Mak JSM, Chung CHS, Chung JPW, Kong GWS, Saravelos SH, Cheung LP, Li TC. The effect of endometrial scratch on natural-cycle cryopreserved embryo transfer outcomes: a randomized controlled study. *Reprod Biomed Online* 2017;**35**:28–36. https://doi.org/10.1016/j.rbmo.2017.04.004

165. Olesen MS, Hauge B, Ohrt L, Olesen TN, Roskær J, Bæk V, *et al*. Therapeutic endometrial scratching and implantation after in vitro fertilization: a multicenter randomized controlled trial. *Fertil Steril* 2019;**112**:1015–21. https://doi.org/10.1016/j.fertnstert.2019.08.010

166. Parsaeian Z. *Comparison of Local Endometrial Injury Efficacy in Follicular versus Luteal Phases on the Success Rate of Frozen Embryo Transfer Cycle*. 2018. URL: www.cochranelibrary.com/central/doi/10.1002/central/CN-01906464/full (accessed 28 April 2021).

167. Salama KM. *Endometrial Injury for Unexplained Infertility*. 2013. URL: https://clinicaltrials.gov/ct2/show/NCT02863198 (accessed 12 March 2021).

168. Shokeir T. *Endometrial Injury in Women with Unexplained Infertility*. 2017. URL: https://clinicaltrials.gov/ct2/show/NCT02863198 (accessed 12 March 2021).

169. Sun Y. *The Effectiveness of Endometrial Injury in IVF*. 2015. URL: https://clinicaltrials.gov/show/NCT02409745 (accessed 12 March 2021).

170. Zayed A. *Retrieval Versus Mid-luteal Endometrial Scratching (ES) for Intra Cytoplasmic Sperm Injection (ICSI)*. 2018. URL: https://clinicaltrials.gov/show/NCT03470298 (accessed 12 March 2021).

Appendix 1 PRECIS-2

Domain	Score	Rationale
1. Eligibility criteria	3	Patients only included if they are predicted to be good responders to treatment with SET – differs from the patients who would receive the intervention in usual care
2. Recruitment path	5	Recruitment took place at routine care appointments
3. Setting	5	Identical setting to usual care, including both NHS and privately run fertility units
4. Organisation intervention	4	A slight increase in resources needed to deliver the intervention in order to collect pain scores. However, no additional training was required, although sites referred to an ES procedure SOP. The information sheets were provided to all participants regarding the ES procedure, which may not be provided during usual care
5. Flex of experimental intervention – delivery	4	There was flexible delivery in terms of timing (within limits) and exact protocol, but a SOP was provided to ensure some consistency in the delivery of the ES intervention across centres
6. Flex of experimental intervention – adherence	5	There was no more than 'usual' encouragement to undergo the intervention
7. Follow-up	3	There were some follow-ups above usual care practice. The outcome of IVF cycles would usually be followed up by centres, but we undertook higher intensity follow-up than usual (e.g. 3 month, 6 months and 6 weeks post partum follow-ups) and collected additional information (e.g. AEs) more thoroughly and frequently than would be undertaken in usual care
8. Outcome	5	Achieving pregnancy and a live birth is of obvious importance to participants
9. Analysis	5	ITT analysis used that included all randomised participants with informed consent regardless of circumstances after randomisation

Appendix 2 Internal pilot feasibility criteria

Stop/go criteria		Internal pilot results
Recruitment	Recruitment of fewer than 108 participants (75% of the 144 target) during the internal pilot phase	170 participants recruited
Intervention delivery	Less than 75% of women scheduled to receive their ES procedure have received the ES at the correct time point (mid-luteal phase)	76.7% women received ES who were scheduled to receive it

Appendix 3 Changes to protocol

Substantial amendments

Version 1 (25 February 2016) amended to version 1.1 (29 March 2016)
The first amendment, following a request from the REC, added clarification that the outcome of the pregnancy will be confirmed with the patient's clinician before contacting the participant for follow-up information.

Version 1.1 (29 March 2016) amended to version 2 (9 May 2016)
The second amendment removed the necessity to randomise participants no later than 1 month before the start of their IVF cycle (thus allowing randomisation at any point prior to IVF); added Southampton as a site for the tissue substudy; an inclusion criterion was added that women are expected to receive IVF/ICSI treatment using fresh embryos; the time point of the 9-/10-month follow-up point was changed to 10.5 months (6 weeks post partum); addition of telephone follow-up pro-forma for nurses to use when undertaking follow-up telephone calls; change of research sites listed (Newcastle, Guy's and St Thomas' Hospital, Nottingham, Oxford added, Bradford removed, change of PI in Sheffield and Leeds).

Version 2 (9 May 2016) amended to version 3 (13 June 2016)
The third amendment included changes to the eligibility criteria (removal of cycle programming as an exclusion criterion); removal of stipulation that anti-inflammatory drugs cannot be taken prior to receiving the intervention; clarification of the action sites should take if a participant has not used a barrier method of contraception prior to the intervention; alteration to the subgroup analyses (correction to data collected for day of embryo transfer); change of archiving time frame from 15 to 5 years; unnecessary detail removed from SAE and AE section, and an AE and SAE collection time point added post procedure.

Version 3 (13 June 2016) amended to version 4 (4 October 2016)
The fourth amendment incorporated changes to the duration of the pilot study (from 6 months to 9 months) and changes to the stop/go criteria (removal of the necessity to set up at least four sites by the end of the pilot trial); detail added to follow-up process regarding process of following up pain score text messages following no response from participant; documents summarising both the Endometrial Scratch and E-Freeze[42] trials added to enable potential participants to make an informed choice regarding which trial to participate in; removal of requirement for HFEA consent to disclosure prior to informed consent to the trial; 24-hour post-intervention pain scale text renamed to '1 day'.

Version 4 (4 October 2016) amended to version 5 (20 July 2017)
The fifth amendment incorporated a change of name of 10.5 months' follow-up to 6 weeks post partum; clarification around the definition of a cycle of IVF within the trial (receiving any sort of stimulation); exclusion criteria added for participants undergoing protocols other than long or antagonist (i.e. ultra-long protocols) and for participants planned to undergo a scratch or similar procedure (e.g. endometrial biopsy for the collection of natural killer cells); clarification that a participant should still remain in the trial (and follow-up carried out) if the intervention is not carried out; detail added regarding collection of safety events at the time of pregnancy tests (sites should attempt to collect safety events from notes if the participant is not contactable following a negative pregnancy test); conditions added to list of expected AEs (anaemia, cholecystitis, epistaxis, itchy skin); electronic questionnaires added at 3-month and 6-week post partum follow-up, in addition to the already approved paper versions of the questionnaires at the same time points; summary letter introduced to be sent to participants who achieve a pregnancy in order to inform them when the 3 telephone follow-ups will be undertaken.

Version 5 (20 July 2017) amended to version 6 (13 November 2017) (Dundee only)

Amendment 6 sought approval for a separate protocol for our Dundee site, which, because of issues with participants having to travel vast distances to undertake consent and randomisation activities, requested a postal consent process. Participants were approached at their routine visit to the clinic and, if interested, were asked to sign and return a postal consent form to the unit stating that they were happy to be randomised to the trial. Randomisation then took place, and, if randomised to do so, the participant was scheduled for an ES procedure.

Version 6 (13 November 2017) (Dundee only) and version 5 (20 July 2017) (all other sites), amended to version 7 (28 January 2019) (Dundee only) and version 6 (28 January 2019) (all other sites)

In amendment 7, the number of centres participating in the trial was updated – this was stated in previous versions of the protocol as 'approximately 10', but was updated to the final number of sites ($n = 16$); clarification was provided regarding the follow-up of trial participants who undergo FET, where the first FET (where no previous transfer has been undertaken) will be followed up within the trial; additional information was added regarding the randomisation process, including allocation concealment; the qualitative substudy was added and the systematic review; changes were made to the reporting of AEs – events related to the birth of a baby or the process of birth were no longer to be classed as AEs or collected within the trial, events related to the weight or growth of the baby (low birthweight, very low birthweight, large for gestational age, preterm delivery, very preterm delivery and small for gestational age) were no longer to be classed as AEs, but would still be reported, events were added to the list of expected AEs in pregnancy (back pain, sciatica, bloating, breast tenderness/pain, cold/flu, feeling faint, fatigue/tiredness, frequent urination/nocturia, migraine, hot flushes, hyperemesis, implantation bleed, period pain, symphysis pubis dysfunction/hip pain/pelvis dysplasia, placenta previa (grade 1 and 2), prolapse, reflux/heartburn/indigestion) and during IVF treatment (abdominal swelling associated with mild OHSS, mild OHSS). Possible fetal abnormality was removed as an expected AE.

Non-substantial amendments

The PI at the South Tees trial site was changed from Ms Mohar Goswami to Mr Fayez Mustafa from February 2019, because of the relocation of Ms Goswami.

Appendix 4 List of recruiting sites and their involvement in the trial

Location of fertility unit	Phase of involvement	Site type[a]
Sheffield Teaching Hospitals NHS Foundation Trust	During and after internal pilot	NHS
University Hospitals of Leicester NHS Trust	During and after internal pilot	NHS
University Hospital Southampton NHS Foundation Trust	During and after internal pilot	NHS
Manchester University NHS Foundation Trust	During and after internal pilot	NHS
Leeds Teaching Hospitals NHS Trust	During and after internal pilot	NHS
Newcastle Upon Tyne Hospitals NHS Foundation Trust	During and after internal pilot	NHS
NHS Tayside (Dundee)	After internal pilot	NHS
Birmingham Women's and Children's NHS Foundation Trust	After internal pilot	NHS
Liverpool Women's NHS Foundation Trust	After internal pilot	NHS
Homerton University Hospital NHS Foundation Trust	After internal pilot	NHS
Guy's and St Thomas' NHS Foundation Trust	After internal pilot	NHS
Nurture Fertility (Nottingham)	After internal pilot	Private
Wrightington, Wigan and Leigh NHS Foundation Trust	After internal pilot	NHS
Oxford Fertility	After internal pilot	Private
Gateshead Health NHS Foundation Trust	After internal pilot	NHS
South Tees Hospitals NHS Foundation Trust	After internal pilot	NHS

a Private centres were those that were run outside the NHS by a private company, where staff were not employed by the NHS. NHS centres are fertility units within an NHS trust, but may also treat private patients.

