IVF Add-Ons: The Quantitative and Qualitative Evidence Behind Their Use

In support of submission for the degree of Doctor of Medicine (MD)

Dr Sarah C. Armstrong MBChB MRCOG

University of Sheffield

February 2023

Table of Contents

1. Acknowledgements

I consider myself incredibly lucky to have had Professor Allan Pacey as my MD supervisor. I had the good fortune of meeting Allan in 2015 at a World Health Organisation meeting in Geneva to generate a guideline for the world on fertility. The guideline itself never flourished, but the meeting seeded one of the most important collaborations of my career to date. One of Allan's great talents is clarity of thinking and the ability to make the insurmountable, surmountable. So, with his typical ease, he informed me that being 8 months pregnant, and working full-time clinically in the South West would be no barrier to undertaking an MD in Sheffield: I could do it part-time and remotely, which I duly did! Without Allan's support, encouragement, wisdom, and incredible expertise, I wouldn't be enjoying the varied and fulfilling career I do today. He is a true feminist, lifting me up to reach my potential, whilst also recognising the richness of a rounded life outside of work. Thank you, Allan, for inspiring me to pursue the path less trodden. I really hope to emulate your inspirational leadership.

I would like to thank my secondary supervisor Professor Cindy Farquhar, another extraordinary visionary and role model to me who has shaped my life in so many positive ways. Thank you for giving me your time, encouragement, and support over the years; I have learnt so much from you and will continue to do so. One of your great strengths is ensuring that female leadership is not simply an aberration by paving the way for women.

My sincere thanks go to Dr Liz Williams who provided a listening ear and sage advice when the MD felt overwhelming after the birth of my children. I would like to thank my co-authors and collaborators all of whom I consider friends: particularly Dr Sarah Lensen (University of Melbourne); Dr Elaine Wainwright (University of Aberdeen); Dr Emily Vaughan (University of Bristol); Dr Priya Bhide (Queen Mary University of London); and Professor Adam Balen (University of Leeds). Your enthusiasm, expertise, dedication, and belief in evidence-based medicine has made our work meaningful and impactful to both patients and clinicians.

Last but by no means least I wish to thank my family. My husband Robert, you have given me the liberty and encouragement to pursue my dreams, enabling me to lead a life less ordinary. This has come with sacrifice, and I am more grateful than you can know for your belief in me. My parents Alan and Cathy Armstrong whose undying love, counsel, and confidence in me has made me fearless in the choices I've made. My parents in law Chris and Lynn Axe, who've shared in my successes, offered encouragement, and helped so much with family life. Sophie Smith for your loyalty, kindness, and making it all possible. Finally, my beautiful, kind, and loving children to whom this work is dedicated. Hattie and Tom, you have watched mummy balance the disparate roles of mother, clinician, and researcher over the years and whilst there's no perfect balance, I hope I've shown that it is possible and immensely rewarding with the right love and support.

2. Abstract

Aims and objectives

This thesis has three aims: (i) to generate, analyse and present the highest quality quantitative and qualitative evidence on *in vitro fertilisation* (IVF) add-ons; (ii) to advocate for evidence-based medicine regarding add-ons, and; (iii) to be a source of trusted information to help inform policy, guidelines and recommendations for governmental regulators, professional societies, and those undergoing IVF.

Methods

Papers 2 and 3 are Cochrane systematic reviews exploring the efficacy and safety behind two commonly used add-ons (time-lapse incubation (TLS), and granulocyte-macrophage colony stimulating factor (GM-CSF) containing culture media). Paper 1 is a narrative review of the quantitative evidence surrounding laboratory add-ons. Papers 4 and 5 are the protocol and write-up of a qualitative semi-structured interview study exploring why patients and professionals opt to use add-ons.

Results

Paper 1: This review of laboratory add-ons did not find any high-quality evidence to support their routine use.

Paper 2: TLS is no more or less effective than conventional incubation for clinical outcomes that reflect effectiveness and safety.

Paper 3: GM-CSF containing culture media is no more or less effective than culture media not containing GM-CSF for clinical outcomes that reflect effectiveness and safety.

Papers 4 and 5: Patients often made decisions about add-ons based on hope, minimising considerations of safety, efficacy, or cost, whereas professionals sought the best outcomes for their patients and wanted to avoid them wasting their money. The driving forces behind add-on use differed: for patients, a professional opinion was most influential, whereas for professionals it was patient driven.

Conclusions

Papers 1 to 5 represent the highest quality, transparent quantitative and qualitative evidence behind add-ons. They have already become a source of trusted information to help inform policy, guidelines, and recommendations for governmental regulators in the UK and Australia. Ongoing assessment of add-ons will continue to fall to the scientific community to ensure that patients and those caring for them can offer informed consent regarding their use.

3. Papers arising from this thesis

This thesis is submitted in publication format. It comprises a collection of peer reviewed, published papers in high-impact scientific journals comprising Fertility and Sterility, Cochrane Database of Systematic Reviews, and BMJ Open. Four papers are presented in their published formats for conciseness and owing to the placement of figures at pertinent junctures throughout the text. The final paper is presented in submission format, owing to it currently undergoing peer review with BMJ Open.

The thesis comprises of five papers: (i) a summary of the evidence surrounding IVF add-ons (**Paper 1**); (ii) a Cochrane systematic review exploring the efficacy and safety of time-lapse systems for embryo incubation and assessment in assisted reproduction, considered an IVF add-on (**Paper 2**); (iii) a Cochrane systematic review on granulocyte macrophage colony stimulating factor-containing culture media for embryos, considered an IVF add-on (**Paper 3**); (iv) a protocol for the VALUE study, a qualitative semi-structured interview study exploring why patients and professionals opt to use non-evidence based add-ons (**Paper 4**); and (v) the write-up of the VALUE study as presented for publication (**Paper 5**).

Pagination throughout the thesis is present at the bottom right-hand corner of the page throughout and original pagination from the source journals has been removed. All publications presented here have been conceptualised, designed, developed, researched, and drafted by me, Dr Sarah Armstrong, as the lead author. I was the Principal Investigator of all research presented here. Co-authors have commented on draft research project design, have been involved in data extraction and analysis, and have reviewed and commented on final drafts of papers. The permission of each co-author to submit these papers as part of this MD is given in Appendix 2.

3.1 Papers presented as chapters of this thesis

Paper 1: **Armstrong S**, Atkinson M, MacKenzie J, Pacey A, Farquhar C. Add-ons in the laboratory: hopeful, but not always helpful. *Fertility and Sterility* 2019, 112(6):994-999. DOI: https://doi.org/10.1016/j.fertnstert.2019.10.031

Paper 2: **Armstrong S**, Bhide P, Jordan V, Pacey A, Marjoribanks J, Farquhar C. Time-lapse systems for embryo incubation and assessment in assisted reproduction. *Cochrane Database of Systematic Reviews* 2019, Issue 5. Art. No.: CD011320. DOI: 10.1002/14651858.CD011320.pub4.

Paper 3: **Armstrong S**, MacKenzie J, Woodward B, Pacey A, Farquhar C. GM-CSF (granulocyte macrophage colony stimulating factor) supplementation in culture media for women undergoing assisted reproductive technology (ART). *Cochrane Database of Systematic Reviews* 2019, Issue 12. Art. No.:CD013497. DOI: 10.1002/14651858.CD013497.

Paper 4: **Armstrong SC**, Lensen S, Vaughan E, Wainwright E, Peate M, Balen A, Farquhar C, Pacey A. VALUE study: a protocol for a qualitative semi-structured interview study of IVF add-ons use by patients, clinicians and embryologists in the UK and Australia *BMJ Open* 2021;**11:**e047307. DOI: 10.1136/bmjopen-2020-047307

Paper 5: **Armstrong SC,** Vaughan E, Lensen S, Caughey L, Farquhar C, Pacey A, Balen A, Peate M, Wainwright E. Patient and professional perspectives about using in vitro fertilisation add-ons in the UK and Australia: a qualitative study. Undergoing editorial review with BMJ Open.

3.2 Publications related to research associated with this MD

1. **Armstrong S**, Pacey A, Farquhar C, Lensen S *et al.* O-277 Has time-lapse technology finally proven its clinical benefit? Human Reproduction, Vol 37, Issue Supplement_1 2022, DOI: deac106.005.
2. **Armstrong S**, Vaughan E, Lensen S, Caughey L *et al.* O-0178 The VALUE study: a qualitative semi-structured interview study of add-on use by patients, clinicians, and embryologists in the UK and Australia. Human Reproduction, Vol 37, Issue Supplement_1, 2022, DOI: deac104.092.
3. Lensen SF, **Armstrong S**, Gibreel A, Nastri CO, Raine-Fenning N, Martins WP. Endometrial injury in women undergoing in vitro fertilisation (IVF). Cochrane Database of Systematic Reviews 2021, Issue 6. Art. No.: CD009517. DOI: 10.1002/14651858.CD009517.pub4.
4. **Armstrong SC,** Farquhar CM. Adjuvants in assisted reproduction. In: Reproductive Medicine for the MRCOG. Cambridge University Press; 2021. P82-92. https://doi.org/10.1017/9781108861724
5. **Armstrong S**, Subfertility, Reproductive Endocrinology and Assisted Reproduction. *BJOG* 2020;22:168. DOI:10.1111/tog.12658
6. Lensen, S., Osavlyuk D, **Armstrong S** *et al*. A Randomized Trial of Endometrial Scratching before In Vitro Fertilization. *New England Journal of Medicine* 2019 **380**(4): 325-334. DOI: 10.1056/NEJMoa1808737
7. **Armstrong S**, Bhide P, Jordan V, Pacey A, Farquhar C. Time-lapse Systems for ART-a systematic review. *Human Reproduction* 2017 32: i225-i226. https://doi.org/10.1093/humrep/32.Supplement_1.1
8. **Armstrong S,** Bhide P, Jordan V, Pacey A, Farquhar C. Time lapse systems for ART. *Reproductive Biomedicine Online* 2017. Doi: https://doi.org/10.1016/j.rbmo.2017.12.012

4. Presentations delivered

I have presented the findings of the MD at various junctures during the degree which are outlined below alongside prizes they attracted.

2022	Has time-lapse technology finally proven its clinical benefit? Invited speaker for ESHRE, Milan
2022	The VALUE Study: A semi-structured interview study on IVF add-on use. ESHRE, Milan
2022	The VALUE Study. Oxford Scientific Forum for Obstetrics and Gynaecology (OXSFOG) Awarded first prize
2022	The VALUE Study. The Reproductive Medicine Winter Symposium, Turin (Lawrence Shaw Medal)
2022	Add-ons: do they add up? The Fertility Partnership Monthly Online Clinical Seminar (MOCS)
2022	Podcast for Total Fertility 'Do Fertility Add-ons Add Up?'
2021	Using systematic reviews in everyday clinical practice. ESHRE pre-congress course (Virtual)
2020	IVF add-ons, hopeful but not helpful. 3-minute thesis, University of Sheffield
2019	IVF add-ons and the evidence. Grand Round, Royal United Hospitals, Bath (Frist prize, Bath Clinical Society)
2017	Time-lapse systems for ART: a systematic review. 3-minute thesis, University of Sheffield

5. Prizes awarded from work associated with this thesis

2022	First prize OXSFOG
2022	Lawrence Shaw Medal, Reproductive Medicine Winter Symposium £500
2020	Dame Pamela Shaw Award, University of Sheffield Medical School £500
2019	Health Education England Bursary £1065
2018	First prize, Bath Clinical Society Research and Development Award £1000

6. Introduction

The advent of in vitro fertilisation (IVF) in 1978 was a breakthrough for infertile people, but current live birth rates per cycle initiated are more or less static. In 2019 fewer than 26% of women had a baby with each cycle [1, 2]. The decline in fecundability with advancing female age makes IVF a time-sensitive treatment and this in combination with limited state funding means that many patients often pay for IVF themselves. The pressure to improve IVF outcomes has led to a search for additional or adjunct procedures known as "add-ons". Add-ons range hugely in scope and variety. They can be grouped broadly into five categories: (i) add-ons for eggs, sperm, and embryos; (ii) incubators; (iii) medications, including intravenous infusions; (iv) operative procedures; and (v) alternative therapies (Table 1). The purpose of add-ons has been to improve the chance of taking home a baby from IVF, however add-ons have been widely introduced without evaluation and usually represent an additional cost to patients [3]. IVF clinics who offer them have been described as 'mercenary' or 'exploitative' owing to charging vulnerable patients for non-evidence based extras [4-10]. The lack of evidence and concerns about informed consent has further highlighted the debate about their merit [11, 12].