Appendix 5 TIDieR checklist: endometrial scratch arm

Brief name	
1	Provide the name or a phrase that describes the intervention
	Endometrial Scratch (ES)

Why	
2	Describe any rationale, theory, or goal of the elements essential to the intervention
	The goal of the ES procedure is to invoke injury to the endometrium. The exact mechanism by which ES may improve implantation is not yet known, however it is known that implantation is a complex process involving the release of a number of inflammatory mediators including uterine natural killer cells, leukaemia inhibitory factor and interleukin 15. It is possible that ES may lead to the release of inflammatory cells and mediators such as macrophages and dendritic cells, tumour necrosis factor-α, interleukin 15, growth-regulated oncogene-α and macrophage inflammatory protein 1B

What	
3	Materials: Describe any physical or informational materials used in the intervention, including those provided to participants or used in intervention delivery or in training of intervention providers. Provide information on where the materials can be accessed (such as online appendix, URL)
	Those who provided the intervention followed a standard operating procedure (SOP) when undertaking the procedure (see Endometrial Scratch Standard Operating Procedure). Trial participants, prior to randomisation, received a PIS containing information about the procedure (see Appendices 4–6)
4	Procedures: Describe each of the procedures, activities, and/or processes used in the intervention, including any enabling or support activities
	The procedure was undertaken in the mid-luteal phase of the menstrual cycle prior to that in which IVF was being undertaken – see Endometrial Scratch Standard Operating Procedure

Who provided	
5	For each category of intervention provider (such as psychologist, nursing assistant), describe their expertise, background, and any specific training given
	A doctor or nurse at the fertility unit provided the ES, and had prior training, within the fertility unit, in order to undertake it. Fertility units already undertook ES on other patient groups (patients undergoing their second or subsequent cycles of IVF), and therefore were already trained to undertake the ES

How	
6	Describe the modes of delivery (such as face to face or by some other mechanism, such as internet or telephone) of the intervention and whether it was provided individually or in a group
	Face to face, on an individual basis

Where	
7	Describe the type(s) of location(s) where the intervention occurred, including any necessary infrastructure or relevant features
	The intervention was undertaken within fertility units, on an outpatient basis. There was no necessary infrastructure required, except for that inherent to fertility units

When and how much	
8	Describe the number of times the intervention was delivered and over what period of time, including the number of sessions, their schedule, and their duration, intensity or dose
	The intervention was delivered once and lasted around 20 minutes. However, if the ES was not possible (e.g. injury not induced due to the cervix being inaccessible), the procedure could be rescheduled and attempted again. This may have resulted in the participant having to reschedule their IVF cycle

Tailoring	
9	If the intervention was planned to be personalised, titrated or adapted, then describe what, why, when, and how
	Not applicable

Modifications	
10	If the intervention was modified during the course of the study, describe the changes (what, why, when, and how)
	Not applicable

How well	
11	Planned: If intervention adherence or fidelity was assessed, describe how and by whom, and if any strategies were used to maintain or improve fidelity, describe them
	Not applicable
12	Actual: If intervention adherence or fidelity was assessed, describe the extent to which the intervention was delivered as planned
	Those delivering the intervention were asked if the procedure was undertaken 'per protocol' (i.e. as described in the ES SOP). ES could not be undertaken in 10 participants due to the cervix being inaccessible. No other deviations were reported

Appendix 6 Expected adverse events

The following is a list of AEs related to pregnancy/IVF treatments that are expected in this patient population:

- abdominal pain
- abdominal swelling associated with mild OHSS
- anaemia
- back pain/sciatica
- bloating
- breast tenderness/pain
- clicky hip
- cholecystitis
- cold/flu
- conjunctivitis
- constipation
- cough
- diarrhoea
- dizziness/feeling faint
- epistaxis
- facial pain
- fall
- fatigue/tiredness
- frequent urination/nocturia
- gestational diabetes
- headache/migraine
- hot flushes
- hypertension
- hyperemesis
- implantation bleed
- itchy skin
- nausea
- mild OHSS
- vomiting
- palpitations
- period pain
- pelvic girdle pain/symphysis pubis dysfunction/hip pain/pelvis dysplasia
- placenta previa (grades 1 and 2)
- pre-eclampsia
- prolapse
- proteinuria
- rash
- reduced fetal movement
- relfux/heartburn/indigestion
- Group B Streptococcus infection
- urinary tract infection
- vaginal bleeding
- vaginal discharge
- vaginal infection
- viral infection.

Appendix 7 Expected serious adverse events

Expected SAEs are those events that are expected in the patient population or as a result of the routine care/treatment of a patient. These expected SAEs were collected as part of the trial and entered into the CRF, but were not reported to regulatory bodies (NHS REC/sponsor).

It is possible that during their pregnancy, participants will be admitted to hospital for treatment or monitoring of their pregnancy. Therefore, expected SAEs included hospitalisations for the following events.

Events related to the pregnancy/IVF therapy

Expected SAEs include admission to a hospital or other institution for general care not associated with any deterioration in condition including:

- hospitalisation for rest
- hospitalisation for observation or monitoring of pregnancy
- hospitalisation for maternal discomfort
- hyperemesis
- OHSS
- hypertensive disorders of pregnancy
- antepartum haemorrhage
- gestational diabetes mellitus (GDM)
- post partum haemorrhage
- placenta praevia (grades 3 and 4)
- accreta placenta
- placental abruption
- manual removal of placenta.

Appendix 8 Embryo grading techniques per site

Site ID	Site	NEQAS (pre April 2017)	Gardner blastocyst grading system	NEQAS (post April 2017)
ES01	Sheffield	S	S	P
ES02	Leicester	P	✗	S
ES03	Southampton	S	P	✗
ES04	Manchester	S	✗	P
ES05	Leeds	P	P	S
ES06	Dundee	S	✗	P
ES07	Birmingham	✗	S	P
ES08	Newcastle	P	✗	S
ES09	Liverpool	S	✗	P
ES10	Homerton	✗	✗	P
ES11	Guy's and St Thomas' Hospital	✗	P	✗
ES12	Nottingham	P	S	✗
ES13	Wrightington	S	✗	P
ES15	Oxford	S	P	✗
ES16	Gateshead	S	✗	P
ES17	South Tees	S	✗	P

✗, system not used; ID, identification number; P, primary grading system; S, some use of grading system.

Appendix 9 Embryo grading conversion charts

Blastocyst

Gardner blastocyst grading system[a]	New NEQAS[a]	Old NEQAS[a]	Category
A/A	A/A	5/3	Excellent
A/B, B/A	A/B, B/A	5/2, 4/3, 3/4, 4/4	Very good
B/B	B/B	4/2,	Good
A/C, C/A	BC/CB, A/C	3/3, 3/2	Fair + freezable
B/C, C/B	C/C, A/D	3/1, 4/1, 5/1	Fair
		2/3, 2/2, 1/3	
C/C, degree of expansion 2 and X/X	C/D, D/C, D/D, degree of expansion 2 and X/X	2/1, 1/1, 1/2	Poor
No grades provided (X/X/X)	No grades provided (X/X/X)	No grades provided (X/X/X)	Early blastocyst
Degree of expansion 1 and any other TE or ICM grade	Degree of expansion 1 and any other TE or ICM grade	Degree of expansion 1 and any other TE or ICM grade	

ICM, inner cell mass; TE, trophectoderm; X, no grade provided.
a Inner cell mass/trophectoderm.

Cleavage

Category	Embryo grading[a]
Excellent	**D2** 4/4/4, 3/4/4
	D3 8/4/4
Good	**D2** 5/4/4, 5/3/4, 5/4/3, 5/3/3, 4/3/4, 4/4/3, 4/3/3
	D3 10/4/4, 10/4/3, 10/3/4, 10/3/3, 9/4/4, 9/4/3, 9/3/4, 9/3/3, 8/4/3, 8/3/4, 8/3/3, 7/4/4, 7/4/3, 7/3/3, 6/4/4, 6/3/4, 6/4/3, 6/3/3, 7/3/4
Fair	**D2** 5/2/3, 5/2/4, 4/2/3, 3/3/4, 3/3/3, 3/2/3, 4/2/4
	D3 6/2/4, 8/4/2
Poor quality	**D2** all ≥ 6c combinations, 5/3/2, 5/2/2, 4/3/2, 4/2/2, 3/4/3, 3/3/2, 3/2/2
	D3 all ≥ 11c combinations, 10/3/2, 10/2/3, 10/2/2, 9/3/2, 9/2/3, 9/2/2, 8/3/2, 8/2/3, 8/2/2, 7/3/2, 7/2/3, 7/2/2, 6/3/2, 6/2/30, 6/2/2
Very poor quality	**D2** 5/2/1, 5/1/2, 5/1/1, 4/2/1, 4/1/2, 4/1/1, 3/2/1, 3/1/2, 3/1/1, 3/1/3
	D3 all -/1/1 combinations, 5/2/1
Slow	**D2** All 2c combinations
	D3 All ≤ 5 cell combinations except -/1/1, 5/4/4, 5/3/4, 5/4/3, 5/3/3, 5/2/3, 5/3/2, 5/2/2, 5/2/3
Arrested development	Graded as cleavage at day 5 of development

C, cell; D, day.
a cell number/shape score/fragmentation score.

Appendix 10 Example MEDLINE search strategy

1. exp embryo transfer/ or exp fertilization in vitro/ or exp sperm injections, intracytoplasmic/
2. embryo transfer$.tw.
3. in vitro fertili?ation.tw.
4. ivf-et.tw.
5. (ivf or et).tw.
6. icsi.tw.
7. intracytoplasmic sperm injection$.tw.
8. (blastocyst adj2 transfer$).tw.
9. exp reproductive techniques, assisted/ or insemination, artificial/ or exp insemination, artificial, heterologous/ or exp insemination, artificial, homologous/
10. assisted reproducti$.tw.
11. FET.tw.
12. or/1-11
13. (endometri$ adj5 scratch$).tw.
14. (endometri$ adj5 injur$).tw.
15. (endometri$ adj5 trauma$).tw.
16. (endometri$ adj5 biop$).tw.
17. (endometri$ adj5 harm$).tw.
18. (endometri$ adj5 damag$).tw.
19. (endometri$ adj5 inflammation).tw.
20. (endometri$ adj5 wound$).tw.
21. (endometri$ adj5 lesion$).tw.
22. (endometri$ adj5 insult$).tw.
23. or/13-22
24. 12 and 23
26. randomized controlled trial.pt.
27. controlled clinical trial.pt.
28. randomized.ab.
29. placebo.tw.
30. clinical trials as topic.sh.
31. randomly.ab.
32. trial.ti.
33. (crossover or cross-over or cross over).tw.

Appendix 11 Recruitment graph
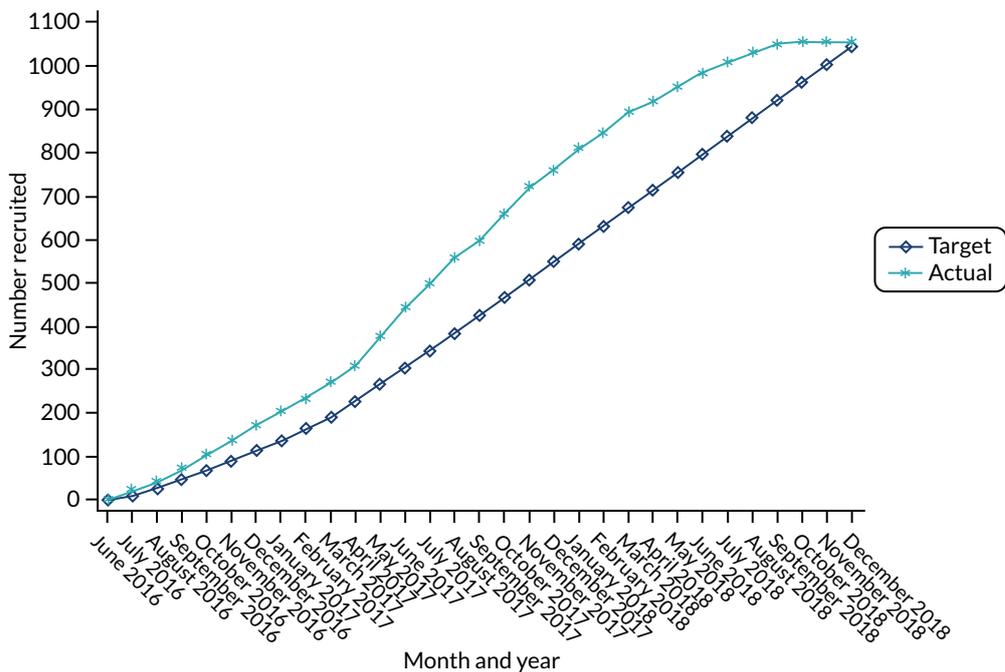
Appendix 12 Post-randomisation ineligibilities

Group	Category	Details
TAU	Randomisation error	Had previous endometrial trauma
ES	Randomisation error	Treated with ultra-long protocol
TAU	Potentially missed at screening	The participant was consented and randomised before the endometrial polyp developed, which was only noticed once she had started IVF cycle and therefore was put on hold until a polypectomy was performed. In the past, this could have been classed as mild endometrial trauma
ES	Potentially missed at screening	Endometrial polyp identified after starting IVF cycle; had embryos frozen and a polypectomy performed. This could have been classed as a mild endometrial trauma
TAU	Potentially missed at screening	The participant had a 5.2 cm fibroid at 12 weeks. There was no report from ACU when she had a baseline scan pre treatment of any evidence of a fibroid so there is a chance it was missed at baseline scan
TAU	Randomisation error	Had a BMI of 35 kg/m^2
ES	Randomisation error	At the point of randomisation, the participant was predicted to be aged $\geq$ 38 at egg collection
ES	Randomisation error	Participant received a myomectomy in 2017 (exact date unknown). The exact method of myomectomy unknown, so it is not known if this would constitute endometrial trauma
ES	Potentially missed at screening	The participant had a degenerating fibroid found in the uterus during pregnancy that was not identified during baseline scan prior to randomisation
ES	Randomisation error	At randomisation, site staff documented that fresh embryo transfer was expected but a reason for not undertaking fresh embryo transfer (see *Table 5*) indicated that FET was planned prior to randomisation

ACU, assisted conception unit.