6.1 Terminology

There is no peer reviewed, published definition of what an add-on is. The UK regulator of IVF, the Human Fertilisation and Embryology Authority (HFEA) describe them as:

> *'Optional additional treatments, also referred to as 'supplementary', 'adjuvants' or 'embryology treatments'. Add-ons often claim to be effective at improving the chances of having a baby (live birth rate) but the evidence to support this for most fertility patients is usually missing or not very reliable. Add-ons are likely to involve an additional cost on top of the cost of a routine cycle of proven fertility treatment. Some treatment add-ons can cost hundreds or thousands of pounds each'*[13].

In 2019, a consensus statement was released by the HFEA regarding the responsible use of treatment add-ons in fertility services. Signatories, amongst others, included the Royal College of Nursing, the HFEA, The European Society of Human Reproduction and Embryology

(ESHRE), and the Royal College of Obstetricians and Gynaecologists [14]. In this statement, add-ons are described as:

> *'optional extras to treatment which claim to improve the chances of having a healthy baby. They cover a range of interventions: genetic tests, drugs, surgery and equipment'.*

The HFEA have created a webpage for patients, in an attempt to help patients navigate the evidence behind add-ons. The website offers information on 11 add-ons, which are described as 'some of the most common', but a full list is not provided [15].

The absence of a definition means that novel add-ons are difficult to quickly identify and thus can lead to a delay in the scientific community undertaking timely assessments of their efficacy and safety. It also makes it difficult for patients to understand whether it is in an optional extra, or an essential part of a standard IVF or ICSI cycle.

6.2 Cost of add-ons

IVF is expensive, with HFEA and NHS resources quoting a cost of at least £5000 per cycle [16, 17]. IVF is funded in specific circumstances by the NHS across all four nations of the UK, however there is enormous disparity in eligibility criteria, leading to inequity in access. In England, funding for IVF is managed by integrated care boards (ICBs), which have recently superseded clinical commissioning groups (CCGs). ICBs set local funding criteria for treatment, alongside what number of cycles of IVF are offered [17, 18]. The autonomy of ICBs and CCGs before them, has created enormous disparity in fertility treatments available across England, with some regions offering the NICE (National Institute for Health and Care Excellence) recommended three cycles of IVF in women aged 39 and under, whereas in others, one or even no cycles are offered [18, 19]. This inequity of care has been coined 'a postcode lottery' by patient campaign groups [20]. As well as a variation in the number of cycles offered by ICBs, there are also disparities in the funding criteria for those seeking treatment. For example, there are variations in the upper age limit of the woman eligible for treatment, with some ICBs limiting treatment to those under the age of 35 [20]. In Scotland, NICE guidance is followed and three NHS funded cycles of IVF are offered to women under the age of 40, however there is a waiting list for treatment [21]. In Northern Ireland, the Health and Social Care Board (HSC) are

responsible for commissioning of fertility services. A motion calling for the HSC to provide three full cycles of treatment has been approved in principle, but in reality, additional finance needs to be made available to make this happen, even in a phased approach [22]. In Wales, two full cycles of NHS funded IVF are available to women who meet the Welsh Health Specialised Services Committee access criteria [23].

The postcode lottery of IVF funding means that over half of people seeking treatment are left to self-fund their IVF cycle, and face additional costly decisions regarding whether to utilise add-on therapies, some of which can cost many hundreds of pounds [24]. Investigations by the HFEA have revealed that 70% of licensed clinics offer at least one add-on, with prices varying enormously from one clinic to the next [25]. There was also a significant variation in the information being offered to patients regarding add-ons, with some being less open than others about the lack of evidence of effectiveness [26].

Assisted reproduction is a rapid paced area of medicine, where demand for treatment is growing, and this is accompanied by accelerated innovation, often driven by industry and pharmaceutical companies [24]. The increasing demand for fertility treatments over recent years has been accompanied by a high up-take of add-ons, with up to 70% of those seeking fertility treatment opting to use at least one in their cycle [1, 24, 27, 28]. Novel add-on therapies arrive on the market and are taken up by clinics with the aim of increasing the chance of pregnancy or live birth, whilst providing revenue to the clinic. However, the evidence behind these add-on therapies is lacking, with most treatments unsupported by good quality evidence [8, 29].

Table 1: Add-on therapies currently available in the UK.
This table is reproduced in its published format within Paper 4, pages 135/6 of this thesis.

Category	Add-on	Description
Add-ons for eggs, sperm, and embryos	Egg activation with calcium ionophore	Aims to stimulate egg activation utilising calcium ionophore. Egg activation is a vital step in fertilisation.
	Intrauterine culture	Aims to allow early embryo development to take place within the woman's uterus, opposed to within an incubator. A device with the early embryo is placed inside the woman.
	EmbryoGlue™	EmbryoGlue™ is a embryo culture medium that contains hyaluronate. Hyaluronate is present naturally in the endometrium at the time of implantation. EmbryoGlue™ is offered as the medium used to transfer the embryo into the uterus.
	Elective freeze-all cycles	All embryos created from a cycle of IVF are frozen, with none transferred 'fresh'. Embryos are then replaced in future frozen embryo transfer cycles in the belief that it will improve livebirth chances.
	Assisted hatching	A technique using acid, lasers or other tools to thin, or make a hole in the zona pellucida of the embryo in order to help the embryo 'hatch'.
	Preimplantation genetic testing for aneuploidy (PGT-A)	PGT-A involves biopsying the embryo to remove cells in order to check for chromosomal abnormalities. The aim is to allow patients to choose euploid embryos for transfer, reducing the risk of miscarriage.
	Sperm DNA test	Sperm DNA testing is a non-invasive procedure performed on a semen sample, usually before treatment as an additional diagnostic test. It aims to assess for DNA damage.
	Embryo culture media containing growth factors (BlastGen® EmbryoGen™)	Culture media for the growth of embryos containing GM-CSF. Aiming to improve the development of embryos to blastocyst stage.
	Intracytoplasmic morphologically selected sperm injection (IMSI)	Use of a high-powered microscope to select sperm for ICSI (intra-cytoplasmic sperm injection).
	SpermSlow	A solution containing hyaluronic acid which aims to slow the sperm to allow the embryologist to select the best sperm for ICSI.
Incubator	Time-lapse imaging (Embryoscope, Primovision, CAREmaps)	A process that enables the assessment of the developing embryo without the need to remove it for assessment outside of the incubator. A time-lapse sequence of images can be created, and computer software applied which can help the embryologist determine the best embryo to replace.
Medications including intravenous infusions	Intravenous immunoglobulin (IVIG)	A blood product containing antibodies given through an intravenous infusion. Aims to reduce miscarriage and failed implantation.

	Tumour necrosis factor alpha blocking agents e.g. infliximab, adalimumab, etanercept	Given as an intravenous or subcutaneous infusion with the aim of reducing miscarriage and failed implantation.
	Intralipid infusion	Intralipid is a fat emulsion of soybean oil and water which is administered intravenously with the aim of reducing miscarriage and failed implantation.
	Quad therapy: aspirin, heparin, progesterone and prednisolone	A combination of medications which aims to improve chances of implantation and reduce miscarriage
	Platelet rich plasma	A blood product infused into the uterus, or infected into the ovaries with the aim of improving implantation, and improving egg quality.
	Testosterone or androgens (DHEA, androderm patch)	Androgenic hormones given orally to improve the quality and maturity of eggs.
Operative procedures	Endometrial scratching	A procedure carried out before IVF where the endometrium is deliberately scratched with the aim of improving implantation and livebirth rates.
	Endometrial receptivity array (ERA)	A genetic test undertaken on a sample of endometrium to help with timing of embryo transfer with the aim of improving implantation and livebirth rates.
Alternative therapies	Chinese herbal medicine	The use of herbal medicine to improve fertility treatment outcomes
	Acupuncture	The practice of inserting small needles into the skin to improve fertility treatment outcomes.

6.3 Professional bodies, regulatory bodies and add-ons

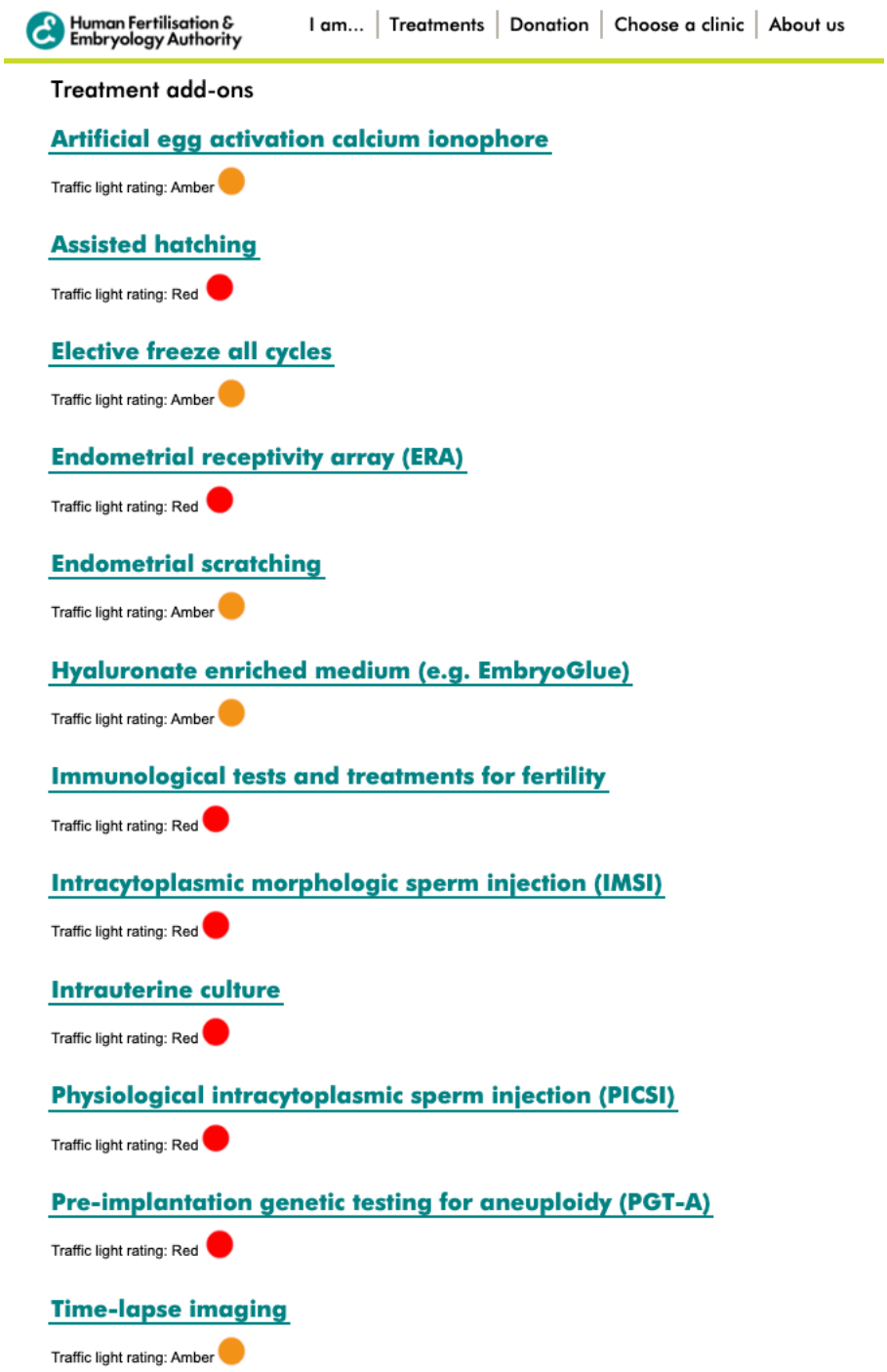
In 2009 and then again in 2015, the British Fertility Society (BFS) undertook scoping reviews on medical add-ons for IVF in response to the surge in uptake of add-ons by patients [30, 31]. As a result, the BFS published a summary of recommendations for good clinical practice for clinicians working in the field of reproductive medicine [30]. In 2017, the rising number of concerns and questions raised by patients regarding add-ons led the HFEA to produce evidence-based information on an initial nine popular add-ons regarding efficacy, safety, and cost. The information took the form of a patient-facing webpage with a traffic light system to denote the quality of evidence on efficacy and safety. Red denoted an add-on with no evidence of benefit or safety, amber for one with a growing body of evidence which is showing promise, but further trials are needed, and green to denote an add-on with more than one high quality study that demonstrated efficacy and safety [26]. The HFEA have modified the webpage twice over the years, adding to the number of add-ons which now sits at 12, and adjusting the traffic-light system to only denote evidence of efficacy (Figure 1) [13]. The decision to drop evidence of safety from the traffic light system was made owing to difficulty in distinguishing efficacy from safety within one colour, and also due to a paucity of evidence surrounding long-term safety of add-ons.

The HFEA's involvement in producing patient resources regarding add-ons has been highly controversial. The HFEA is the UK's regulator for IVF, a statutory body, ensuring that licenced IVF clinics abide by UK law with regards to fertility [32]. IVF add-ons are of course legal, having been licensed and approved for use in the UK by the HFEA itself, which begs the question regarding its conflicts of interest. Europe's largest fertility society, ESHRE, has written a critical review of the HFEA's add-on website efforts, which revealed that ESHRE believe certain add-ons to be effective [33]. ESHRE are in the midst of producing a guideline on add-ons that will include "four diagnostic procedures, 11 laboratory interventions, three 'selective' treatments, and 15 initiatives of clinical management." Further details of the add-ons included are not available at this stage [33].

The BFS website offers patient information under their 'quick guides' on some add-ons, including time-lapse imaging and PGT-A [34]. The information regarding time-lapse has been prepared by The Association of Reproductive and Clinical Scientists, and states that 'most

research concludes that time-lapse offers a benefit to patients' [34]. This is at odds with the information offered by the HFEA traffic light patient information, which regards this technology as 'amber' [13]. This may represent a somewhat confusing contradiction for patients.

Figure 1: HFEA website traffic light ratings of treatment add-ons [15]



Reproduced with permission from HFEA (Appendix 1)

6.4 Evidence behind IVF add-ons and the media

In 2016, a paper was published in the BMJ that compiled a list of interventions offered in addition to standard IVF and sought to examine the evidence behind improvement in live birth rate [35]. The authors identified 27 add-ons, of which only 4 revealed evidence of a benefit in live birth rates. The paper attracted a wave of criticism, most notably in the form of a letter signed by over 60 leading clinicians and scientists in the UK [36]. The criticism focused on the lack of scientific robustness of the paper, owing to the mixing of categories of treatment, not all of which were considered to be add-ons. For example, they were criticised for including investigations of ovarian reserve, and essential treatments such as surgical sperm retrieval for men with male-factor infertility, under the term 'add-on'. The authors of the criticism went on to underline that the BFS and RCOG are fully aware of the lack of evidence of add-ons, using intralipid (an infusion given to patients with supposedly abnormal levels of natural killer cells) as an example [30]. Finally, criticism was levelled at the use of live birth as primary outcome measure, stating that 'using live birth as the sole indicator of an evidence base oversimplifies a hugely complex process and fails to recognize the significant scientific research underlying decisions to bring treatments into clinical practice' [36]. They concluded that the paper was highly inaccurate and misrepresents the fertility sector, calling it 'misleading and deeply unhelpful to patients'.

This paper garnered traction in the press, which was closely followed in the same year by a BBC Panorama programme exploring add-ons in the UK. The show sent reporters with secret cameras into private fertility clinics posing as prospective patients. Add-ons were discussed in a positive light, with professionals filmed making claims made that they would improve livebirth chances [4]. Thus, 2016 represented the beginning of a period of controversy for IVF clinics who offered add-ons. The fertility sector was portrayed as exploiting patients and using add-ons to make revenue [6]. **Paper 1,** a narrative review and opinion piece, was written during this period of intense controversy. It was aimed at gathering the highest quality available evidence on the most popular add-ons for those involved in IVF [3]. **Paper 1** also aimed to make the evidence-based voices of professionals within the sphere of assisted reproduction heard. Following 2016 Panorama exposé, it felt important to reflect the sentiment that add-ons do not serve patients, or clinics, and that evidence-based medicine remains crucially important within assisted reproduction.

6.5 Existing literature regarding time-lapse incubation and assessment of embryos

Here the biological plausibility for time-lapse in IVF is summarised, alongside the existing systematic review evidence to 2016.

Background

Embryo incubation is a vital step in IVF procedures, providing an early embryo with the conditions required to develop in the first few days of its life [37]. Embryo development is a dynamic process, moving through the fertilisation stage to cleavage stage and onto blastocyst stage in some cases [37]. Embryologists apply a tiered grading system based on embryo morphology to assess the quality of the embryo, and thus its potential for implantation and a successful pregnancy [37-42].

The minimum data set required for the accurate description of embryo morphology was established in 2011 by Alpha Scientists in Reproductive Medicine and European Society of Human Reproduction and Embryology (ESHRE) Special Interest Group of Embryology [41]. A consensus on timings of observations of fertilized oocytes and embryos was established to compare results between different laboratories. This essentially involves a daily check of morphology following the fertilisation check that takes place approximately 17 hours following incubation or injection [41].

Traditionally, these morphological checks have been achieved by removing embryos from the controlled environment of the incubator and placing them under a microscope for assessment. This exposes embryos to changes in temperature, pH, mechanical disturbance, changes in humidity and gas composition [43].

Time-lapse imaging of human embryos during incubation has been researched and described in the scientific literature since the 1990s. It was in 1997 that one of the first studies to report on time-lapse assessment of fertilization and early embryo development following ICSI was published [44]. The technology proved a useful tool for studying embryo development for research purposes, but it wasn't until 2009 that time-lapse systems (TLS) for embryo

incubation and assessment came into commercially available use [45]. Since then, there has been widespread uptake of TLS in IVF clinics around the world [46].

Potential advantages of TLS

TLS hold two potential advantages. The first lies in its capability to obtain frequent detailed digital images of developing embryos, which can be compiled to create a time-lapse sequence of embryo development. This negates the need to disturb the embryos for morphological assessment. It also allows for an undisturbed culture environment for embryos and for the assessment by embryologists of the timing and synchronicity of early mitotic divisions and abnormal cleavage patterns that generate morphokinetic parameters [45]. The second is the use of software programmes that utilize complex algorithms based on a combination of morphokinetic parameters and selection and de-selection criteria to aid the embryologist in selecting the optimal embryo for transfer [43, 47, 48]. These algorithms will be referred to as 'embryo selection software' in this chapter.

Potential disadvantages of TLS

TLS involve exposing embryos to light during the acquisition of digital images of embryos, often as frequently as every five to 10 minutes [47]. The total dose of ultraviolet radiation is likely to be very low, however there is potential for harm, which may be borne out in clinical outcomes such as miscarriage and stillbirth [47]. However, the overall light exposure is thought to be lower than with traditional embryo assessment under a light microscope [49]. Several studies have assessed the potential for harm using surrogate outcomes such as fertilization, embryo development, blastocyst formation and implantation potential, and all conclude that TLS have no detrimental effect on embryo development [43, 50-53]. Appropriately powered randomised controlled trials focusing on the clinical outcomes associated with harm (miscarriage and stillbirth) were not found in the published literature at the time of reviewing the existing evidence in 2016.

TLS are more costly than conventional incubators, and currently the cost of TLS is often being passed onto patients in the UK [29]. In 2016 this was estimated to be an extra £850 added to the cost of a conventional IVF cycle to use the technology [29]. There does not seem to have been any health economic assessments of the use of TLS technology.

Available evidence on clinical outcomes

Three observational studies reveal higher pregnancy rates with the use of TLS utilising embryo selection software [43, 54, 55]. However, it is worth noting the methodological flaws in these studies: (i) one study divided oocytes opposed to women between interventions, which makes the study vulnerable to unit of analysis errors. Use of multiple observations per woman (for example when dividing a woman's eggs between interventions) leads to unpredictable bias in the estimate of treatment difference and exaggerates the apparent sample size. This exaggeration leads to spuriously narrow confidence intervals and low *P*-values; (ii) another study was a retrospective cohort study and used a combination of donor and autologous oocytes. Alongside the unit of analysis error, there is also the problem of utilising donor eggs that are likely to be from fertile patients, with an inherently higher chances of success with IVF; and (iii) the third study was a small retrospective cohort with no control.

More recently several RCTs have been published which assess clinically important outcomes such as pregnancy, livebirth, miscarriage, and stillbirth. The first two published RCTs randomised embryos or oocytes and compared TLS with a conventional incubator, however assessment of embryo quality was undertaken by routine morphology in both arms [50, 56]. Both studies reported no differences in clinical pregnancy rates between the interventions, based on the absence of a statistically significant difference.

Eight RCTs that randomised couples were found, the largest of which was undertaken in Spain [57]. A total of 856 couples were randomised to TLS utilising embryo selection software versus conventional incubation and assessment. It revealed higher ongoing pregnancy rates (RR 1.23, 95% CI 1.06-1.43) with the use of TLS. Two further smaller RCTs also undertook the same comparison both of which are interim results of ongoing studies [58, 59]. One study reported no difference in ongoing pregnancy rates between the two interventions (TLS: 93/144 vs control: 101/140, p=>0.05) [59]. The other had similar findings (TLS: 14/24 vs control 11/25) [58]. However,

it is worth noting that all studies did not follow intention to treat protocols[1] in reporting of their results.

Three RCTs were found that compared TLS without use of the embryo selection software versus conventional incubation and assessment. One large study of 364 couples, undertaken in Sweden, reported no difference in pregnancy rates between the two arms of the study (TLS 72/240 vs control 39/124, *p*=0.87) [60]. Likewise, two further small studies, one undertaken in Turkey, and the other in the USA, reached the same conclusion regarding pregnancy rates (TLS 24/33 vs control 23/31, *p*=0.89; TLS 3/24 vs control 3/25, *p*=1.0) [61, 62].

Finally, two RCTs undertaken in the USA compared TLS utilising embryo selection software versus TLS without embryo selection software. Both studies reported no improvement in clinical pregnancy rates with the addition of embryo selection software [63, 64].

Systematic reviews

A systematic review published by Cochrane in 2015 included three RCTs and concluded no evidence of a difference to choose between TLS and standard incubation and assessment of embryos [65]. The quality of the evidence was low to moderate and only one trial reported pregnancy rates.

Four further systematic reviews have been found that examine clinical outcomes [66-69]. One includes only two RCTs and is therefore out of date [68]. Another included 13 eligible studies, however none of them were RCTs and the majority were retrospective cohort studies [69]. This review concluded that there is currently limited high-quality evidence to support the routine clinical use of TLS for selection of human pre-implantation embryos [69]. Another review included six RCTs and concluded that there was currently 'insufficient evidence to support that time-lapse imaging is superior to conventional methods for embryo incubation and selection' [67]. Most recently, an author group undertook a systematic review of TLS utilising TLS embryo selection software [66]. They concluded that TLS using embryo selection software

[1] An intention to treat protocol means that all data from patients randomised are analysed according to what group they were initially randomised to, regardless of whether they dropped out, deviated from the protocol, or were non-compliant after randomisation. The advantage of an intention to treat analysis is that it avoids bias that can arise from non-random attrition of participants.

was associated with a significantly higher ongoing pregnancy rate, a significantly lower early pregnancy loss and a significantly higher live birth rate in comparison to control.

At the start of this MD, the need for a comprehensive, high quality systematic review of RCTs on TLS was evident, given the disparity in results from existing systematic reviews, and the fact that no existing systematic review had included all available RCTs discovered in this search of the literature. **Paper 2** presented in this thesis therefore provides this comprehensive review.

6.6 Existing literature regarding Granulocyte-macrophage colony stimulating factor (GM-CSF) containing culture media

Background

Culture media are the solutions that human pre-implantation embryos are immersed in for several days during their development in IVF or ICSI treatment. In the early days of IVF, culture media were formulated 'in house', with Louise Brown, the first child born of IVF, having been cultured in Earle's simple salt solution with pyruvate supplemented with the patient's serum [70]. At that time there was full transparency regarding the composition of these media, but in the 1980s, commercially available IVF culture media entered the market, and with it came competition and secrecy regarding its balance and composition [71]. There are now many different commercially available media, with variable formations, ranging from simple salt solutions to complex media containing amino acids, vitamins and growth factors, such as GM-CSF [72]. A body of evidence exists that supports the importance of the culture conditions, specifically culture media, on the impact on pre and post implantation development and possibly the health of offspring, in particular, birth weight [71, 73-76]. Therefore, it is possible that the differences in composition of culture media may affect the success rates of IVF.

GM-CSF is a cytokine, also known as a growth factor, produced by epithelial cells under the influence of estrogens, in the human uterus and Fallopian tubes [77, 78]. There are GM-CSF receptors on human embryos and the ovary [79, 80]. Both human and animal studies have demonstrated its survival-promoting effects on embryos, thought to be due to it exerting a positive control over various genetic paths, such as cell proliferation, progression to blastocyst, zona pellucida hatching, embryo implantation into the endometrium, and

suppression of apoptosis [77, 80-82]. GM-CSF has been found at reduced concentrations in patients with recurrent miscarriage, and experiments involving GM-CSF knockout mice revealed they underwent increased miscarriages, intrauterine growth restriction, impaired placentation, fetal malformations, and the offspring suffered impaired growth and higher post-natal mortality [83, 84].