Appendix 13 Self-reported pain rating of endometrial scratch procedure

Timing post ES procedure	Pain rating summary	
	Including ambiguous scores	**Excluding ambiguous scores**
Within 30 minutes	$n = 450$	$n = 450$
Mean (SD)	4.1 (2.4)	4.1 (2.4)
Median (IQR)	4.0 (2.0–6.0)	4.0 (2.0–6.0)
Minimum, maximum	0.0, 10.0	0.0, 10.0
Day 1	$n = 408$	$n = 402$
Mean (SD)	1.8 (2.5)	1.8 (2.5)
Median (IQR)	1.0 (0.0–3.0)	1.0 (0.0–3.0)
Minimum, maximum	0.0, 10.0	0.0, 10.0
Day 7	$n = 397$	$n = 390$
Mean (SD)	0.8 (2.1)	0.8 (2.0)
Median (IQR)	0.0 (0.0–0.0)	0.0 (0.0–0.0)
Minimum, maximum	0.0, 10.0	0.0, 10.0

The higher the score the greater the pain: 0 (no pain) to 10 (worst pain imaginable).

Appendix 14 Subgroup results of the effect of endometrial scratch on live birth (odds ratio scale)

Characteristics/subgroup	TAU n/N (%)	ES n/N (%)	Adjusted OR (95% CI)	Interaction test: p-value
Treatment protocol used				
Antagonist	118/297 (39.7)	116/291 (39.9)	1.02 (0.73 to 1.43)	
Long protocol	70/197 (35.5)	83/205 (40.5)	1.24 (0.82 to 1.88)	0.479
Fertilisation method				
IVF	105/264 (39.8)	110/241 (45.6)	1.30 (0.91 to 1.87)	
ICSI	80/205 (39.0)	84/209 (40.2)	1.08 (0.72 to 1.62)	
IVF and ICSI split	3/7 (42.9)	5/15 (33.3)	0.57 (0.08 to 3.90)	0.603
Embryo transferred				
Single	157/374 (42.0)	169/381 (44.4)	1.13 (0.84 to 1.51)	
Double	31/90 (34.4)	30/74 (40.5)	1.34 (0.70 to 2.57)	0.633
Nature of embryo transferred				
Fresh	173/425 (40.7)	183/416 (44.0)	1.16 (0.88 to 1.53)	
Frozen	15/39 (38.5)	16/39 (41.0)	1.27 (0.50 to 3.21)	0.857
Day of embryo transfer[a]				
2	8/27 (29.6)	5/29 (17.2)	0.56 (0.15 to 2.03)	
3	16/57 (28.1)	17/68 (25.0)	0.89 (0.40 to 2.02)	
4	6/11 (54.5)	2/6 (33.3)	0.37 (0.04 to 3.13)	
5	158/367 (43.1)	175/352 (49.7)	1.32 (0.98 to 1.78)	0.334
Cycle programming				
No	73/173 (42.2)	61/173 (35.3)	0.77 (0.49 to 1.20)	
Yes	45/124 (36.3)	55/118 (46.6)	1.55 (0.92 to 2.62)	0.045
History of miscarriage				
0–2	191/519 (36.8)	200/517 (38.7)	1.09 (0.85 to 1.41)	
≥ 3	4/6 (66.7)	2/6 (33.3)	0.26 (0.02 to 3.07)	0.256

a Day 6 excluded because there were only two women in the TAU group.

Appendix 15 Subgroup results of the effect of endometrial scratch on live birth (relative risk scale)

Characteristic/subgroup	TAU n/N (%)	ES n/N (%)	Adjusted RR (95% CI)	Interaction test: p-value
Treatment protocol				
Antagonist	118/297 (39.7)	116/291 (39.9)	1.02 (0.83 to 1.24)	
Long protocol	70/197 (35.5)	83/205 (40.5)	1.14 (0.89 to 1.45)	0.477
Fertilisation method				
IVF	105/264 (39.8)	110/241 (45.6)	1.16 (0.95 to 1.42)	
ICSI	80/205 (39.0)	84/209 (40.2)	1.05 (0.83 to 1.33)	
IVF and ICSI split	3/7 (42.9)	5/15 (33.3)	0.72 (0.24 to 2.17)	0.604
Embryo transferred				
Single	157/374 (42.0)	169/381 (44.4)	1.07 (0.91 to 1.26)	
Double	31/90 (34.4)	30/74 (40.5)	1.20 (0.82 to 1.76)	0.591
Nature of embryo transferred				
Fresh	173/425 (40.7)	183/416 (44.0)	1.09 (0.93 to 1.27)	
Frozen	15/39 (38.5)	16/39 (41.0)	1.15 (0.67 to 1.96)	0.846
Day of embryo transfer[a]				
2	8/27 (29.6)	5/29 (17.2)	0.62 (0.23 to 1.68)	
3	16/57 (28.1)	17/68 (25.0)	0.92 (0.51 to 1.63)	
4	6/11 (54.5)	2/6 (33.3)	0.58 (0.18 to 1.92)	
5	158/367 (43.1)	175/352 (49.7)	1.16 (0.99 to 1.35)	0.373
Cycle programming				
No	73/173 (42.2)	61/173 (35.3)	0.86 (0.66 to 1.12)	
Yes	45/124 (36.3)	55/118 (46.6)	1.29 (0.95 to 1.74)	0.047
History of miscarriage				
0–2	191/519 (36.8)	200/517 (38.7)	1.06 (0.91 to 1.23)	
≥ 3	4/6 (66.7)	2/6 (33.3)	0.53 (0.15 to 1.84)	0.282

a Day 6 excluded because there were only two women in the TAU group.

Appendix 16 Effects of endometrial scratch on the secondary outcomes in intention-to-treat population (worst case)

Secondary outcome	TAU (N = 525), n (%)	ES (N = 523), n (%)	Unadjusted treatment effect (95% CI)			p-value
			AD	OR	RR	
Implantation rate	258 (49.1)	253 (48.4)	−0.8% (−6.8% to 5.3%)	0.97 (0.76 to 1.24)	0.98 (0.87 to 1.11)	0.804
Clinical pregnancy rate	213 (40.6)	223 (42.6)	2.1% (−3.9% to 8.0%)	1.09 (0.85 to 1.39)	1.05 (0.91 to 1.21)	0.497
Miscarriage rate	43 (8.2)	32 (6.1)	−2.1% (−5.2% to 1.0%)	0.73 (0.45 to 1.17)	0.75 (0.48 to 1.16)	0.193
Multiple birth rate	11 (2.1)	6 (1.1)	−0.9% (−2.5% to 0.6%)	0.54 (0.20 to 1.48)	0.55 (0.20 to 1.47)	0.224
Preterm delivery rate[a]	20 (3.8)	14 (2.7)	−1.1% (−3.3% to 1.0%)	0.69 (0.35 to 1.39)	0.70 (0.36 to 1.38)	0.301
Ectopic pregnancy rate	2 (0.4)	1 (0.2)	−0.2% (−0.8% to 0.5%)	0.50 (0.05 to 5.54)	0.50 (0.05 to 5.52)	0.565
Stillbirth rate	1 (0.2)	1 (0.2)	0.0% (−0.5% to 0.5%)	1.00 (0.06 to 16.09)	1.00 (0.06 to 16.01)	0.998

a Two born babies had missing gestational age so were assumed not to be preterm deliveries.
p-values not adjusted for multiple hypothesis tests.
There were 20 out of 525 (3.8%) biochemical pregnancies in the TAU group and 11 out of 523 (2.1%) in the ES group.
There were no pregnancies of unknown location reported.

Appendix 17 Effects of endometrial scratch on negative secondary outcomes in pregnant women

Secondary outcome	TAU (N = 258), n (%)	ES (N = 253), n (%)	Unadjusted treatment effect (95% CI)			p-value
			AD	OR	RR	
Miscarriage rate	43 (16.7)	32 (12.6)	−4.0% (−10.1% to 2.1%)	0.72 (0.44 to 1.19)	0.76 (0.50 to 1.16)	0.199
Multiple birth rate	11 (4.3)	6 (2.4)	−1.9% (−5.0% to 1.2%)	0.55 (0.20 to 1.50)	0.56 (0.21 to 1.48)	0.233
Preterm delivery rate[a]	20 (7.8)	14 (5.5)	−2.2% (−6.5% to 2.1%)	0.70 (0.34 to 1.41)	0.71 (0.37 to 1.38)	0.314
Ectopic pregnancy rate	2 (0.8)	1 (0.4)	−0.4% (−1.7% to 0.9%)	0.51 (0.05 to 5.64)	0.51 (0.05 to 5.59)	0.574
Stillbirth rate	1 (0.4)	1 (0.4)	0.0% (−1.1% to 1.1%)	1.02 (0.06 to 16.39)	1.02 (0.06 to 16.22)	0.989

a Two born babies had missing gestational age so were assumed not to be preterm deliveries.
p-values not adjusted for multiple hypothesis tests.
There were 20 out of 258 (7.8%) biochemical pregnancies in the TAU group and 11 out of 253 (4.3%) in the ES group.
There were no pregnancies of unknown location reported.