Potential advantages of GM-CSF

It has been proposed that the addition of GM-CSF to culture media may improve IVF and ICSI success rates by accelerating the development of embryos whilst *in vitro* and *in vivo*; improving their progression to blastocyst, increasing their cell number, promoting earlier hatching, and increasing implantation rates [77, 85]. There is a suggestion that GM-CSF containing culture media may be of particular benefit in women or couples undergoing IVF who have suffered recurrent failed cycles, or miscarriages [71].

The safety of GM-CSF containing culture media was addressed in an RCT in 2010 that randomised the donated oocytes of 73 women undergoing IVF or ICSI to GM-CSF containing media and control media [86]. All oocytes were fertilized by the same chromosomally normal sperm donor and resultant embryos were examined for chromosomal abnormalities on day 3 of development through fluorescence *in-situ* hybridization. There was no difference in aneuploidy or chromosomal constitution between the two groups [86]. None of the donated embryos entering the trial were transferred to the patients owing to the trial being confined to research on embryos, and not a clinical trial.

Potential disadvantages of GM-CSF

One of the key principles that renders GM-CSF potentially beneficial is its ability to suppress apoptosis. However, apoptosis is a normal biological mechanism during embryo development, the purpose of which is to deal with abnormal cells and act as a repair mechanism [71]. Removal or suppression of apoptosis can lead to rapid development of embryos, which does not always represent improved quality. A mouse embryo study raised safety issues for culture media that promote fast growth secondary to loss of genomic imprinting [87]. In addition, there is a phenomenon called 'large offspring syndrome' which has been associated with the addition of growth factors to culture media in an animal study [88].

GM-CSF containing culture media are available from one supplier in the UK under the trade names EmbryoGen® and BlastGen™ [89]. The cost of these culture media is higher than standard culture media which do not contain growth factors and are therefore often treated as an additional cost, or an 'add-on' to an IVF or ICSI cycle. There is evidence that this culture media is rarely funded by the NHS and fertility care providers often charge patients between £420 to £440 to use GM-CSF containing culture media [90, 91]. Finally, the commercial nature of embryo culture media means that the exact concentrations and proportions of various components, including GM-CSF, are not known. This leads to difficulty in conducting trials, as the components of interventions are not clear.

Available evidence on clinical outcomes

A prospective observational study conducted in 1999 on human embryos cultured in GM-CSF containing culture media, from 99 couples undergoing IVF, revealed improved embryological markers for development: they found more reached the blastocyst stage, there was improved hatching initiation and a rise in the number of cells in the blastocyst [77]. However, it wasn't until 2001 that a study examining pregnancy rates was published [92]. This study examined the effects of GM-CSF in culture media versus control on 154 women in a prospective observational study and recorded that the study group had a significantly improved clinical pregnancy rate compared to the control group (46.1% versus 30.8%, p <0.05) in both IVF and ICSI cycles, with the difference being more pronounced when only IVF was used (66.7% versus 37.3%, p <0.05) [92].

Since these initial studies, a number of RCTs examining clinical outcomes have been published, the first of which was undertaken in the USA in 2003 [93]. This group investigated the effect of GM-CSF containing culture media versus control on 72 couples undergoing IVF. The study revealed increased blastocyst cell numbers and more expanded blastocysts in the intervention arm, but no difference in pregnancy rates. A large multi-centre RCT, sponsored by Cooper Surgical, Trumbull, USA, the company that manufactures GM-CSF containing culture media, recruited more than 1300 women from 14 fertility clinics [94]. Women undergoing IVF or ICSI were randomised to have their embryos cultured in GM-CSF containing culture media or control culture media. Authors revealed an improved live birth rate for the

intervention arm (OR 1.35, CI 1.03-1.78, *p*=0.03), but subgroup analysis reveals that this effect was lost if high human serum albumin culture media was used [94]. A pilot study of 69 patients with previously unsuccessful IVF cycles revealed improved pregnancy rates in the GM-CSF arm versus control, but these differences did not reach statistical significance [95].

Systematic reviews

To date, there appears to be only one published systematic review focusing on GM-CSF containing culture media. This was published in 2012 [82]. It included the studies described above, except one as it was not yet published at the time of writing. The primary outcome was livebirth, however only one included study reported on this. The quality of the studies is reported as 'median' which is difficult to interpret. The review was narrative and meta-analysis of the RCTs was not undertaken. The authors of the systematic review conclude that the beneficial nature of GM-CSF in terms of clinical pregnancy and live birth rates has yet to be resolved and that further properly conducted RCTs are required [82].

There is a narrative review on the quality of various culture media, including a section on GM-CSF containing media [71]. It similarly concludes that well powered RCTs assessing live birth and long-term follow up of the offspring are needed to determine the benefit and safety of GM-CSF supplemented media for the general IVF population [71].

Therefore, to address the shortcomings described above, **Paper 3** presented here is the Cochrane review on GM-CSF containing culture media which brings together all available RCTs in a high-quality systematic review.

6.7 Existing qualitative literature regarding the decision-making process surrounding the use of IVF add-ons

In 2018 the HFEA commissioned YouGov to undertake the largest survey of fertility patients in the UK to date [27]. The survey was conducted online and was completed by 1017 patients or partners. It was broad in scope and touched on the use of add-ons, revealing that time-lapse imaging was amongst the most popular, used by 22% of those surveyed. Over three quarters of those who had used a treatment add-on were satisfied that the costs were open and transparent, however some paid slightly more than expected. The survey also revealed

that the use of IVF add-ons had increased between 2013 and 2018. When looking at time-lapse imaging, it was used by 14% of those who had undergone IVF two to five years previously, and 22% of those who had undergone IVF in the past two years. Overall, for those who had treatment in the past two years, three quarters had used at least one type of treatment add-on (74%).

Understanding the reasons behind why patients opt to use add-ons is an area that has not been widely studied or published. It is acknowledged that infertile women/couples embarking on treatment are a vulnerable group, and with desperation, comes the desire to optimise treatment to improve the chances of having their dream child [96, 97]. Some clinicians from the reproductive medicine community, as well as professional societies such as ESHRE believe that it is unethical to deny a patient's right to autonomy in opting to use add-ons [33, 96]. This is argued on the basis that withholding a potentially helpful intervention based solely on lack of proof to its effectiveness at a given moment in time, is unfair because conclusive evidence is likely to take many years to come to light, during which time, the window for treatment may have passed [96]. In addition, it is argued that add-ons have emotional and psychological benefits to patients during a time of intense vulnerability and desperation [98]. Perhaps using add-ons enables individuals to gain emotional closure and inner assurance that they tried all possible treatment options in their efforts to have a baby [96].

However, the principle of autonomy rests upon the patient having informed consent, that is, understanding the potential benefits and risks of any given add-on [99]. This is a challenge for patients who are faced with strong marketing information from Industry, and potentially from fertility clinics as well. It is thought too that fertility treatment providers may be inadvertently contributing to the confusion by using anecdotal examples of patients they have treated locally.

There have been no published studies undertaking semi-structured interviews to establish the level of counselling regarding the efficacy, cost-effectiveness, and pros and cons of add-ons. Therefore, **papers 4 and 5** of this MD thesis outline the protocol and the write-up of the large qualitative semi-structured interview study 'VALUE'. VALUE shows how patients and professionals weigh up the pros and cons of add-ons in the emotive sphere of IVF.

6.8 Overarching aims of the MD

This body of work has three main overarching aims: (i) to generate or gather, analyse and present the highest quality quantitative and qualitative evidence on IVF add-ons; (ii) to advocate for evidence-based medicine regarding add-ons in what is otherwise a very commercial sphere of medicine, and; (iii) to be a source of trusted information to help inform policy, guidelines and recommendations for governmental regulators (HFEA in the UK), Royal Colleges, professional societies, patient charities, and most importantly, provide trusted information to those undergoing IVF.

Quantitative and qualitative evidence

The primary aim for this MD is to provide a summary of the highest quality, reliable quantitative and qualitative evidence for add-ons in one place. Papers 2 and 3 presented here represent many years of diligent and determined hard work to gather the quantitative evidence on two commonly used add-ons (time-lapse incubation and assessment, and GM-CSF containing culture media) into Cochrane systematic reviews. Cochrane reviews are considered the most robust, reliable, transparent, and trusted form of systematic review for clinical topics owing to their methodology. Key reasons behind their credibility include the prospective publication of a peer-reviewed protocol, relying on at least two authors to select studies, assess risk of bias, and interpret the reliability of results. The impact of these papers is evident in the number of citations they have received (243 to date), as well as being used by organisations such as by the HFEA for their add-ons website [13, 100]. Paper 1 provides a narrative overview of the highest quality evidence on the most commonly used add-ons and served as a catalyst for many other research groups to explore the ethics of providing non-evidence based treatment in IVF, as evidenced through citations [100]. Its publication represented a sea change in the way add-ons were viewed and interrogated, which carries on today. Papers 4 and 5 set out the protocol and results of a large novel qualitative study 'VALUE' exploring why patients and professionals opt to use add-ons. Great care was taken to ensure the robust methodology and conduct of VALUE so that its results might be reliable and credible, and therefore helpful to the scientific community and patients.

Advocate for evidence-based medicine

The sphere of IVF remains a commercial area of medicine, with over 57% of licenced clinics in the UK being privately owned [101]. Commercial entities develop IVF add-ons with the aim of improving the success rates of IVF, but they also represent a source of revenue [102]. Research has shown that add-ons are often advertised on IVF clinic websites, with the majority claiming clinical benefit from the add-on [103, 104] Therefore, I wanted this thesis to represent a voice for evidence-based medicine; to shine a light on the best available evidence on add-ons and engage in debates and discussions at conferences, in the literature, and on podcasts (see section 4). I want this work to be a voice for healthcare professions and those undergoing IVF, who desire evidence-based care and reliable information upon which to make informed choices. The information from papers 1, 2 and 3 in this thesis will help by shedding light on evidence regarding potential benefits and risks of add-ons. Papers 4 and 5 reveal the disparate factors that are at play when patients and professionals make decisions regarding add-ons.

Source of information for policy and guidance

The third and final aim of this thesis is that the evidence generated from this work is used to help guide decision making to benefit patients, their families, and professionals caring for them. It is satisfying to know that the information from the Cochrane reviews presented here (papers 2 and 3) is already being used by regulatory bodies in the UK and Australia. In the UK, the HFEA's Scientific and Clinical Advances Advisory Committee (SCAAC) have used the results in assigning traffic light ratings to GM-CSF containing culture media and time-lapse systems for embryo incubation [13]. In Australia, The Victorian Assisted Reproductive Treatment Authority (VARTA) has used hyperlinks to papers 2 and 3 on its patient facing website regarding IVF add-ons [100]. It is also being used by the European Society of Human Reproduction and Embryology (ESHRE) to inform an inaugural guideline entitled 'Add-ons in Reproductive Medicine' which is currently in draft format [101].

I hope that the qualitative results from papers 4 and 5 will help regulatory bodies, medical societies, and patient organisations to recognise the far-reaching consequences of infertility and all the factors and emotions that are at play when making choices regarding IVF add-ons. I would like such organisations to take stock of the results of VALUE to consider how patients' best interests might best be served regarding add-on use.

Paper 1: Add-ons in the laboratory: hopeful, but not always helpful

The idea for paper 1 came about after spending time working on the two Cochrane reviews included in this thesis (papers 2 and 3). It became apparent that there was a gap in the literature for a summary of the highest quality evidence on a broad variety of commonly used add-ons. The summary written by Heneghan *et al* three years previously had shone a light on the lack of evidence to support add-ons, being published in the high-impact BMJ, but it received a lot of criticism [35]. The criticism was levelled at its inclusion of techniques not considered to be add-ons, and due to its authors having no clinical or scientific link with IVF (see section 6.4) [36].

Hence, I wanted this review to be authored by clinicians and embryologists with credibility in the field and published in a journal which would reach the desired audience and a wide readership. I gathered a group of motivated and experienced co-authors and drafted paper 1. It was accepted by Fertility and Sterility, a major journal in the sphere of IVF, with an impact factor of 8.1[102]. The paper has been cited 13 times to date.

Paper 1 aims to set out the highest level of evidence for six commonly used add-ons using systematic reviews of randomised controlled trials where available, followed by randomised controlled trials where systematic reviews are not available. It questions the ethics of having a menu of non-evidence based add-ons that patients are presented with, and calls for closer collaboration between clinicians, embryologists, and patients to determine what our research priorities should be.