Appendix 18 Forest plot of effects of endometrial scratch on secondary outcomes (relative risk scale)

Secondary outcome	TAU, *n/N* (%)	ES, *n/N* (%)		Unadjusted RR (95% CI)	*p*-value
ITT					
Implantation rate	258/525 (49.1)	253/523 (48.4)		0.98 (0.87 to 1.11)	0.804
Clinical pregnancy rate	213/525 (40.6)	223/523 (42.6)		1.05 (0.91 to 1.21)	0.497
Miscarriage rate	43/525 (8.2)	32/523 (6.1)		0.75 (0.48 to 1.16)	0.193
Multiple birth rate	11/525 (2.1)	6/523 (1.1)		0.55 (0.20 to 1.47)	0.224
Preterm delivery rate	20/525 (3.8)	14/523 (2.7)		0.70 (0.36 to 1.38)	0.301
Ectopic pregnancy rate	2/525 (0.4)	1/523 (0.2)		0.50 (0.05 to 5.52)	0.565
Stillbirth rate	1/525 (0.2)	1/523 (0.2)		1.00 (0.06 to 16.01)	0.998
Pregnant women (ITT)					
Miscarriage rate	43/258 (16.7)	32/253 (12.6)		0.76 (0.50 to 1.16)	0.199
Multiple birth rate	11/258 (4.3)	6/253 (2.4)		0.56 (0.21 to 1.48)	0.233
Preterm delivery rate	20/258 (7.8)	14/253 (5.5)		0.71 (0.37 to 1.38)	0.314
Ectopic pregnancy rate	2/258 (0.8)	1/253 (0.4)		0.51 (0.05 to 5.59)	0.574
Stillbirth rate	1/258 (0.4)	1/253 (0.4)		1.02 (0.06 to 16.22)	0.989

0.00 0.25 0.50 0.75 1.00 1.50 2.00 2.50 3.00 3.50

Appendix 19 All unexpected adverse events and serious adverse events during the entire trial

	TAU (N = 537)		ES (N = 458)		
AE classification	**n (%)**	**n/TPy (IR)**	**n (%)**	**n/TPy (IR)**	**IRR (95% CI)**
Unexpected AEs	94 (17.5)	125/334.1 (0.37)	90 (19.7)	120/295.0 (0.42)	1.15 (0.85 to 1.54)
SAEs	50 (9.3)	59/334.1 (0.17)	37 (8.1)	49/295.0 (0.16)	0.96 (0.63 to 1.46)
Unexpected SAEs	26 (4.8)	28/334.1 (0.08)	22 (4.8)	26/295.0 (0.09)	1.06 (0.60 to 1.88)
Expected SAEs	28 (5.2)	31/334.1 (0.09)	20 (4.4)	23/295.0 (0.08)	0.85 (0.48 to 1.50)
Unexpected AE category					
Pain	20 (3.7)	21	14 (3.1)	19	
OHSS	5 (0.9)	5	7 (1.5)	7	
GI-related issues	10 (1.9)	12	11 (2.4)	13	
Infection/inflammation	18 (3.4)	18	22 (4.8)	24	
Bleeding/blood-related events	35 (6.5)	39	21 (4.6)	22	
Urinary-related issues	7 (1.3)	7	3 (0.7)	3	
Cardiac-related issues	7 (1.3)	7	13 (2.8)	14	
Vaginal bleeding-related issues	0 (0.0)	0	1 (0.2)	1	
Pregnancy-specific issues	9 (1.7)	10	8 (1.7)	9	
Placenta-related issues	7 (1.3)	8	4 (0.9)	4	
Skin-related issues	5 (0.9)	5	0 (0.0)	0	
Events in baby/fetus	5 (0.9)	5	4 (0.9)	4	
Mental health	1 (0.2)	2	1 (0.2)	1	
Neuro-related issues	1 (0.2)	1	1 (0.2)	1	
Uterine abnormality	0 (0.0)	0	4 (0.9)	4	
Other	16 (3.0)	17	21 (4.6)	21	
Seriousness of SAE					
Life-threatening	5 (0.9)	5	2 (0.4)	2	
Inpatient hospitalisation	36 (6.7)	41	29 (6.3)	37	
Prolongs hospitalisation	12 (2.2)	12	8 (1.7)	10	
Persistent or significant disability/incapacity	1 (0.2)	1	0 (0.0)	0	
Frequency of SAE					
Isolated	43 (8.0)	48	34 (7.4)	39	
Intermittent	5 (0.9)	5	5 (1.1)	8	
Continuous	2 (0.4)	2	1 (0.2)	1	
Missing[a]	0 (0.0)	0	1 (0.2)	1	
Unknown	4 (0.7)	4	0 (0.0)	0	

AE classification	TAU (N = 537)		ES (N = 458)		IRR (95% CI)
	n (%)	n/TPy (IR)	n (%)	n/TPy (IR)	
Intensity of SAE					
Mild	18 (3.4)	18	7 (1.5)	7	
Moderate	29 (5.4)	33	23 (5.0)	31	
Severe	8 (1.5)	8	6 (1.3)	10	
Missing[a]	0 (0.0)	0	1 (0.2)	1	
Outcome of SAE					
Recovered	45 (8.4)	51	32 (7.0)	40	
Improved	7 (1.3)	7	5 (1.1)	7	
Unchanged	1 (0.2)	1	1 (0.2)	1	
Missing[a]	0 (0.0)	0	1 (0.2)	1	
Relationship to ES					
Possible	N/A	N/A	1 (0.2)	1	
Unlikely	N/A	N/A	10 (2.2)	11	
Unrelated	N/A	N/A	25 (5.5)	35	
Missing[a]	N/A	N/A	1 (0.2)	1	
Not assessable	N/A	N/A	1 (0.2)	1	

GI, gastrointestinal; N/A, not applicable; n/TPy (IR), number of all repeated events/total follow-up time in years (IR per person-year).
a The frequency, intensity, and outcome of one unexpected SAE that resulted in inpatient hospitalisation could not be ascertained (SAE was identified from the final trial questionnaire and the site could not contact the participant to ascertain the details).

Appendix 20 Admissions to hospital after accident and emergency visit

Reason for admission	TAU, *n*	ES, *n*	Total, *N*	Unit cost (£)[a]
OHSS	0	4	4	2033
Abdominal pain	1	2	2	2565
Bleeding	2	2	4	2033
Ectopic pregnancy	1	1	2	1868
Infection	2	2	4	2560
Baby unwell	4	1	5	2560
Complications before giving birth	0	6	6	2676
Complications with pregnancy	2	3	5	2033
Enlarged liver	1	0	1	2033
Gallstones	1	1	2	2609
Hand surgery	0	1	1	2255
Lumbago	0	1	1	2033
Vaginal abscesses	0	1	1	2033

a　Source of unit costs: Department of Health and Social Care Reference Costs (2018/19)[115]

Appendix 21 Planned hospital visits: day cases

Reason for visit	TAU, *n*	ES, *n*	Total, *N*	Unit cost (£)
A&E	2	0	2	166[a]
Abatacept infusions	1	0	1	147[a]
Amniotic fluid	0	1	1	133[a]
Antenatal clinic	3	8	11	331[a]
Assisted delivery	1	1	2	2112[a]
Asthma	1	0	1	242[a]
BCG vaccine	2	0	2	17[a]
Bleed comparison between participant and baby	0	1	1	133[a]
Bleeding	1	0	1	1221[a]
Blood test	11	15	26	2[a]
Blood transfusion	1	0	1	530[a]
Bowel biopsy	0	1	1	677[a]
Breast clinic	1	2	3	262[a]
Breast abscess: consultancy review	0	1	1	133[a]
Breastfeeding support	1	1	2	1319[a]
C-section	0	1	1	4186[a]
CTG monitoring of baby	1	0	1	532[a]
Colonoscopy	2	2	4	590[a]
Consultant	47	33	80	133[a]
Counsellor	1	0	1	16[b]
Dental surgery	1	0	1	1596[a]
Drug treatment for Crohn's disease	0	1	1	141[a]
ECG	0	2	2	73[a]
Early pregnancy unit	1	1	2	331[a]
Endocrinology appointment	6	2	8	152[a]
Epilepsy outpatient	0	1	1	587[a]
Eye unit	1	0	1	95[a]
Gallstones	1	0	1	2452[a]
Gestational diabetes	2	2	4	318[a]
Heart scan	1	0	1	73[a]
Hip scan	1	0	1	138[a]
Hysterosalpingogram	1	0	1	1073[a]
Hysterosalpingogram	0	1	1	1073[a]
Hysteroscopy	2	2	4	1073[a]
IVF	62	85	147	3218[c]

Reason for visit	TAU, n	ES, n	Total, N	Unit cost (£)
Induction	0	2	2	3390[a]
Jaundice clinic	1	0	1	2345[a]
Labour	2	2	4	2101[a]
Laparoscopy	13	7	20	2452[a]
Lumps in groin	1	0	1	2560[a]
Mastitis	0	1	1	2560[a]
Maternity triage	1	0	1	331[a]
Maxillofacial biopsy	1	0	1	126[a]
Medical care of MS	0	3	3	612[a]
Melanoma review	1	0	1	112[a]
Mental health	0	1	1	307[a]
Midwife	6	4	10	97[a]
Mole checked	1	2	3	618[a]
Nasal septum deviation straightening	0	1	1	778[a]
Neurosurgeon	0	1	1	167[a]
Nurse consultation	0	1	1	129[a]
Paediatrician	1	0	1	267[a]
Pancreatitis	0	1	1	1698[a]
Pelvic ultrasound	1	0	1	374[a]
Physiotherapy	3	0	3	55[a]
Post-natal check-up	4	0	4	2560[a]
Post-surgery review	0	1	1	144[a]
Renal clinic	0	1	1	410[a]
Review of child's pleural effusion	0	1	1	976[a]
Salpingectomy	2	1	3	1109[a]
Smear using colposcopy	0	2	2	2452[a]
Sweep	0	2	2	97[a]
Tongue tie clinic	1	2	3	2560[a]
Tongue tie cut (private)	1	0	1	2560[a]
UVB therapy	0	1	1	474[a]
Ultrasound: kidneys	1	0	1	55[a]
Ultrasound scan	43	47	90	365[a]
Unspecific bowel problem in child	0	1	1	208[a]
Upper eyelid blepharoplasty	0	1	1	95[a]
Urology	1	0	1	105[a]

BCG, bacillus Calmette–Guérin; CTG, cardiotocography; C-section, caesarean section; ECG, electrocardiogram;
MS, multiple sclerosis; UVB, ultra-violet B.
a Source of unit costs: Department of Health and Social Care Reference Costs (2018/19).[115]
b Source of unit costs: Agenda for Change 2018/19.[116]
c Source of unit cost: Endometrial Scratch Trial.

Appendix 22　Planned hospital visits: overnight stay

Reason for visit	TAU, *n*	ES, *n*	Total, *N*	Unit cost (£)[a]
Antenatal clinic	2	1	3	331
Baby unwell	0	2	2	2345
Bleeding	1	0	1	1221
C-section	11	10	21	4186
Cervical stitch	1	0	1	4915
Colonoscopy	0	1	1	590
Discharge + headaches	0	1	1	1221
Hysteroscopy	0	1	1	1073
Induction	17	19	36	3990
Infection	0	1	1	2345
Jaundice	1	0	1	2345
Labour	21	33	54	2101
Neonatal	0	2	2	4446
Post party sepsis	1	0	1	4239
Pregnancy related	1	0	1	1221
Salpingectomy	0	1	1	1109
Sepsis	0	1	1	4239
Surgery to hand	0	1	1	2201
Thrombocytopenia	1	0	1	1254

C-section, caesarean section.
a   Source of unit costs: Department of Health and Social Care Reference Costs (2018/19).[115]

Appendix 23 List of reasons for admission

	TAU, *n*	ES, *n*	Total, *N*	Unit cost
A&E	4	3	7	£166[a]
ACU	54	56	110	£1,831[a]
Acupuncture	30	30	60	£63[a]
Antenatal	4	3	7	£331[a]
Aromatherapy	0	1	1	£52[b]
Audiology	0	1	1	£200[a]
Bloods	1	2	3	£2[a]
Botox	1	0	1	£116[a]
Breast clinic	0	1	1	£1,572[a]
Breastfeeding clinic	7	6	13	£2,560[a]
Children's clinic	1	0	1	£172[a]
Chiropractor	6	1	7	£63[a]
Colposcopy	1	0	1	£590[a]
Consultant	8	6	14	£133[a]
Counselling	5	5	10	£15.92[c]
Day surgery	0	1	1	£778[a]
Dental	3	0	3	£104[d]
Dermatology	1	1	2	£112[a]
Endocrinology	3	1	4	£161[a]
Eye clinic	1	0	1	£95[a]
Flu jab	1	1	2	£9.44[e]
Fracture clinic	1	1	2	£167[a]
GP	1	1	2	£39.23[d]
Gynaecology	1	1	2	£133[a]
Haematology	1	0	1	£167[a]
Hand surgery	0	1	1	£2201[a]
Health visitor	0	1	1	£68[a]
Hypnotherapist	2	0	2	£80[f]
Jaundice	1	0	1	£2345[a]
Labour	0	1	1	£2101[a]
Massage	8	1	9	£63[a]
Mental health	1	0	1	£155[a]
Midwife	2	3	5	£63[a]
Miscarriage	0	1	1	£331[a]
NHS 111 telephone helpline	1	0	1	£7.33[a]
NHS Direct	1	0	1	£7.33[a]