Add-ons in the laboratory: hopeful, but not always helpful

Sarah Armstrong, M.B.Ch.B.,[a] Monique Atkinson, B.Sc., M.B.B.S.,[b] Jeanette MacKenzie, B.Sc., M.C.E,[c] Allan Pacey, Ph.D.,[a] and Cynthia Farquhar, M.B.Ch.B., M.D., M.P.H.[c,d]

[a] Department of Oncology and Metabolism, The University of Sheffield, Sheffield, United Kingdom; [b] Westmead Fertility Centre, Sydney, New South Wales, Australia; [c] Fertility Plus, Auckland District Health Board, Auckland, New Zealand; and [d] Department of Obstetrics and Gynaecology, University of Auckland, Auckland, New Zealand

All the steps in an in vitro fertilization cycle are important but none more so than those that occur in the laboratory. To improve the chance of success, adjuncts, commonly referred to as 'add-ons', are offered. Yet as with other new interventions, add-ons in the laboratory require justification by well-designed studies prior to being offered as routine practice. Add-ons aim to improve the chance of a take-home baby, but, their safety and efficacy is less than clear. In addition, the financial burden from the use of add-ons is often borne by the couple. This review of the most commonly used laboratory add-ons did not find any high-quality evidence to support their use in routine practice. (Fertil Steril® 2019;112:994–9. ©2019 by American Society for Reproductive Medicine.)

Key Words: Laboratory add-ons, in vitro fertilization

The interventions that occur in the fertility laboratory are of critical importance to the outcomes of an in vitro fertilization (IVF) cycle. It is not surprising that many scientists and clinicians seek improvements in outcomes by proposing adjuncts (or add-ons) to the traditional laboratory procedures involved in IVF. In the early days of IVF, few innovations were comprehensively evaluated. Some readers may remember gamete intrafallopian transfer, zygote intrafallopian tube transfer, and preimplantation genetic screening on a subset of chromosomes using fluorescence in situ hybridization, all now relegated to history.

The IVF laboratory now faces many rapid changes and innovations with some advances being driven by industry. Most laboratory add-ons fall outside the pharmaceutical regulations and are therefore often not required to have rigorous evaluations before being marketed. As a result, many are unsupported by well-designed research (1, 2). Add-ons may rapidly become established as part of normal working practice well before high quality studies such as randomized controlled trials (RCTs) have been completed. Finally, the cost of add-ons is often borne by patients in the belief that if it is offered, then it must have been properly evaluated. Clearly, it is important that add-ons are properly evaluated so that benefits and harms can be reported.

The UK's Human Fertilisation and Embryology Authority (HFEA) defines IVF add-ons for patients as, "… optional extras you may be offered on top of your normal fertility treatment, often at an additional cost. They're sometimes emerging techniques that may have shown some promising results in initial studies, or they may have been around for a number of years, but haven't necessarily been proven to improve pregnancy or birth rates…" (3).

In this article we discuss the some of the most commonly offered laboratory add-ons, outlining the best available clinical evidence on their use (Fig 1). We have specifically not discussed preimplantation genetic testing for aneuploidy (PGT-A) as this has been debated extensively already.

TIME-LAPSE IMAGING

Embryos develop in an incubator, moving through the fertilization stage to cleavage stage and on to blastocyst stage in some cases. Embryologists check the developing embryos in order to select those most likely to implant and develop into a baby (4–9). Traditionally, embryos are removed from the controlled environment of the incubator and briefly placed under a microscope to be examined by an embryologist. In contrast, time-lapse systems (TLS) allow the embryologist to monitor the developing embryo without removing it from the incubator, and to select the best embryo for transfer based on morphology, and

Received September 18, 2019; accepted October 25, 2019.
S.A. has nothing to disclose. M.A. has nothing to disclose. J.M. has nothing to disclose. A.P. has nothing to disclose. C.F. has nothing to disclose.
Correspondence: Sarah Armstrong M.B.Ch.B., M.R.C.O.G., Department of Oncology and Metabolism, Academic Unit of Reproductive and Developmental Medicine, Level 4, The Jessop Wing, Sheffield S10 2SF, United Kingdom (E-mail: scarmstrong1@sheffield.ac.uk).

Fertility and Sterility® Vol. 112, No. 6, December 2019 0015-0282/$36.00

https://doi.org/10.1016/j.fertnstert.2019.10.031

FIGURE 1

Category	Add-on	HFEA traffic light scoring
Gamete, endometrial and embryological	Time-lapse imaging of embryos*	🟡
	Assisted hatching*	🔴
	EmbryoGlue*	🟡
	Sperm DNA testing*	Not considered by HFEA
	Egg activation with calcium ionophore*	🟡
	Physiological intracytoplasmic sperm injection (PICSI)*	🔴
	Intracytoplasmic morphologic sperm injection (IMSI)*	🔴
	Preimplantation genetic screening (PGS) (on subset of chromosomes) §	🔴
	Endometrial receptivity array†	🟡
Surgical procedures	Endometrial scratching†	🟡
Drug therapies	Reproductive immunology‡	🔴

Key
- 🟢 Evidence of clinical effectiveness and safety
- 🟡 Conflicting clinical effectiveness
- 🔴 Evidence of clinical ineffectiveness

Human Fertilisation and Embryology Authority (HFEA) traffic light ratings of commonly used in vitro fertilization add-ons. *Add-ons are covered in this article. †Add-ons covered Lensen et al. in this issue's Views and Reviews. ‡Add-ons covered by Kamath et al. in this issue's Views and Reviews. §Not covered in this issue's Views and Reviews.

Armstrong. Add-ons in the laboratory. Fertil Steril 2019.

on the timing and synchronicity of early mitotic divisions and abnormal cleavage patterns that generate morphokinetic parameters (10). There has been widespread uptake of TLS in IVF clinics worldwide (11).

Potential advantages of TLS include: the availability of detailed digital images of developing embryos, which can be compiled to create a time-lapse sequence of their development; achievement of an undisturbed culture environment for embryos, which avoids exposing embryos to mechanical disturbance or changes in temperature, pH, humidity and gas composition (12); and the availability of embryo selection software, with complex algorithms based on a combination of morphokinetic parameters and selection and de-selection criteria which help the embryologist to select the optimal embryo for transfer (12–14). Potential disadvantages include the increased osmolarity in drops as some TLS cannot be humified, and the cost, which is approximately $1000 per cycle although not all clinics charge patients (1, 15).

A 2019 Cochrane systematic review (16) included nine RCTs (2955 women). Compared with conventional incubation it was unclear if there was a difference in live birth rate with the use of the time-lapse images with an algorithm (odds ratio [OR] 1.12, 95% confidence interval [CI] 0.92–1.36; 3 RCTs). Overall the review concluded that there was insufficient good quality evidence of differences in live birth or ongoing pregnancy, miscarriage, stillbirth or clinical pregnancy to choose between TLS, with or without embryo selection software, and conventional incubation. The evidence was low or very low quality overall. The HFEA consider that the evidence is conflicting for TLS and that there is certainly not enough evidence to show that time-lapse imaging improves birth rates (3).

ASSISTED HATCHING

The zona pellucida is an acellular glycoprotein coat involved in different processes during fertilization and embryo development. Once the blastocyst reaches the uterus, the embryo needs to exit the zona pellucida so that it can interact with the endometrium and implant. It has been suggested that zona pellucida hardening may occur as a result of in vitro culture (17). The escape of the embryo may be inhibited by a thickened or hardened zona pellucida (18) and it is suggested that failure of the zona pellucida to rupture following blastocyst expansion may be a contributing factor in failure of embryo implantation.

Assisted hatching can be performed with acid, a laser or mechanically. Acid tyrodes digests the zona pellucida leaving a breach. Partial zona dissection uses a micropipette to mechanically slice through the zona pellucida. Laser-assisted hatching uses photoablation to make a very precise and accurately controlled opening in the zona pellucida.

Possible advantages of assisted hatching are that it may assist the embryo to hatch by allowing the embryo to come away from the zona pellucida more freely. However, possible disadvantages are that it may cause damage to the embryo and may increase the risk of a multiple pregnancy (19). The reason for the increase in multiple pregnancies is difficult to explain as most of the studies transfer more than one embryo but it is possible that monozygotic twinning occurs because of embryo splitting occurring during artificial zona hatching (20).

A Cochrane review of 31 RCTs in 2012 demonstrated that although assisted hatching does appear to offer a significantly increased chance of achieving a clinical pregnancy (OR 1.13, 95% CI 1.01–1.27; moderate quality evidence) there was no increase in the live birth rate (9 RCTs; OR 1.03, 95% CI 0.85–1.26; moderate quality evidence) (19). When restricting analysis of clinical pregnancy rate to those trials that went on to report live birth, which overall were better quality trials with a lower risk of bias, the clinical pregnancy result showed insufficient evidence of a difference between the assisted hatching and the control group. A different systematic review in 2016 had similar results with no difference in live birth rate (OR 1.09, 95% CI 0.92–1.30) but a small increase in clinical pregnancies OR 1.16, 95% CI 1.00–1.36) and multiple pregnancy rates (OR 1.50, 95% CI 1.11–2.01) with assisted hatching (20).

Overall, there was some evidence of increased multiple pregnancy in the assisted hatching group in both systematic

reviews but as more than one embryo was usually transferred this is difficult to interpret (19, 20). Only the Cochrane review reported on monozygotic twinning but only 6 of the 31 trials reported this outcome and the data is inconclusive (19). The prevalence of multiple pregnancy with monozygotic splitting after elective single embryo transfer was 1.36% and was associated with embryo manipulations including assisted hatching (20). Another cohort study suggested that blastocyst transfer is more likely to lead to monozygotic twinning (21).

The quality of the evidence is overall poor to moderate secondary to selective reporting in several studies, and significant statistical heterogeneity between trials. The National Institute for Health and Care Excellence (NICE, UK) and the HFEA recommends against the use of assisted hatching, due to lack of evidence of benefit and safety (3, 22). In 2014 the American Society for Reproductive Medicine (ASRM) published guidelines on assisted hatching, they stated there is good evidence that assisted hatching slightly improves clinical pregnancy rates in poor prognosis patients but insufficient evidence to conclude that assisted hatching improves live birth rates. The ASRM recommends that assisted hatching should not be recommended routinely for all patients undergoing IVF (23).

HYALURONIC ACID

The rate of human implantation is low (10%-30%), and failed implantation is a common cause of IVF failure even in younger women with euploid embryos (24). Any improvement in the implantation rate may maximize the chance of achieving a pregnancy and livebirth. Improved implantation rates may also lead to a reduction in multiple embryo transfer in IVF and multiple pregnancy (25).

Hyaluronic acid (HA) is used as a supplement to conventional embryo transfer medium, with the aim of improving implantation rates via the following proposed mechanisms (26): indirectly promote angiogenesis and improve cell-to-cell and cell-to matrix adhesion, thus assisting in embryo apposition and attachment to the endometrium; enhance embryo transfer and prevent expulsion of embryos from the uterine cavity after transfer because of HA's high viscosity; and act as a receptor mediator, as the primary HA receptor is CD44, which is also expressed in the pre-implantation embryo and in the endometrium.

A 2015 Cochrane review included 16 RCTs (n=3,687) using transfer medium supplemented with HA versus transfer medium not supplemented with HA and reported moderate quality evidence of improved live birth rates with the intervention (OR 1.41, 95% CI 1.17–1.69; 6 RCTs, n = 1,950, moderate-quality evidence). However multiple pregnancy rates were also increased, possibly due to combination of adherence compound as most studies transferred multiple embryos (25). A subsequent RCT compared HA versus conventional transfer medium in 581 cycles and found no evidence of an improvement in live birth rates (26).

A raised multiple pregnancy rate is the expected natural consequence of increased implantation and pregnancy rates. This suggests that clinics using HA supplemented embryo transfer medium should adopt an elective single embryo transfer policy, closely monitor their multiple pregnancy rate and ensure that patients are aware of the increased chance of multiple pregnancy if multiple embryos are transferred (2). The HFEA considers the evidence to be conflicting and suggests that further high-quality studies are needed before patients, clinicians and embryologists can be confident of the benefits of hyaluronic acid.

MEASUREMENT OF SPERM DNA FRAGMENTATION

The quality of a sperm sample is not completely described by reporting the number and motility of spermatozoa (27, 28). Tests of sperm DNA fragmentation attempt to provide further information about sperm quality with the aim of improving outcomes such as live birth. A number of sperm DNA fragmentation tests have been developed and these include the: sperm chromatin structure assay; sperm chromatin dispersiontest; terminal deoxynucleotidyl transferase mediated deoxyuridine triphosphate nick end labelling; and single cell gel electrophoresis assay (29-36).

Several underlying causes of sperm DNA fragmentation have been proposed. These include defective apoptosis, excessive reactive oxygen species production; and decreased seminal antioxidants. Exogenous factors have also been implicated including toxic effects of drugs, cigarette smoking, pollution, high testicular temperature associated with systemic fever or varicoceles, and advanced age (29). A potential benefit of identifying high sperm DNA fragmentation could be that it motivates the couple to avoid such exposures.

Ultimately, we must consider if measurement of sperm DNA fragmentation will bring benefit to a couple by influencing clinical management. This must be weighed against the potential disadvantages of the additional cost incurred by the patient for running the sperm DNA fragmentation test, and then the additional cost of any subsequent management strategies proposed. Multiple systematic reviews and meta-analyses have been performed regarding the clinical utility of sperm DNA fragmentation tests (29-36). Although there is some low quality evidence that antioxidant supplementation might improve live birth rates in subfertile men (37), it should be noted as well that these studies have been performed in couples attending fertility clinics, rather than only including men with high DNA fragmentation levels. The question of whether DNA fragmentation tests should guide the decision for treatment is also unclear as although the chance of spontaneous conception is low with a DNA fragmentation index >20%, and approaches zero for DNA fragmentation index >30% to 40% (38, 39), there are no trials comparing the likelihood of spontaneous pregnancy to pregnancy following medically assisted reproduction with different DNA fragmentation thresholds (29).

The ASRM (40) and the British Fertility Society (41) have both concluded that current methods for assessing sperm DNA integrity do not reliably predict treatment outcomes and hence cannot be recommended routinely for clinical use. The limitation of sperm DNA fragmentation

measurements in predicting pregnancy may be secondary to the fact that multiple other factors also influence pregnancy. In particular the quality and age of the oocyte is known to influence the ability for repairing DNA damage in both parental genomes after fertilization (42). The inability to draw firm conclusions using meta-analyses stems from significant heterogeneity across studies. Factors such as the inclusion criteria, type of DNA fragmentation test, timing of performing DNA fragmentation test relative to the couple receiving IVF/intracytoplasmic sperm injection (ICSI) treatment and cut-off thresholds vary across individual studies (29, 43).

ARTIFICIAL OOCYTE ACTIVATION

Total fertilization failure (TFF) is a devastating outcome for couples undergoing IVF or ICSI. In ICSI cycles, TFF has been reported to occur in about 3% of cycles (44). Risk factors for TFF can include globoozospermia. However, TFF can also be unpredictable, following previous cycles in which normal fertilization has been observed (45). The underlying cause for fertilization failure can be either sperm or oocyte factors. Ultimately however it is due to a disruption in the normal sequence of molecular events which occur during fertilization. One of the critical steps leading to fertilization includes oocyte activation. This is a physiological process normally triggered on entrance of the spermatozoon into the oocyte, releasing PLCzeta. Oscillatory rises in calcium are then observed which are thought to trigger downstream events leading to fertilization and further embryo development (46).

Methods to artificially induce oocyte activation have been proposed as a possible treatment for fertilization failure. Different protocols exist but have been based on either chemical, mechanical or electrical stimulation of the oocyte following ICSI. These all ultimately lead to an influx of calcium into the ooplasm, theoretically mimicking the natural oocyte activation which occurs with normal fertilization (44).

Systematic reviews of RCTs concluded that there is insufficient clinical evidence to recommend its use in practice (47). This intervention requires further evaluation, particularly with studies which include live birth rate as an outcome.

The process of oocyte activation is thought to ultimately influence normal embryo development, epigenetic imprinting and pregnancy outcome. The significant deviations from a normal oscillatory calcium changes, as well as probably unknown molecular events occurring in artificial oocyte activation, are potentially causes for abnormal pregnancy, obstetric or neonatal outcomes (44). There is limited data available regarding this. One descriptive study including 47 children (from 237 cycles) reported that the birth characteristics and congenital malformations detected within 3 months of age are within the expected range (48). Interpretation of this however should be cautious owing to the small number of children included.

The HFEA state that oocyte activation with calcium ionophores may improve fertilization rates in ICSI cycle where failed fertilization has previously been observed. However, they acknowledge there are no RCTs to demonstrate that it is effective or follow up studies on the safety of this technique (49).

ADVANCED SPERM SELECTION TECHNIQUES

Advanced sperm selection techniques use methods to select healthy, mature, and genetically sound sperm for fertilization in the expectation that this will improve the outcomes of traditional IVF or ICSI treatment cycles. These methods include: ability to bind to hyaluronic acid (this is also known as physiological intracytoplasmic sperm injection [PICSI]); sperm selection according to surface charge (also known as the Zeta potential) (49); and according to sperm apoptosis using magnetic-activated cell sorting. These techniques are used in some centers around the world, but their effectiveness is unclear.

PICSI is a technique that co-incubates sperm with HA, a natural polymeric secretion of the cervical mucus and the cumulus-oocyte complex. Previous laboratory experiments have suggested that sperm which express the receptors to bind to HA have better morphology and motility as well as lower rates of sperm DNA fragmentation and better chromatin structure. Hence, PICSI might better identify sperm for use in ICSI treatment.

A Cochrane Review of eight studies using different techniques that included hyaluronic acid-selected sperm (or PICSI), Zeta potential and magnetic activating cell sorting did not report an increase in the likelihood of live birth (50). The largest randomized study of HA selected sperm (PICSI) did not increase the chances of a live birth (51). The evidence from the Cochrane review suggests that sperm selected by hyaluronic acid binding (PICSI) may have little or no effect on live birth or clinical pregnancy but may reduce miscarriage (50). The effect of Zeta sperm selection and magnetic activated cell sorting on live birth, clinical pregnancy, and miscarriage was uncertain due principally to the very low quality of the evidence for these techniques.

Further high-quality studies are required to evaluate whether any of these advanced sperm selection techniques can be recommended for use in routine practice. The HFEA currently states that there is no evidence that this add-on is effective and safe (3). To our knowledge there are no ASRM guidelines on this topic.

There is little data from RCTs on congenital abnormality in pregnancies utilizing advanced sperm selection techniques. Risks associated with the use of ICSI also apply to advanced sperm selection techniques; there are no significant additional risks to the patient or embryo. It is also unclear whether those with high sperm DNA fragmentation or other etiologies of subfertility, might benefit from these advanced sperm selection techniques. Further research should focus on the impact on congenital abnormalities and the impact in these subgroups.

INTRACYTOPLASMIC MORPHOLOGICALLY SELECTED SPERM INJECTION

Like PICSI, intracytoplasmic morphologically selected sperm injection (IMSI) it a technique employed to select the best sperm with fertilization potential. IMSI employs motile sperm organellar morphology examination to define the quality of six spermatozoa organelles: acrosome, post acrosomal

lamina, neck, mitochondria, tail and nucleus (51). This is done under higher magnification (×6000 to ×13,000) and requires more operator time. The disadvantage is the length of time required to examine and select the spermatozoa (52). Otherwise the risks faced by patients and embryos reflect those of routine ICSI (3).

A Cochrane review reported in 2013 very low-quality evidence with no improvement in clinical pregnancy, live birth rate or miscarriage rate with IMSI when compared with routine ICSI (53). Only two further randomized trials have been published since this review and neither reported an increase in live birth. (54, 55). A systematic review of both randomized and non-randomized studies suggested benefit in the non-randomized studies but not in the randomized studies (56). This serves as a reminder regarding the importance of performing well designed RCTs in order to allow correct conclusions to be drawn.

None of the trials included in the Cochrane review reported on congenital abnormalities. There has been one retrospective analysis of a cohort of babies born following IMSI compared to those born following ICSI published as a conference abstract. This demonstrated that IMSI led to a greater proportion of babies being born small for gestational age (< 2500g) compared to ICSI. No other difference in perinatal outcome was significant (57).

CONCLUSIONS

Laboratory techniques are critical for the success of medically assisted reproduction cycles. In the search for improvement in outcomes many innovations have been introduced, unfortunately often with little evidence of improved outcomes. Greater collaboration between clinicians, embryologists and patients to develop research priorities for the laboratory with the intention of developing a rational approach to evaluating innovations in the laboratory that might lead to improved outcomes is required. Currently, we have a menu of add-ons in the laboratory that has confused and mislead over the past decade. We could do better as our patients deserve it.

REFERENCES

1. Heneghan C, Spencer EA, Bobrovitz N, Collins DRJ, Nunan D, Plüddemann A, et al. Lack of evidence for interventions offered in UK fertility centres. BMJ 2016;355:392–4.
2. Harper J, Jackson E, Sermon K, Aitken RJ, Harbottle S, Mocanu E, et al. Adjuncts in the IVF laboratory: where is the evidence for 'add-on' interventions? Hum Reprod 2017;32:485–91.
3. Human Fertilization and Embryology Authority. Treatment add-ons 2018. Available at: https://www.hfea.gov.uk/treatments/explore-all-treatments/treatment-add-ons/. May 1, 2019.
4. Hong KH, Werner MD, Franasiak JM, Forman EJ, Prodoehl A, Upham K, et al. Embryologist interpretation of time-lapse imaging parameters at the blastocyst stage do not alter selection among transferred euploid blastocysts. Fertil Steril 2014;102:e305.
5. Scott L. The biological basis of non-invasive strategies for selection of human oocytes and embryos. Hum Reprod Update 2003;9:237–49.
6. Scott L. Pronuclear scoring as a predictor of embryo development. Reprod Biomed Online 2003;6:201–14.
7. Shoukir Y, Campana A, Farley T, Sakkas D. Early cleavage of in-vitro fertilized human embryos to the 2-cell stage: a novel indicator of embryo quality and viability. Hum Reprod 1997;12:1531–6.
8. Alpha-Scientists. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. Hum Reprod 2011;26:1270–83.
9. Cutting R, Morroll D, Roberts SA, Pickering S, Rutherford A. Elective single embryo transfer: guidelines for practice British Fertility Society and Association of Clinical Embryologists. Hum Fertil 2008;11:131–46.
10. Bhide P, Maheshwari A, Cutting R, Seenan S, Patel A, Khan K, et al. Time lapse imaging: is it time to incorporate this technology into routine clinical practice? Hum Fertil 2017;20:74–9.
11. Vitrolife. General 2015. Available at: http://www.vitrolife.com/en/Products/EmbryoScope-Time-Lapse-System/Support-Material/FAQ/General/. May 23, 2018.
12. Meseguer M, Rubio I, Cruz M, Basile N, Marcos J, Requena A. Embryo incubation and selection in a time-lapse monitoring system improves pregnancy outcome compared with a standard incubator: a retrospective cohort study. Fertil Steril 2012;98:1481–9.e10.
13. Armstrong S, Vail A, Mastenbroek S, Jordan V, Farquhar C. Time-lapse in the IVF-lab: how should we assess potential benefit? Hum Reprod 2015;30:3–8.
14. Conaghan J, Chen AA, Willman SP, Ivani K, Chenette PE, Boostanfar R, et al. Improving embryo selection using a computer-automated time-lapse image analysis test plus day 3 morphology: results from a prospective multicenter trial. Fertil Steril 2013;100:412–9.e5.
15. Swain JE, Schoolcraft WB, Bossert N, Batcheller AE. Media osmolality changes over 7 days following culture in a non-humidified benchtop incubator. Fertil Steril 2016;106:e362.
16. Armstrong S, Bhide P, Jordan V, Pacey A, Marjoribanks J, Farquhar C. Time-lapse systems for embryo incubation and assessment in assisted reproduction. Cochrane Database Syst Rev 2019;5:CD011320.
17. Cohen J. Assisted hatching of human embryos. J In Vitro Fert Embryo Transf 1991;8:179–90.
18. Cohen J, Alikani M, Trowbridge J, Rosenwaks Z. Implantation enhancement by selective assisted hatching using zona drilling of human embryos with poor prognosis. Hum Reprod 1992;7:685–91.
19. Carney SK, Das S, Blake D, Farquhar C, Seif MM, Nelson L. Assisted hatching on assisted conception (in vitro fertilization (IVF) and intracytoplasmic sperm injection (ICSI). Cochrane Database Syst Rev 2012;12:CD001894.
20. Li D, Yang DL, An J, Jiao J, Zhou YM, Wu QJ, et al. Effect of assisted hatching on pregnancy outcomes: a systematic review and meta-analysis of randomized controlled trials. Sci Rep 2016;6:31228.
21. Mateizel I, Santos-Ribeiro S, Done E, Van Landuyt L, Van de Velde H, Tournaye H, et al. Do ARTs affect the incidence of monozygotic twinning? Hum Reprod 2016;31:2435–41.
22. NICE. Fertility problems: assessment and treatment (CG156) 2017. Available at: https://www.nice.org.uk/guidance/CG156. May 1, 2019.
23. American Society for Reproductive, Society for Assisted Reproductive Technology. Role of assisted hatching in in vitro fertilization: a guideline. Fertil Steril 2014;102:348–51.
24. Schoolcraft WB, Katz-Jaffe MG. Comprehensive chromosome screening of trophectoderm with vitrification facilitates elective single-embryo transfer for infertile women with advanced maternal age. Fertil Steril 2013;100:615–9.
25. Bontekoe S, Heineman MJ, Johnson N, Blake D. Adherence compounds in embryo transfer media for assisted reproductive technologies. Cochrane Database Syst Rev 2014:CD007421.
26. Fancsovits P, Lehner A, Murber A, Kaszas Z, Rigo J, Urbancsek J. Effect of hyaluronan-enriched embryo transfer medium on IVF outcome: a prospective randomized clinical trial. Arch Gynecol Obstet 2015;291:1173–9.
27. Leuchtenberger C, Schrader F, Weir DR, Gentile DP. The desoxyribosenucleic acid (DNA) content in spermatozoa of fertile and infertile human males. Chromosoma 1953;6:61–78.
28. De Jonge C. Semen analysis: looking for an upgrade in class. Fertil Steril 2012;97:260–6.
29. Cissen M, Wely MV, Scholten I, Mansell S, Bruin JP, Mol BW, et al. Measuring Sperm DNA Fragmentation and Clinical Outcomes of Medically Assisted Reproduction: A Systematic Review and Meta-Analysis. PLoS One 2016;11:e0165125.