	TAU, *n*	ES, *n*	Total, *N*	Unit cost
Neonatal	0	1	1	£2,345[a]
Neurology	2	0	2	£169[a]
OP: BMI	0	1	1	£84[a]
Obstetrician	3	0	3	£135[a]
Occupational health	0	3	3	£43[d]
Ophthalmic	0	1	1	£98[a]
Optician	1	0	1	£31.44[d]
Osteopathy	3	2	5	£55[a]
Outpatient	3	2	5	£127[a]
Physiotherapy	0	1	1	£55[a]
Podiatry	2	0	2	£54[a]
Post natal	2	1	3	£2560[a]
Psychotherapy	6	1	7	£337[a]
Radiology	1	1	2	£143[a]
Reflexology	6	5	11	£35[g]
Reiki	1	0	1	£40[h]
Respiratory medicine	0	1	1	£146[a]
Rheumatology	0	1	1	£89[a]
SCBU	1	0	1	£2560[a]
Smoking cessation	1	0	1	£36[d]
Speech and language	2	0	2	£107[a]
Surgical assessment unit	0	1	1	£133[a]
Tongue tie	0	2	2	£2560[a]
Ultrasound: kidney	0	1	1	£55[a]
Ultrasound: hand	0	1	1	£55[a]
Ultrasound scan	9	7	16	£365[a]
Yoga	0	1	1	£6

ACU, assisted conception unit; OP-BMI – outpatient body mass index; SCBU, special care baby unit.
a Source of unit costs: Department of Health and Social Care Reference Costs (2018/19).[115]
b Source of costs: private holistic treatment provider.[117]
c Source of unit costs: Agenda for Change 2018/19.[116]
d Source of unit costs: Curtis and Burns 2019.[106]
e Source of unit costs: NICE guidance.[118]
f Source of costs: private hypnotherapy provider.[119]
g Source of costs: private reflexology provider.[120]
h Source of costs: private reiki provider.[121]

Appendix 24 Frequency of NHS and private ultrasound scans

Scan type	Group	Baseline	3 months, *n* (%)	6 weeks post partum, *n* (%)
Ultrasound on NHS	TAU	0	145 (98.0)	72 (49.7)
	ES	0	168 (98.2)	93 (58.5)
Private ultrasound	TAU	0	50 (33.8)	11 (7.6)
	ES	0	50 (29.2)	14 (8.8)

Appendix 25　Participants' and companions' occupations

Occupation/field of occupation	TAU		ES	
	Participant	Companion	Participant	Companion
Account manager	5	1	1	1
Accountant	2	2	3	1
Accounts assistant	1	1	0	0
Administrator	5	1	3	0
Advisor	0	0	1	0
Analyst	0	1	0	1
Architect	0	0	1	0
Armed forces	0	2	1	2
Assistant buyer	0	0	1	0
Assistant manager	1	0	0	0
Beauty therapist	1	0	1	0
BSL interpreter	0	0	1	0
Business analyst	1	1	0	1
Business development	1	3	3	1
Business owner	1	2	1	1
Call centre	1	1	1	1
Care assistant	3	0	5	2
Catering/chef	3	2	1	2
Charities	0	0	2	0
Civil servant	3	0	5	1
Claims handler	1	0	0	0
Class 2 driver	0	4	0	3
Cleaner	2	0	0	0
Clinical psychologist	3	0	0	0
Commercial officer	1	0	0	0
Company director	0	2	1	2
Council worker	3	1	0	1
Counsellor	1	0	1	0
CRM assistant	0	0	1	0
Customer experience	4	0	3	0
Dental hygienist	1	0	0	0
Dental nurse	1	0	1	0
Dental receptionist	0	0	1	0
Dental technician	1	0	1	0

Occupation/field of occupation	TAU		ES	
	Participant	Companion	Participant	Companion
Dentist	0	0	3	1
Dietitian	2	0	1	0
Dispenser hospital	0	0	1	0
Doctor/GP	2	2	4	0
Dog groomer/vet nurse	0	0	1	1
Engineer	0	11	1	9
Electrician	0	4	0	1
Finance	2	6	1	3
Firefighter	0	1	1	1
Fraud investigator	0	0	1	1
Graphic designer	0	1	1	1
Hairdresser/stylist/barber	1	0	2	1
Health visitor	0	0	1	0
Higher education	0	0	1	0
Housing development	1	0	0	0
HR	3	0	1	0
Insolvency practice	0	0	1	0
Insurance	0	1	0	1
Interior designer	1	0	0	0
IT	0	6	4	10
Journalist/illustrator/writer	0	2	0	1
Lecturer	0	1	2	1
Legal secretary	2	0	0	0
Logistics operative	1	0	0	0
Management consultant	1	0	0	0
Manager	14	13	11	15
Managing director	0	0	1	3
Manual	0	19	1	17
Marketing	2	0	4	1
Mental health worker	1	0	1	0
Midwife	2	0	1	1
Mortgage advisor	1	0	0	0
Nurse	4	2	9	0
Nursery nurse	2	0	1	0
Occupational therapist	1	0	0	1
Optical/optometrist	2	1	0	1
Orthoptisa	1	0	0	0
Osteopath	0	0	1	0

Occupation/field of occupation	TAU		ES	
	Participant	Companion	Participant	Companion
Parent liaison officer	1	0	0	0
Pharmacist	1	1	0	1
Pharmacy assistant	1	0	2	0
Phlebotomist	1	0	0	0
Physiotherapist	2	0	0	0
Play assistant	1	0	1	0
Police investigator	0	0	1	0
Police officer	1	0	2	5
Production operation	0	0	1	0
Programme manager	2	3	2	1
Public relations	0	0	1	0
Receptionist	0	0	1	0
Recruitment	1	1	0	3
Retail	1	2	2	1
Retail manager	1	2	1	1
Sales	3	4	2	4
Sales director	1	2	0	0
Sales manger	0	1	1	2
Scientist	0	0	3	1
Sedation co-ordinator	0	0	1	0
Self-employed	3	9	1	4
Skills training	1	0	1	0
Social worker	2	1	1	0
Solicitor	0	0	0	2
Speech and language	0	0	1	0
Supervisor	2	0	0	0
Support worker	0	2	1	0
Surveyor	0	0	0	1
Talent acquisition officer	0	0	1	0
Teacher	13	1	12	5
Teaching assistant	1	0	1	0
Town planner	0	0	1	0
Trainer	1	1	0	1
Translator	0	0	0	1
Transport worker	0	1	2	4
TV/entertainment	0	1	1	1
Youth support co-ordinator	1	0	0	0

BSL, British sign language; CRM, customer relationship management; HR, human resources; IT, information technology.

Appendix 26 Studies excluded at full-paper review stage

Study (first author)	Reason for exclusion
Shawki[72]	Abstract with insufficient detail
Shahrokh-Tehraninejad 2015[122]	Not in English
Abdelmagied[123]	Not RCT
Chang[124]	Not RCT
Chawla[125]	Not RCT
Du[126]	Not RCT
Lensen[11]	Not RCT
Lensen[127]	Not RCT
Liang[128]	Not RCT
McKenzie[129]	Not RCT
Mitic[130]	Not RCT
Moghissi[131]	Not RCT
Mol[132]	Not RCT
Nastri[133]	Not RCT
Nayar[134]	Not RCT
Shokeir[135]	Not RCT
Siristatidis[136]	Not RCT
Vanvitelli[137]	Not RCT
Wise[138]	Not RCT
Wise[139]	Not RCT
Zhou[140]	Not RCT
Checa[70]	Results not presented
Genod[141]	Results not presented
Geslevich[68]	Results not presented
Lee[142]	Results not presented
Marsh[69]	Results not presented
Martins[143]	Results not presented
Martins[144]	Results not presented
Shehata[71]	Results not presented
Bahaa[145]	Wrong intervention
Checa[146]	Wrong intervention
Gibreel[147]	Wrong intervention
Gibreel[148]	Wrong intervention
Movahedi[149]	Wrong intervention
Shokeir[150]	Wrong intervention

Study (first author)	Reason for exclusion
Soliman[151]	Wrong intervention
Vilela[152]	Wrong intervention
Xu[153]	Wrong intervention
Yoldemir[96]	Wrong intervention
Aflatoonian[73]	Wrong population
Alalfy[154]	Wrong population
Baum[155]	Wrong population
Chi[156]	Wrong population
Farrag[157]	Wrong population
Frantz[74]	Wrong population
Gomaa[158]	Wrong population
Guven[79]	Wrong population
Hare[159]	Wrong population
Hebeisha[80]	Wrong population
Helmy[160]	Wrong population
Hilton[81]	Wrong population
Hurst[78]	Wrong population
Kandavel[161]	Wrong population
Levin[162]	Wrong population
Maged[163]	Wrong population
Mak[164]	Wrong population
Mehrafza[82]	Wrong population
Merriam[83]	Wrong population
Olesen[165]	Wrong population
Parsaeian[166]	Wrong population
Salama[167]	Wrong population
Sherbiny[76]	Wrong population
Sherif[84]	Wrong population
Shokeir[168]	Wrong population
Shriver[77]	Wrong population
Spandorfer[85]	Wrong population
Sun[169]	Wrong population
Vidal[86]	Wrong population
Zayed[170]	Wrong population
Zhao[75]	Trial never commenced

Appendix 27 Forest plot of live birth rates, excluding Mahran *et al.*[20]

In this figure, 'df' is defined as 'degrees of freedom' and 'M–H' is defined as 'Mantel–Haenszel test'.

Study or subgroup	ES Events	ES Total	Control Events	Control Total	Weight (%)	OR M–H, fixed, 95% CI	OR M–H, fixed, 95% CI
Eskew[21]	19	39	19	27	4.2	0.40 (0.14 to 1.13)	
Izquierdo Rodriguez[22]	43	77	36	78	5.7	1.48 (0.78 to 2.78)	
Lensen[14]	81	298	74	278	20.3	1.03 (0.71 to 1.49)	
Liu[18]	25	70	29	72	6.7	0.82 (0.42 to 1.62)	
Mackens[28]	23	70	29	78	6.7	0.83 (0.42 to 1.63)	
Nastri[87]	2	8	4	10	1.0	0.50 (0.07 to 3.85)	
Polanski[88]	25	51	23	60	3.9	1.55 (0.73 to 3.30)	
Pye[35]	202	523	195	525	43.4	1.06 (0.83 to 1.37)	
Yeung[89]	32	105	32	104	8.1	0.99 (0.55 to 1.78)	
Total (95% CI)		1241		1232	100.0	1.03 (0.87 to 1.21)	
Total events	452		441				

Heterogeneity: $\chi^2 = 16.93$, df = 8 ($p = 0.54$); $I^2 = 0\%$
Test for overall effect: $Z = 0.33$ ($p = 0.74$)

0.1 0.2 0.5 1.0 2.0 5.0 10.0
Favours control   Favours ES

Appendix 28 Forest plot assessing the timing of endometrial scratch on live birth rates including all trials

In this figure, 'df' is defined as 'degrees of freedom' and 'M–H' is defined as 'Mantel–Haenszel test'.