30. Collins JA, Barnhart KT, Schlegel PN. Do sperm DNA integrity tests predict pregnancy with in vitro fertilization? Fertil Steril 2008;89:823–31.
31. Deng C, Li T, Xie Y, Guo Y, Yang QY, Liang X, et al. Sperm DNA fragmentation index influences assisted reproductive technology outcome: A systematic review and meta-analysis combined with a retrospective cohort study. Andrologia 2019;51:e13263.
32. Evenson D, Wixon R. Meta-analysis of sperm DNA fragmentation using the sperm chromatin structure assay. Reprod Biomed Online 2006;12:466–72.
33. Li Z, Wang L, Cai J, Huang H. Correlation of sperm DNA damage with IVF and ICSI outcomes: a systematic review and meta-analysis. J Assist Reprod Genet 2006;23:367–76.
34. Osman A, Alsomait H, Seshadri S, El-Toukhy T, Khalaf Y. The effect of sperm DNA fragmentation on live birth rate after IVF or ICSI: a systematic review and meta-analysis. Reprod Biomed Online 2015;30:120–7.
35. Simon L, Zini A, Dyachenko A, Ciampi A, Carrell DT. A systematic review and meta-analysis to determine the effect of sperm DNA damage on in vitro fertilization and intracytoplasmic sperm injection outcome. Asian J Androl 2017;19:80–90.
36. Zhao J, Zhang Q, Wang Y, Li Y. Whether sperm deoxyribonucleic acid fragmentation has an effect on pregnancy and miscarriage after in vitro fertilization/intracytoplasmic sperm injection: a systematic review and meta-analysis. Fertil Steril 2014;102:998–1005.e8.
37. Smits RM, Mackenzie-Proctor R, Yazdani A, Stankiewicz MT, Jordan V, Showell MG. Antioxidants for male subfertility. Cochrane Database Syst Rev 2019;3:CD007411.
38. Evenson DP, Jost LK, Marshall D, Zinaman MJ, Clegg E, Purvis K, et al. Utility of the sperm chromatin structure assay as a diagnostic and prognostic tool in the human fertility clinic. Hum Reprod 1999;14:1039–49.
39. Spano M, Bonde JP, Hjollund HI, Kolstad HA, Cordelli E, Leter G. Sperm chromatin damage impairs human fertility. The Danish First Pregnancy Planner Study Team. Fertil Steril 2000;73:43–50.
40. Practice Committee of the American Society for Reproductive M. The clinical utility of sperm DNA integrity testing: a guideline. Fertil Steril 2013; 99:673–7.
41. Tomlinson M, Lewis S, Morroll D, British Fertility S. Sperm quality and its relationship to natural and assisted conception: British Fertility Society guidelines for practice. Hum Fertil 2013;16:175–93.
42. Ashwood-Smith MJ, Edwards RG. DNA repair by oocytes. Mol Hum Reprod 1996;2:46–51.
43. Pacey A. Is sperm DNA fragmentation a useful test that identifies a treatable cause of male infertility? Best Pract Res Clin Obstet Gynaecol 2018;53:11–9.
44. Anifandis G, Michopoulos A, Daponte A, Chatzimeletiou K, Simopoulou M, Messini CI, et al. Artificial oocyte activation: physiological, pathophysiological and ethical aspects. Syst Biol Reprod Med 2019;65:3–11.
45. Karabulut S, Aksunger O, Ata C, Sagiroglu Y, Keskin I. Artificial oocyte activation with calcium ionophore for frozen sperm cycles. Syst Biol Reprod Med 2018;64:381–8.
46. Yeste M, Jones C, Amdani SN, Patel S, Coward K. Oocyte activation deficiency: a role for an oocyte contribution? Hum Reprod Update 2016;22: 23–47.
47. Murugesu S, Saso S, Jones BP, Bracewell-Milnes T, Athanasiou T, Mania A, et al. Does the use of calcium ionophore during artificial oocyte activation demonstrate an effect on pregnancy rate? A meta-analysis. Fertil Steril 2017;108:468–82.e3.
48. Mateizel I, Verheyen G, Van de Velde H, Tournaye H, Belva F. Obstetric and neonatal outcome following ICSI with assisted oocyte activation by calcium ionophore treatment. J Assist Reprod Genet 2018;35:1005–10.
49. Human Fertilization and Embryology Authority. Treatmen add-ons. Available at: https://www.hfea.gov.uk/treatments/explore-all-treatments/treatment-add-ons/.
50. McDowell S, Kroon B, Ford E, Hook Y, Glujovsky D, Yazdani A. Advanced sperm selection techniques for assisted reproduction. Cochrane Database Syst Rev 2014;10:CD010461.
51. Bartoov B, Berkovitz A, Eltes F, Kogosowski A, Menezo Y, Barak Y. Real-time fine morphology of motile human sperm cells is associated with IVF-ICSI outcome. J Androl 2002;23:1–8.
52. De Vos A, Polyzos NP, Verheyen G, Tournaye H. Intracytoplasmic morphologically selected sperm injection (IMSI): a critical and evidence-based review. Basic Clin Androl 2013;23:10.
53. Teixeira DM, Barbosa MAP, Ferriani RA, Navarro PA, Raine-Fenning N, Nastri CO, Martins WP. Regular (ICSI) versus ultra-high magnification (IMSI) sperm selection for assisted reproduction. Cochrane Database Syst Rev 2013:CD010167.
54. Leandri RD, Gachet A, Pfeffer J, Celebi C, Rives N, Carre-Pigeon F, et al. Is intracytoplasmic morphologically selected sperm injection (IMSI) beneficial in the first ART cycle? A multicentric randomized controlled trial. Andrology 2013;1:692-697.
55. La Sala GB, Nicoli A, Fornaciari E, Falbo A, Rondini I, Morini D, et al. Intracytoplasmic morpho-logically selected sperm injection versus conventional intracyto-plasmic sperm injection: a randomized controlled trial. Reprod Biol Endocrinol 2015;13:97.
56. Duran-Retamal M, Morris G, Achilli C, Gaunt M, Theodorou E, Saab W, et al. Live birth and miscarriage rate following intracytoplasmic morphologically selected sperm injection vs intracytoplasmic sperm injection: An updated systematic review and meta-analysis. Acta Obstet Gynecol Scand 2019.
57. Junca A, Dumont M, Cornet D, Douard S, De Mouzon J, Prisant N. Is intracytoplasmic morphologically sperm injection (IMSI) detrimental for pregnancy outcome? Fertil Steril 2010;94:S31.

Paper 2: Time-lapse imaging for embryo incubation and assessment in assisted reproduction

The first iteration of this Cochrane systematic review was first authored by myself and published in 2015 [103]. It included only three RCTs, involving 994 women. In 2018 the review was completely re-run as part of this MD. This time, it included 8 RCTs, with a total of 2303 women. The systematic review was undertaken according to strict Cochrane methodological criteria, and the review underwent peer review with the Gynaecology and Fertility Cochrane Group Editorial Board. The conclusion of the review didn't change from that of the 2015 iteration. There was still insufficient evidence of differences in livebirth, miscarriage, stillbirth, or clinical pregnancy to choose between time-lapse systems (TLS), with or without embryo selection software, and conventional incubation.

Following the publication of this review, two important things occurred. Firstly, it was recognised that a large, high quality RCT had been published since the publication of the Cochrane review, and secondly, feedback was received from authors of one of the included RCTs explaining that an error had been made in how the trial was classified. Following reflection on these two issues with Prof. Allan Pacey it was concluded that it would serve patients, clinicians, and the scientific community best if the review was re-run from the beginning so as to capture all recently published RCTs, and to correct the error made in classifying an included study as one that utilised embryo selection software [60].

The updated review was undertaken over the course of 6 months and was published in May 2019. The 2019 review is reproduced below in its published format, minus the following sections for the sake of brevity: characteristics of studies awaiting assessment; characteristics of ongoing studies; and the appendices, which include the search strategies and the feedback section. These can be read in the full version of the published review [104].

Cochrane Library

Cochrane Database of Systematic Reviews

Time-lapse systems for embryo incubation and assessment in assisted reproduction (Review)

Armstrong S, Bhide P, Jordan V, Pacey A, Marjoribanks J, Farquhar C

Armstrong S, Bhide P, Jordan V, Pacey A, Marjoribanks J, Farquhar C.
Time-lapse systems for embryo incubation and assessment in assisted reproduction.
Cochrane Database of Systematic Reviews 2019, Issue 5. Art. No.: CD011320.
DOI: 10.1002/14651858.CD011320.pub4.

www.cochranelibrary.com

[Intervention Review]

Time-lapse systems for embryo incubation and assessment in assisted reproduction

Sarah Armstrong[1], Priya Bhide[2], Vanessa Jordan[3], Allan Pacey[4], Jane Marjoribanks[3], Cindy Farquhar[3]

[1]Department of Oncology & Metabolism, University of Sheffield, Sheffield, UK. [2]Homerton University Hospital NHS Foundation Trust, London, UK. [3]Department of Obstetrics and Gynaecology, University of Auckland, Auckland, New Zealand. [4]Department of Oncology & Metabolism, Academic Unit of Reproductive and Developmental Medicine, The University of Sheffield, Sheffield, UK

Contact address: Sarah Armstrong, Department of Oncology & Metabolism, University of Sheffield, Academic Unit of Reproductive and Developmental Medicine, Level 4, The Jessop Wing, Sheffield, S10 2SF, UK. sarahcarmstrong@yahoo.co.uk.

Editorial group: Cochrane Gynaecology and Fertility Group.
Publication status and date: New search for studies and content updated (no change to conclusions), published in Issue 5, 2019.

Citation: Armstrong S, Bhide P, Jordan V, Pacey A, Marjoribanks J, Farquhar C. Time-lapse systems for embryo incubation and assessment in assisted reproduction. *Cochrane Database of Systematic Reviews* 2019, Issue 5. Art. No.: CD011320. DOI: 10.1002/14651858.CD011320.pub4.

ABSTRACT

Background

Embryo incubation and assessment is a vital step in assisted reproductive technology (ART). Traditionally, embryo assessment has been achieved by removing embryos from a conventional incubator daily for quality assessment by an embryologist, under a microscope. In recent years time-lapse systems (TLS) have been developed which can take digital images of embryos at frequent time intervals. This allows embryologists, with or without the assistance of embryo selection software, to assess the quality of the embryos without physically removing them from the incubator.

The potential advantages of a TLS include the ability to maintain a stable culture environment, therefore limiting the exposure of embryos to changes in gas composition, temperature, and movement. A TLS has the potential advantage of improving embryo selection for ART treatment by utilising additional information gained through continuously monitoring embryo development. Use of a TLS often adds significant extra cost to ART treatment.

Objectives

To determine the effect of a TLS compared to conventional embryo incubation and assessment on clinical outcomes in couples undergoing ART.

Search methods

We used standard methodology recommended by Cochrane. We searched the Cochrane Gynaecology and Fertility (CGF) Group Trials Register, CENTRAL, MEDLINE, Embase, CINAHL, and two trials registers on 7 January 2019 and checked references of appropriate papers.