Study or subgroup	ES Events	Total	Control Events	Total	Weight (%)	OR M–H, random, 95% CI	OR M–H, random, 95% CI
ES in menstrual cycle in which IVF is undertaken							
Mackens[28]	23	70	29	78	10.2	0.83 (0.42 to 1.63)	
Subtotal (95% CI)		70		78	10.2	0.83 (0.42 to 1.63)	
Total events	23		29				
Heterogeneity: not applicable							
Test for overall effect: $Z = 0.55$ ($p = 0.58$)							
ES in menstrual cycle prior to which IVF is undertaken							
Eskew[21]	19	39	19	27	7.6	0.40 (0.14 to 1.13)	
Izquierdo Rodriguez[22]	43	77	36	78	10.5	1.48 (0.78 to 2.78)	
Lensen[14]	81	298	74	278	12.4	1.03 (0.71 to 1.49)	
Liu[18]	25	70	29	72	10.2	0.82 (0.42 to 1.62)	
Mahran[20]	134	209	56	209	12.1	4.88 (3.22 to 7.40)	
Nastri[87]	2	8	4	10	3.3	0.50 (0.07 to 3.85)	
Polanski[88]	25	51	23	60	9.6	1.55 (0.73 to 3.30)	
Pye[35]	202	523	195	525	13.1	1.06 (0.83 to 1.37)	
Yeung[89]	32	105	32	104	10.9	0.99 (0.55 to 1.78)	
Subtotal (95% CI)		1380		1363	89.8	1.21 (0.76 to 1.92)	
Total events	563		468				
Heterogeneity: $\tau^2 = 0.37$; $\chi^2 = 51.49$, df = 8 ($p < 0.00001$); $I^2 = 84\%$							
Test for overall effect: $Z = 0.80$ ($p = 0.42$)							
Total (95% CI)		1450		1441	100.0	1.17 (0.76 to 1.79)	
Total events	586		467				
Heterogeneity: $\tau^2 = 0.35$; $\chi^2 = 53.15$, df = 9 ($p < 0.00001$); $I^2 = 83\%$							
Test for overall effect: $Z = 0.70$ ($p = 0.48$)							
Test for subgroup differences: $\chi^2 = 0.82$, df = 1 ($p = 0.37$); $I^2 = 0\%$							

0.1 0.2 0.5 1.0 2.0 5.0 10.0
Favours control Favours ES

Appendix 29 Forest plot assessing the timing of endometrial scratch on live birth rates and excluding Mahran *et al.*,[20] using a random-effects model

In this figure, 'df' is defined as 'degrees of freedom' and 'M–H' is defined as 'Mantel–Haenszel test'.

Study or subgroup	ES Events	Total	Control Events	Total	Weight (%)	OR M–H, random, 95% CI	OR M–H, random, 95% CI
ES in menstrual cycle in which IVF is undertaken							
Mackens[28]	23	70	29	78	6.0	0.83 (0.42 to 1.63)	
Subtotal (95% CI)		70		78	6.0	0.83 (0.42 to 1.63)	
Total events	23		29				
Heterogeneity: not applicable							
Test for overall effect: Z = 0.55 (p = 0.58)							
ES in menstrual cycle prior to which IVF is undertaken							
Eskew[21]	19	39	19	27	2.6	0.40 (0.14 to 1.13)	
Izquierdo Rodriguez[22]	43	77	36	78	6.9	1.48 (0.78 to 2.78)	
Lensen[14]	81	298	74	278	20.4	1.03 (0.71 to 1.49)	
Liu[18]	25	70	29	72	6.0	0.82 (0.42 to 1.62)	
Nastri[87]	2	8	4	10	0.7	0.50 (0.07 to 3.85)	
Polanski[88]	25	51	23	60	4.8	1.55 (0.73 to 3.30)	
Pye[35]	202	523	195	525	44.5	1.06 (0.83 to 1.37)	
Yeung[89]	32	105	32	104	8.0	0.99 (0.55 to 1.78)	
Subtotal (95% CI)		1171		1154	94.0	1.04 (0.88 to 1.24)	
Total events	429		412				
Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 6.50$, df = 8 (p = 0.48); $I^2 = 0\%$							
Test for overall effect: Z = 0.50 (p = 0.62)							
Total (95% CI)		1241		1232	100.0	1.03 (0.87 to 1.22)	
Total events	452		441				
Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 6.93$, df = 8 (p = 0.54); $I^2 = 0\%$							
Test for overall effect: Z = 0.35 (p = 0.73)							
Test for subgroup differences: $\chi^2 = 0.43$, df = 1 (p = 0.51); $I^2 = 0\%$							

0.2   0.5   1.0   2.0   5.0
Favours control   Favours ES

Appendix 30 Forest plot assessing the timing of endometrial scratch on live birth rates and excluding Mahran et al.,[20] using a fixed-effects model

In this figure, 'df' is defined as 'degrees of freedom' and 'M–H' is defined as 'Mantel–Haenszel test'.

Study or subgroup	ES Events	ES Total	Control Events	Control Total	Weight (%)	OR M–H, fixed, 95% CI	OR M–H, fixed, 95% CI
ES in menstrual cycle in which IVF is undertaken							
Mackens[28]	23	70	29	78	6.7	0.83 (0.42 to 1.63)	
Subtotal (95% CI)		70		78	6.7	0.83 (0.42 to 1.63)	
Total events	23		29				
Heterogeneity: not applicable							
Test for overall effect: $Z = 0.55$ ($p = 0.58$)							
ES in menstrual cycle prior to which IVF is undertaken							
Eskew[21]	19	39	19	27	4.2	0.40 (0.14 to 1.13)	
Izquierdo Rodriguez[22]	43	77	36	78	5.7	1.48 (0.78 to 2.78)	
Lensen[14]	81	298	74	278	20.3	1.03 (0.71 to 1.49)	
Liu[18]	25	70	29	72	6.7	0.82 (0.42 to 1.62)	
Nastri[87]	2	8	4	10	1.0	0.50 (0.07 to 3.85)	
Polanski[88]	25	51	23	60	3.9	1.55 (0.73 to 3.30)	
Pye[35]	202	523	195	525	43.4	1.06 (0.83 to 1.37)	
Yeung[89]	32	105	32	104	8.1	0.99 (0.55 to 1.78)	
Subtotal (95% CI)		1171		1154	93.3	1.04 (0.88 to 1.24)	
Total events	429		412				
Heterogeneity: $\chi^2 = 6.50$, df = 7 ($p = 0.48$); $I^2 = 0\%$							
Test for overall effect: $Z = 0.48$ ($p = 0.63$)							
Total (95% CI)		1241		1232	100.0	1.03 (0.87 to 1.21)	
Total events	452		441				
Heterogeneity: $\chi^2 = 6.93$, df = 8 ($p < 0.54$); $I^2 = 0\%$							
Test for overall effect: $Z = 0.33$ ($p = 0.74$)							
Test for subgroup differences: $\chi^2 = 0.42$, df = 1 ($p = 0.52$); $I^2 = 0\%$							

0.2　0.5　1.0　2.0　5.0
Favours control　Favours ES

Appendix 31 Forest plot of clinical pregnancy rates excluding Mahran *et al.*[20]

In this figure, 'df' is defined as 'degrees of freedom' and 'M–H' is defined as 'Mantel–Haenszel test'.

Study or subgroup	ES Events	ES Total	Control Events	Control Total	Weight (%)	OR M–H, fixed, 95% CI	OR M–H, fixed, 95% CI
Eskew[21]	21	39	19	27	3.2	0.49 (0.17 to 1.39)	
Izquierdo Rodriguez[22]	48	77	44	78	5.0	1.28 (0.67 to 2.43)	
Karimzade[17]	9	77	26	79	6.9	0.27 (0.12 to 0.62)	
Lensen[14]	94	298	87	278	18.9	1.01 (0.71 to 1.44)	
Liu[18]	29	70	30	72	5.3	0.99 (0.51 to 1.93)	
Mackens[28]	35	70	33	78	4.8	1.36 (0.71 to 2.61)	
Maged[19]	61	150	41	150	7.5	1.82 (1.12 to 2.96)	
Nastri[87]	3	8	4	10	0.7	0.90 (0.13 to 6.08)	
Polanski[88]	33	51	34	60	3.4	1.40 (0.65 to 3.02)	
Pye[35]	223	523	213	525	37.4	1.09 (0.85 to 1.39)	
Yeung[89]	42	105	38	104	7.0	1.16 (0.66 to 2.02)	
Total (95% CI)		1468		1461	100.0	1.08 (0.93 to 1.26)	
Total events	598		569				

Heterogeneity: $\chi^2 = 18.67$, df = 10 ($p = 0.04$); $I^2 = 46\%$
Test for overall effect: $Z = 1.06$ ($p = 0.29$)

0.1 0.2 0.5 1.0 2.0 5.0 10.0
Favours control   Favours ES

Appendix 32 Forest plot of clinical pregnancy rates by subgroup (timing of endometrial scratch)

In this figure, 'df' is defined as 'degrees of freedom' and 'M–H' is defined as 'Mantel–Haenszel test'.

Study or subgroup	ES Events	ES Total	Control Events	Control Total	Weight (%)	OR M–H, random, 95% CI
ES in menstrual cycle in which IVF is undertaken						
Karimzade[17]	9	77	26	79	7.3	0.27 (0.12 to 0.62)
Mackens[28]	35	70	33	78	8.6	1.36 (0.71 to 2.61)
Subtotal (95% CI)		147		157	15.8	0.62 (0.13 to 3.04)
Total events	44		59			
Heterogeneity: $\tau^2 = 1.17$; $\chi^2 = 8.99$, df = 1 ($p = 0.003$); $I^2 = 89\%$						
Test for overall effect: $Z = 0.59$ ($p = 0.56$)						
ES in menstrual cycle prior to which IVF is undertaken						
Eskew[21]	21	39	19	27	6.1	0.49 (0.17 to 1.39)
Izquierdo Rodriguez[22]	48	77	44	78	8.6	1.28 (0.67 to 2.43)
Lensen[14]	94	298	87	278	10.5	1.01 (0.71 to 1.44)
Liu[18]	29	70	30	72	8.4	0.99 (0.51 to 1.93)
Maged[19]	61	150	41	150	9.7	1.82 (1.12 to 2.96)
Mahran[20]	147	209	71	209	10.1	4.61 (3.05 to 6.96)
Nastri[87]	3	8	4	10	2.9	0.90 (0.13 to 6.08)
Polanski[88]	33	51	34	60	7.7	1.40 (0.65 to 3.02)
Pye[35]	223	523	213	525	11.0	1.09 (0.85 to 1.39)
Yeung[89]	42	105	38	104	9.2	1.16 (0.66 to 2.02)
Subtotal (95% CI)		1530		1513	84.2	1.33 (0.91 to 1.95)
Total events	701		581			
Heterogeneity: $\tau^2 = 0.26$; $\chi^2 = 46.02$, df = 9 ($p < 0.00001$); $I^2 = 80\%$						
Test for overall effect: $Z = 1.48$ ($p = 0.14$)						
Total (95% CI)		1677		1670	100.0	1.18 (0.82 to 1.72)
Total events	745		640			
Heterogeneity: $\tau^2 = 0.31$; $\chi^2 = 59.87$, df = 11 ($p < 0.00001$); $I^2 = 82\%$						
Test for overall effect: $Z = 0.88$ ($p = 0.38$)						
Test for subgroup differences: $\chi^2 = 0.84$, df = 1 ($p = 0.36$); $I^2 = 0\%$						

OR M–H, random, 95% CI: 0.01 — 0.10 — 1.00 — 10.00 — 100.00
Favours control Favours ES

Appendix 33 Forest plot of clinical pregnancy rates by subgroup (timing of endometrial scratch) with Mahran *et al.*[20] excluded, using a random-effects model

In this figure, 'df' is defined as 'degrees of freedom' and 'M–H' is defined as 'Mantel–Haenszel test'.

Study or subgroup	ES Events	ES Total	Control Events	Control Total	Weight (%)	OR M–H, random, 95% CI	OR M–H, random, 95% CI
ES in menstrual cycle in which IVF is undertaken							
Karimzade[17]	9	77	26	79	6.0	0.27 (0.12 to 0.62)	
Mackens[28]	35	70	33	78	8.6	1.36 (0.71 to 2.61)	
Subtotal (95% CI)		147		157	14.6	0.62 (0.13 to 3.04)	
Total events	44		59				
Heterogeneity: $\tau^2 = 1.17$; $\chi^2 = 8.99$, df = 1 ($p = 0.003$); $I^2 = 89\%$							
Test for overall effect: $Z = 0.59$ ($p = 0.56$)							
ES in menstrual cycle prior to which IVF is undertaken							
Eskew[21]	21	39	19	27	4.3	0.49 (0.17 to 1.39)	
Izquierdo Rodriguez[22]	48	77	44	78	8.7	1.28 (0.67 to 2.43)	
Lensen[14]	94	298	87	278	15.3	1.01 (0.71 to 1.44)	
Liu[18]	29	70	30	72	8.3	0.99 (0.51 to 1.93)	
Maged[19]	61	150	41	150	11.8	1.82 (1.12 to 2.96)	
Nastri[87]	3	8	4	10	1.5	0.90 (0.13 to 6.08)	
Polanski[88]	33	51	34	60	6.8	1.40 (0.65 to 3.02)	
Pye[35]	223	523	213	525	18.5	1.09 (0.85 to 1.39)	
Yeung[89]	42	105	38	104	10.2	1.16 (0.66 to 2.02)	
Subtotal (95% CI)		1321		1304	85.4	1.13 (0.97 to 1.33)	
Total events	554		510				
Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 7.32$, df = 8 ($p = 0.50$); $I^2 = 0\%$							
Test for overall effect: $Z = 1.55$ ($p = 0.12$)							
Total (95% CI)		1468		1461	100.0	1.06 (0.84 to 1.35)	
Total events	598		569				
Heterogeneity: $\tau^2 = 0.07$; $\chi^2 = 18.67$, df = 10 ($p = 0.04$); $I^2 = 46\%$							
Test for overall effect: $Z = 0.51$ ($p = 0.61$)							
Test for subgroup differences: $\chi^2 = 0.55$, df = 1 ($p = 0.46$); $I^2 = 0\%$							

Scale: 0.01 0.10 1.00 10.00 100.00

Favours control Favours ES

Appendix 34 Forest plot of clinical pregnancy rates by subgroup (timing of endometrial scratch) with Mahran *et al.*[20] excluded, using a fixed-effects model

In this figure, 'df' is defined as 'degrees of freedom' and 'M–H' is defined as 'Mantel–Haenszel test'.

Study or subgroup	ES Events	Total	Control Events	Total	Weight (%)	OR M–H, fixed, 95% CI
ES in menstrual cycle in which IVF is undertaken						
Karimzade[17]	9	77	26	79	6.9	0.27 (0.12 to 0.62)
Mackens[28]	35	70	33	78	4.8	0.36 (0.71 to 2.61)
Subtotal (95% CI)		147		157	11.7	0.72 (0.44 to 1.17)
Total events	44		59			
Heterogeneity: $\chi^2 = 8.99$, df = 1 ($p = 0.003$); $I^2 = 89\%$						
Test for overall effect: $Z = 1.34$ ($p = 0.18$)						
ES in menstrual cycle prior to which IVF is undertaken						
Eskew[21]	21	39	19	27	3.2	0.49 (0.17 to 1.39)
Izquierdo Rodriguez[22]	48	77	44	78	5.0	1.28 (0.67 to 2.43)
Lensen[14]	94	298	87	278	18.9	1.01 (0.71 to 1.44)
Liu[18]	29	70	30	72	5.3	0.99 (0.51 to 1.93)
Maged[19]	61	150	41	150	7.5	1.82 (1.12 to 2.96)
Nastri[87]	3	8	4	10	0.7	0.90 (0.13 to 6.08)
Polanski[88]	33	51	34	60	3.4	1.40 (0.65 to 3.02)
Pye[35]	223	523	213	525	37.4	1.09 (0.85 to 1.39)
Yeung[89]	42	105	38	104	7.0	1.16 (0.66 to 2.02)
Subtotal (95% CI)		1321		1304	88.3	1.13 (0.97 to 1.33)
Total events	554		510			
Heterogeneity: $\chi^2 = 7.32$, df = 8 ($p = 0.50$); $I^2 = 0\%$						
Test for overall effect: $Z = 1.55$ ($p = 0.12$)						
Total (95% CI)		1468		1461	100.0	1.08 (0.93 to 1.26)
Total events	598		569			
Heterogeneity: $\chi^2 = 18.67$, df = 10 ($p = 0.04$); $I^2 = 46\%$						
Test for overall effect: $Z = 1.06$ ($p = 0.29$)						
Test for subgroup differences: $\chi^2 = 3.07$, df = 1 ($p = 0.08$); $I^2 = 67.5\%$						

Appendix 35 Forest plot of all trials reporting ongoing pregnancy rates

In this figure, 'df' is defined as 'degrees of freedom' and 'M–H' is defined as 'Mantel–Haenszel test'.

Study or subgroup	ES Events	Total	Control Events	Total	Weight (%)	OR M–H, random, 95% CI	OR M–H, random, 95% CI
Izquierdo Rodriguez[22]	43	77	36	78	24.6	1.48 (0.78 to 2.78)	
Karimzade[17]	7	77	23	79	18.1	0.24 (0.10 to 0.61)	
Lensen[14]	85	298	80	278	31.5	0.99 (0.69 to 1.42)	
Yeung[89]	33	105	32	104	25.8	1.03 (0.57 to 1.85)	
Total (95% CI)		557		539	100.0	0.86 (0.49 to 1.48)	
Total events	168		171				

Heterogeneity: $\tau^2 = 0.22$; $\chi^2 = 10.49$, df = 3 ($p = 0.01$); $I^2 = 71\%$
Test for overall effect: $Z = 0.56$ ($p = 0.58$)
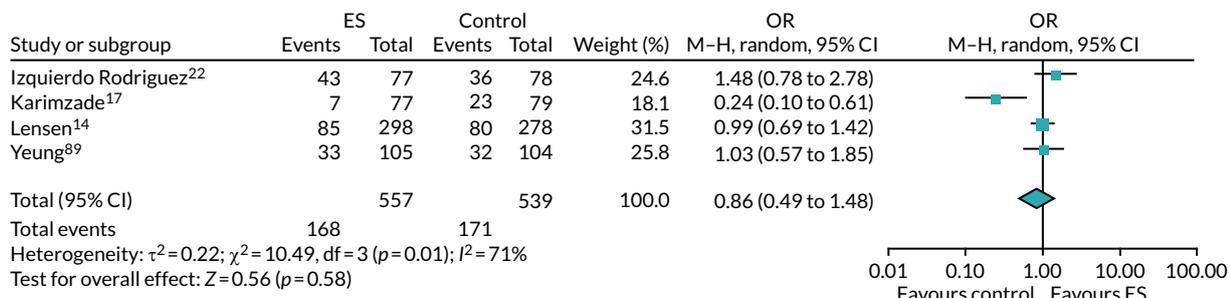
Appendix 36 Forest plot of all trials reporting ongoing pregnancy rates, stratified by the timing of endometrial scratch

In this figure, 'df' is defined as 'degrees of freedom' and 'M–H' is defined as 'Mantel–Haenszel test'.

Study or subgroup	ES Events	ES Total	Control Events	Control Total	Weight (%)	OR M–H, random, 95% CI	OR M–H, random, 95% CI
ES in menstrual cycle in which IVF was undertaken							
Karimzade[17]	7	77	23	79	18.1	0.24 (0.10 to 0.61)	
Subtotal (95% CI)		77		79	18.1	0.24 (0.10 to 0.61)	
Total events	7		23				
Heterogeneity: not applicable							
Test for overall effect: $Z = 3.02$ ($p = 0.003$)							
ES in menstrual cycle prior to which IVF is undertaken							
Izquierdo Rodriguez[22]	43	77	36	78	24.6	1.48 (0.78 to 2.78)	
Lensen[14]	85	298	80	278	31.5	0.99 (0.69 to 1.42)	
Yeung[89]	33	105	32	104	25.8	1.03 (0.57 to 1.85)	
Subtotal (95% CI)		480		460	81.9	1.08 (0.82 to 1.42)	
Total events	161		148				
Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 1.19$, df = 2 ($p = 0.55$); $I^2 = 0\%$							
Test for overall effect: $Z = 0.52$ ($p = 0.60$)							
Total (95% CI)		557		539	100.0	0.86 (0.49 to 1.48)	
Total events	168		171				
Heterogeneity: $\tau^2 = 0.22$; $\chi^2 = 10.49$, df = 3 ($p = 0.01$); $I^2 = 71\%$							
Test for overall effect: $Z = 0.56$ ($p = 0.58$)							
Test for subgroup differences: $\chi^2 = 9.27$, df = 1 ($p = 0.002$); $I^2 = 89.2\%$							

Scale: 0.01 0.10 1.00 10.00 100.00
Favours control   Favours ES

Appendix 37 Forest plot of all trials reporting miscarriage rates

In this figure, 'df' is defined as 'degrees of freedom' and 'M–H' is defined as 'Mantel–Haenszel test'.

Study or subgroup	ES Events	Total	Control Events	Total	Weight (%)	OR M–H, fixed, 95% CI	OR M–H, fixed, 95% CI
Eskew[21]	4	21	0	19	0.5	10.03 (0.50 to 199.86)	
Izquierdo Rodriguez[22]	5	48	8	44	8.8	0.52 (0.16 to 1.74)	
Lensen[14]	4	85	5	80	5.8	0.74 (0.19 to 2.86)	
Liu[18]	1	29	1	30	1.1	1.04 (0.06 to 17.38)	
Mackens[28]	11	35	3	33	2.5	4.58 (1.15 to 18.31)	
Maged[19]	11	61	9	41	10.4	0.78 (0.29 to 2.10)	
Mahran[20]	13	147	14	71	20.3	0.39 (0.17 to 0.89)	
Nastri[87]	1	3	0	4	0.3	5.40 (0.15 to 188.83)	
Pye[35]	32	223	43	213	44.4	0.66 (0.40 to 1.09)	
Yeung[89]	13	42	7	39	5.9	2.05 (0.72 to 5.84)	
Total (95% CI)		694		574	100.0	0.86 (0.63 to 1.17)	
Total events	95		90				

Heterogeneity: $\chi^2 = 17.14$, df = 9 ($p = 0.05$); $I^2 = 47\%$
Test for overall effect: $Z = 0.96$ ($p = 0.34$)

Favours ES   Favours control
0.01   0.10   1.00   10.00   100.00

Appendix 38 Forest plot of trials reporting miscarriage rates, with Mahran *et al.*[20] excluded

In this figure, 'df' is defined as 'degrees of freedom' and 'M–H' is defined as 'Mantel–Haenszel test'.

Study or subgroup	ES Events	Total	Control Events	Total	Weight (%)	OR M–H, random, 95% CI
Eskew[21]	4	21	0	19	3.1	10.03 (0.50 to 199.86)
Izquierdo Rodriguez[22]	5	48	8	44	12.9	0.52 (0.16 to 1.74)
Lensen[14]	4	85	5	80	11.1	0.74 (0.19 to 2.86)
Liu[18]	1	29	1	30	3.5	1.04 (0.06 to 17.38)
Mackens[28]	11	35	3	33	10.7	4.58 (1.15 to 18.31)
Maged[19]	11	61	9	41	16.0	0.78 (0.29 to 2.10)
Nastri[87]	1	3	0	4	2.3	5.40 (0.15 to 188.83)
Pye[35]	32	223	43	213	25.5	0.66 (0.40 to 1.09)
Yeung[89]	13	42	7	38	15.0	1.99 (0.70 to 5.67)
Total (95% CI)		547		502	100.0	1.12 (0.64 to 1.96)
Total events	82		76			

Heterogeneity: $\tau^2 = 0.25$; $\chi^2 = 13.43$, df = 8 ($p = 0.10$); $I^2 = 40\%$
Test for overall effect: $Z = 0.14$ ($p = 0.68$)

Appendix 39 Forest plot of all trials reporting miscarriage rates, stratified by the timing of miscarriage

In this figure, 'df' is defined as 'degrees of freedom' and 'M–H' is defined as 'Mantel–Haenszel test'.

Study or subgroup	ES Events	ES Total	Control Events	Control Total	Weight (%)	OR M–H, fixed, 95% CI	OR M–H, fixed, 95% CI
Early miscarriage							
Izquierdo Rodriguez[22]	5	48	8	44	8.2	0.52 (0.16 to 1.74)	
Maged[19]	11	61	9	41	9.7	0.78 (0.29 to 2.10)	
Subtotal (95% CI)		109		85	17.9	0.66 (0.31 to 1.42)	
Total events	16		17				
Heterogeneity: $\chi^2 = 0.26$, df = 1 ($p = 0.61$); $I^2 = 0\%$							
Test for overall effect: $Z = 1.06$ ($p = 0.29$)							
Late miscarriage							
Izquierdo Rodriguez[22]	5	48	8	44	8.2	0.52 (0.16 to 1.74)	
Lensen[14]	4	85	5	80	5.4	0.74 (0.19 to 2.86)	
Mackens[28]	11	35	3	33	2.3	4.58 (1.15 to 18.31)	
Mahran[20]	13	147	14	71	18.9	0.39 (0.17 to 0.89)	
Nastri[87]	1	3	0	4	0.3	5.40 (0.15 to 188.83)	
Pye[35]	32	223	43	213	41.4	0.66 (0.40 to 1.09)	
Yeung[89]	13	42	7	38	5.6	1.99 (0.70 to 5.67)	
Subtotal (95% CI)		583		483	82.1	0.81 (0.58 to 1.14)	
Total events	79		80				
Heterogeneity: $\chi^2 = 14.03$, df = 6 ($p = 0.03$); $I^2 = 57\%$							
Test for overall effect: $Z = 1.22$ ($p = 0.22$)							
Total (95% CI)		692		568	100.0	0.78 (0.58 to 1.07)	
Total events	95		97				
Heterogeneity: $\chi^2 = 14.41$, df = 8 ($p = 0.07$); $I^2 = 44\%$							
Test for overall effect: $Z = 1.55$ ($p = 0.12$)							
Test for subgroup differences: $\chi^2 = 0.22$, df = 1 ($p = 0.64$); $I^2 = 0\%$							

0.01 0.10 1.00 10.00 100.00
Favours ES   Favours control

Appendix 40 Forest plot of all trials reporting multiple pregnancy rate, using a random-effects model

In this figure, 'df' is defined as 'degrees of freedom' and 'M–H' is defined as 'Mantel–Haenszel test'.

Study or subgroup	ES Events	Total	Control Events	Total	Weight (%)	OR M–H, random, 95% CI	OR M–H, random, 95% CI
Izquierdo Rodriguez[22]	4	69	7	71	13.8	0.56 (0.16 to 2.02)	
Lensen[14]	4	94	3	87	9.7	1.24 (0.27 to 5.73)	
Liu[18]	10	29	9	30	18.8	1.23 (0.41 to 3.67)	
Maged[19]	31	61	21	41	35.9	0.98 (0.45 to 2.17)	
Yeung[89]	13	42	8	38	21.7	1.68 (0.61 to 4.65)	
Total (95% CI)		295		267	100.0	1.09 (0.68 to 1.75)	
Total events	62		48				

Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 1.87$, df = 4 ($p = 0.76$); $I^2 = 0\%$
Test for overall effect: $Z = 0.36$ ($p = 0.72$)

0.01　0.10　1.00　10.00　100.00
Favours control　Favours ES
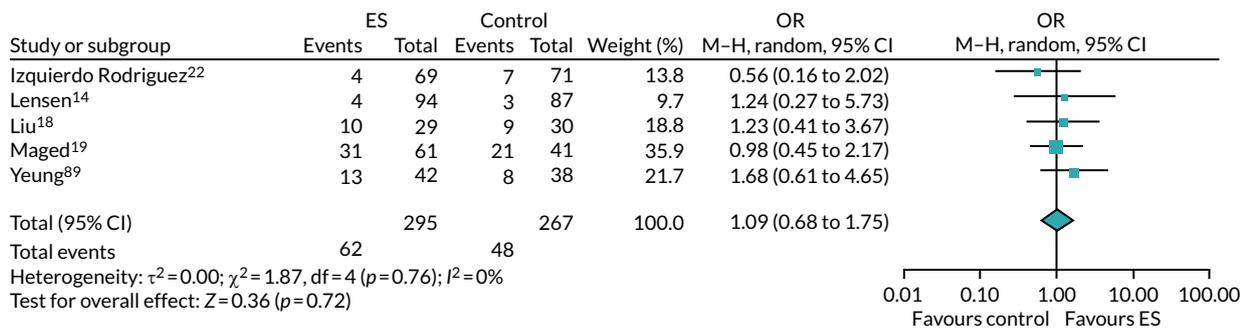
Appendix 41 Forest plot of all trials reporting multiple pregnancy rate, using a fixed-effects model

In this figure, 'df' is defined as 'degrees of freedom' and 'M–H' is defined as 'Mantel–Haenszel test'.

Study or subgroup	ES Events	ES Total	Control Events	Control Total	Weight (%)	OR M–H, fixed, 95% CI
Izquierdo Rodriguez[22]	4	69	7	71	19.4	0.56 (0.16 to 2.02)
Lensen[14]	4	94	3	87	8.9	1.24 (0.27 to 5.73)
Liu[18]	10	29	9	30	17.3	1.23 (0.41 to 3.67)
Maged[19]	31	61	21	41	36.9	0.98 (0.45 to 2.17)
Yeung[89]	13	42	8	38	17.3	1.68 (0.61 to 4.65)
Total (95% CI)		295		267	100.0	1.09 (0.68 to 1.74)
Total events	62		48			

Heterogeneity: $\chi^2 = 1.87$, df = 4 ($p = 0.76$); $I^2 = 0\%$
Test for overall effect: $Z = 0.35$ ($p = 0.72$)
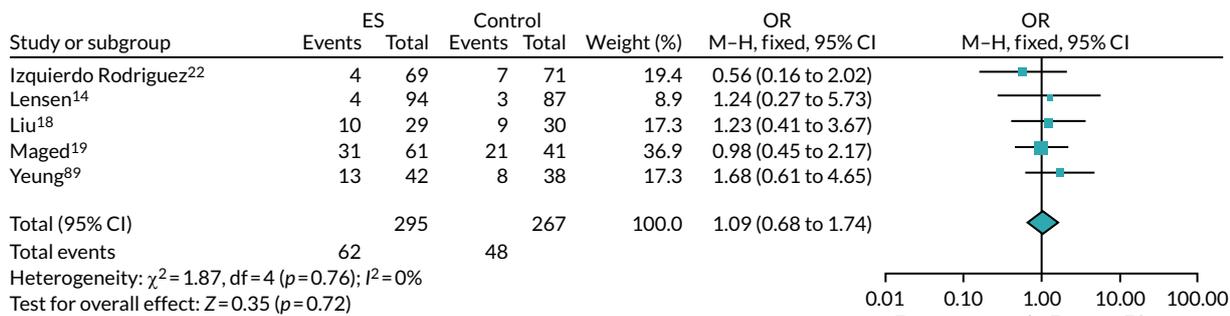
Appendix 42 Forest plot of all trials reporting ectopic pregnancy rate (random effects)

In this figure, 'df' is defined as 'degrees of freedom' and 'M–H' is defined as 'Mantel–Haenszel test'.

| Study or subgroup	ES		Control		Weight (%)	OR M–H, random, 95% CI	OR M–H, random, 95% CI
	Events	Total	Events	Total			
Izquierdo Rodriguez[22]	4	69	8	71	53.5	0.48 (0.14 to 1.69)	
Lensen[14]	0	94	1	87	8.1	0.31 (0.01 to 7.59)	
Liu[18]	1	29	0	20	7.9	2.16 (0.08 to 55.68)	
Mackens[28]	1	35	0	33	8.0	2.91 (0.11 to 74.06)	
Mahran[20]	0	147	1	71	8.1	0.16 (0.01 to 3.96)	
Pye[35]	1	253	2	258	14.4	0.51 (0.05 to 5.64)	
Total (95% CI)		627		540	100.0	0.56 (0.22 to 1.39)	
Total events	7		12				

Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 2.45$, df = 9 ($p = 0.78$); $I^2 = 0\%$
Test for overall effect: $Z = 1.25$ ($p = 0.21$)
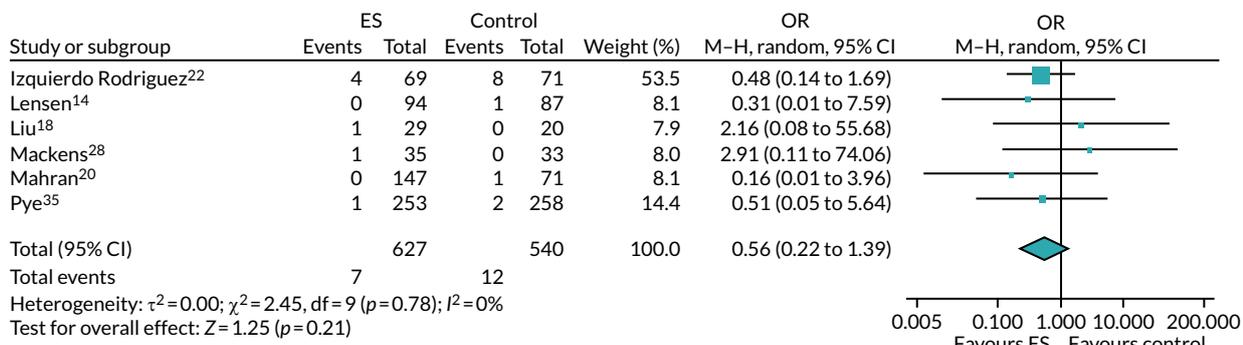
Appendix 43 Forest plot of all trials reporting ectopic pregnancy rate (fixed effects)

In this figure, 'df' is defined as 'degrees of freedom' and 'M–H' is defined as 'Mantel–Haenszel test'.

Study or subgroup	ES Events	ES Total	Control Events	Control Total	Weight (%)	OR M–H, fixed, 95% CI	OR M–H, fixed, 95% CI
Izquierdo Rodriguez[22]	4	69	8	71	53.0	0.48 (0.14 to 1.69)	
Lensen[14]	0	94	1	87	11.1	0.31 (0.01 to 7.59)	
Liu[18]	1	29	0	20	4.0	2.16 (0.08 to 55.68)	
Mackens[28]	1	35	0	33	3.5	2.91 (0.11 to 74.06)	
Mahran[20]	0	147	1	71	14.4	0.16 (0.01 to 3.96)	
Pye[35]	1	253	2	258	14.1	0.51 (0.05 to 5.64)	
Total (95% CI)		627		540	100.0	0.57 (0.24 to 1.35)	
Total events	7		12				

Heterogeneity: $\chi^2 = 2.45$, df = 5 ($p = 0.78$); $I^2 = 0\%$
Test for overall effect: $Z = 1.27$ ($p = 0.20$)

0.005　0.100　1.000　10.000　200.000
Favours ES　Favours control