Selection criteria

We included randomised controlled trials (RCTs) comparing TLS, with or without embryo selection software, versus conventional incubation with morphological assessment; and TLS with embryo selection software versus TLS without embryo selection software among couples undergoing ART.

Data collection and analysis

We used standard methodological procedures recommended by Cochrane. The primary review outcomes were live birth or ongoing pregnancy, miscarriage and stillbirth, and cumulative live birth or ongoing pregnancy rate. The secondary outcomes were clinical

pregnancy and cumulative clinical pregnancy. We assessed the quality of the evidence using GRADE methodology. We made the following comparisons.

TLS with conventional morphological assessment of still TLS images versus conventional incubation and assessment

TLS utilising embryo selection software versus TLS with conventional morphological assessment of still TLS images

TLS utilising embryo selection software versus conventional incubation and assessment

Main results

We included nine RCTs (N = 2955 infertile couples). The quality of the evidence ranged from very low to low. The main limitations were high risk of bias in the included studies, imprecision, indirectness, and inconsistency. There were no data on cumulative live birth or ongoing pregnancy rate or cumulative clinical pregnancy rate.

TLS with conventional morphological assessment of still TLS images versus conventional incubation and assessment

It is unclear whether there is any difference between interventions in rates of live birth or ongoing pregnancy (odds ratio (OR) 0.91, 95% confidence interval (CI) 0.67 to 1.23, 3 RCTs, N = 826, I^2 = 33%, low-quality evidence) or in miscarriage rates (OR 1.90, 95% CI 0.99 to 3.61, 3 RCTs, N = 826, I^2 = 0%, low-quality evidence). The evidence suggests that if the rate of live birth or ongoing pregnancy associated with conventional incubation and assessment is 35%, the rate with the use of TLS with conventional morphological assessment of still TLS images would be between 27% and 40%, and if the miscarriage rate with conventional incubation is 4%, the rate associated with conventional morphological assessment of still TLS images would be between 4% and 14%. It is unclear whether there is a difference between the interventions in rates of stillbirth (OR 1.00, 95% CI 0.13 to 7.49, 1 RCT, N = 76, low-quality evidence) or clinical pregnancy (OR 1.06, 95% CI 0.79 to 1.41, 4 RCTs, N = 875, I^2 = 0%, low-quality evidence).

TLS utilising embryo selection software versus TLS with conventional morphological assessment of still TLS images

All findings for this comparison were very uncertain due to the very low-quality of the evidence. No data were available on live birth, but one RCT reported ongoing pregnancy. It is unclear whether there is any difference between the interventions in rates of ongoing pregnancy (OR 0.61, 95% CI 0.32 to 1.20, 1 RCT, N = 163); miscarriage (OR 1.39, 95% CI 0.64 to 3.01, 2 RCTs, N = 463, I^2 = 0%); or clinical pregnancy (OR 0.97, 95% CI 0.67 to 1.42, 2 RCTs, N = 463, I^2 = 0%). The evidence suggests that if the rate of ongoing pregnancy associated with TLS with conventional morphological assessment of still TLS images is 47%, the rate associated with TLS utilising embryo selection software would be between 22% and 52%, and if the miscarriage rate associated with conventional morphological assessment of still TLS images is 5%, the rate associated with TLS utilising embryo selection software would be between 4% and 15%. No studies reported stillbirth.

TLS utilising embryo selection software versus conventional incubation and assessment

The findings for this comparison were also very uncertain due to the very low quality of the evidence. It is unclear whether there is any difference between the interventions in rates of live birth (OR 1.12, 95% CI 0.92 to 1.36, 3 RCTs, N = 1617, I^2 = 84%). There was very low-quality evidence that TLS might reduce miscarriage rates (OR 0.63, 95% CI 0.45 to 0.89, 3 RCTs, N = 1617, I^2 = 0%). It is unclear whether there is any difference between the interventions in rates of clinical pregnancy (OR 0.95, 95% CI 0.78 to 1.16, 3 RCTs, N = 1617, I^2 = 89%). The evidence suggests that if the rate of live birth associated with conventional incubation and assessment is 48%, the rate with TLS utilising embryo selection software would be between 46% and 55%, and if the miscarriage rate with conventional incubation and assessment is 11%, the rate associated with TLS would be between 5% and 10%. No stillbirths occurred in the only study reporting this outcome.

Authors' conclusions

There is insufficient good-quality evidence of differences in live birth or ongoing pregnancy, miscarriage and stillbirth, or clinical pregnancy to choose between TLS, with or without embryo selection software, and conventional incubation. As the evidence is of low or very low-quality, our findings should be interpreted with caution.

PLAIN LANGUAGE SUMMARY

Time-lapse systems for embryo incubation and embryo assessment for couples undergoing in vitro fertilisation and intracytoplasmic sperm injection

Review question

Does a time-lapse system (TLS) improve the chances of a pregnancy and live-born baby, and reduce the risk of miscarriage and stillbirth?

Background

In vitro fertilisation (IVF) and intracytoplasmic sperm injection (ICSI) are processes whereby a woman's eggs and a man's sperm are combined to achieve fertilisation outside of the body. Embryos are stored in an incubator and replaced into the woman between day 2

and 5 of development. Usually, embryos are removed from an incubator for assessment, under a microscope, of their quality and stage of development. A TLS can take images of embryos at frequent time intervals, which allows assessment without removing the embryos from the incubator. A TLS can also apply software that assists the embryologist in selecting the best-quality embryo for replacement, potentially improving the chance of a baby.

Study characteristics

The evidence is current to January 2019. We included nine studies (randomised controlled trials, that is studies in which participants are assigned to one of two or more treatment groups using a random method) of 2955 infertile couples undergoing IVF or ICSI. There were three different study designs: (1) TLS with conventional assessment of still TLS images versus conventional incubation and assessment; (2) TLS utilising embryo selection software versus TLS with conventional assessment of still TLS images; and (3) TLS utilising embryo selection software versus conventional incubation and assessment.

What the review found

TLS with conventional assessment of still TLS images versus conventional incubation and assessment

All the evidence for this comparison was low-quality. It is unclear whether there is any difference between the interventions in rates of livebirth or ongoing pregnancy or miscarriage. The evidence suggests that if the rate of livebirth or ongoing pregnancy associated with conventional incubation and assessment is 35%, the rate with use of TLS with conventional morphological assessment of still TLS images would be between 27% and 40%, and if the miscarriage rate with conventional incubation is 4%, the rate associated with conventional morphological assessment of still TLS images would be between 4% and 14%. It is unclear whether there is a difference between interventions in rates of stillbirth or clinical pregnancy.

TLS utilising embryo selection software versus TLS with conventional assessment of still TLS images

All findings for this comparison were very uncertain due to very low-quality evidence. No data were available on livebirth, but one study reported ongoing pregnancy. It is unclear whether there is any difference between interventions in rates of ongoing pregnancy, miscarriage, or clinical pregnancy. The evidence suggests that if the rate of ongoing pregnancy associated with TLS with conventional morphological assessment of still TLS images is 47%, the rate associated with TLS utilising embryo selection software would be between 22% and 52%, and if the miscarriage rate associated with conventional morphological assessment of still TLS images is 5%, the rate associated with TLS utilising embryo selection software would be between 4% and 15%. No studies reported stillbirth.

TLS utilising embryo selection software versus conventional incubation and assessment

All findings for this comparison were very uncertain due to the very low-quality of the evidence. It is unclear whether there is any difference between interventions with respect to rates of livebirth or clinical pregnancy. The evidence suggests lower rates of miscarriage in the TLS group for the outcome of miscarriage. The evidence suggests that if the livebirth rate associated with conventional incubation is 48%, the rate with the use of TLS would be between 46% and 55%, and if the miscarriage rate with conventional incubation is 11%, the rate associated with TLS would be between 5% and 10%.

Overall conclusions

There is no good evidence showing that TLS is more or less effective than conventional methods of embryo incubation. Patients may wish to take part in randomised controlled trials on TLS in order to add to the existing evidence base and to help guide assisted reproductive technology patients in the future.

Quality of the evidence

The quality of the evidence ranged from very low to low. The main limitations were high risk of bias in the included studies, imprecision, indirectness, and inconsistency.

SUMMARY OF FINDINGS

Summary of findings for the main comparison. TLS with conventional morphological assessment of still TLS images compared to conventional incubation and assessment for embryo incubation and assessment in assisted reproduction

TLS with conventional morphological assessment of still TLS images compared to conventional incubation and assessment for embryo incubation and assessment in assisted reproduction

Patient or population: couples undergoing assisted reproductive technology
Setting: fertility clinic
Intervention: TLS with conventional morphological assessment of still TLS images
Comparison: conventional incubation and assessment

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)
	Risk with conventional incubation and assessment	Risk with TLS with conventional morphological assessment of still TLS images			
Live birth or ongoing pregnancy	353 per 1000	332 per 1000 (268 to 402)	OR 0.91 (0.67 to 1.23)	826 (3 RCTs)	⊕⊕⊕⊖ Low[a]
Miscarriage	42 per 1000	77 per 1000 (42 to 137)	OR 1.90 (0.99 to 3.61)	826 (3 RCTs)	⊕⊕⊖⊖ Low[b]
Stillbirth	12 per 1000	12 per 1000 (2 to 86)	OR 1.00 (0.13 to 7.49)	76 (1 RCT)	⊕⊕⊖⊖ Low[c]
Clinical pregnancy	374 per 1000	388 per 1000 (321 to 458)	OR 1.06 (0.79 to 1.41)	875 (4 RCTs)	⊕⊕⊕⊖ Low[d]

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

CI: confidence interval; **OR:** odds ratio; **RCT:** randomised controlled trial; **TLS:** time-lapse system

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate quality: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low quality: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low quality: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[a]We downgraded our assessment of the quality of the evidence for live birth or ongoing pregnancy once for serious risk of performance bias and once for serious imprecision due to wide confidence intervals, compatible with a benefit in either group.
[b]We downgraded our assessment of the evidence for miscarriage once for serious risk of performance bias and once for serious imprecision due to wide confidence intervals and small number of events (total of 48).
[c]We downgraded our assessment of the quality of the evidence for stillbirth once for serious risk of performance bias and once for serious imprecision. Although two studies examined this outcome, one had no events in either arm and was therefore removed from meta-analysis in accordance with Cochrane guidance. This left a single small study with very wide confidence intervals and only four events.
[d]We downgraded our assessment of the quality of the evidence for clinical pregnancy once for serious risk of performance bias and once for serious imprecision, due to wide confidence intervals compatible with a benefit in either group.

Summary of findings 2. TLS utilising embryo selection software compared to TLS with conventional morphological assessment of still TLS images for embryo incubation and assessment in assisted reproduction

TLS utilising embryo selection software compared to TLS with conventional morphological assessment of still TLS images for embryo incubation and assessment in assisted reproduction

Patient or population: couples undergoing assisted reproductive technology
Setting: fertility clinic
Intervention: TLS utilising embryo selection software
Comparison: TLS with conventional morphological assessment of still TLS images

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with TLS with conventional morphological assessment of still TLS images	Risk with TLS utilising embryo selection software				
Live birth or ongoing pregnancy	472 per 1000	353 per 1000 (222 to 517)	OR 0.61 (0.32 to 1.20)	163 (1 RCT)	Very low[a]	The outcome was ongoing pregnancy; no live-birth data were available.
Miscarriage	54 per 1000	74 per 1000 (35 to 147)	OR 1.39 (0.64 to 3.01)	463 (2 RCTs)	Very low[b]	
Stillbirth	No studies reported this outcome.	-	-	-	-	
Clinical pregnancy	537 per 1000	529 per 1000 (437 to 622)	OR 0.97 (0.67 to 1.42)	463 (2 RCTs)	Very low[c]	

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

CI: confidence interval; **OR:** odds ratio; **RCT:** randomised controlled trial; **TLS:** time-lapse system

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate quality: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low quality: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low quality: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[a]We downgraded our assessment of the quality of the evidence for live birth or ongoing pregnancy once for serious risk of performance bias and twice for very serious imprecision due to there being only one RCT with a small number of events (64) and wide confidence intervals compatible with a benefit in either group.
[b]We downgraded our assessment of the quality of the evidence for miscarriage once for serious risk of performance bias; once for serious indirectness (heterogeneity between the study designs: one included study involved removing embryos for benchtop microscopy daily in both the intervention and control arms, whereas the other study left embryos in the intervention and control arms undisturbed); and once for serious imprecision (wide confidence intervals compatible with a benefit in either group and a low number of events overall (N = 29)).
[c]We downgraded our assessment of the quality of the evidence for clinical pregnancy once for serious risk of performance bias, once for serious indirectness (as described above), and once for serious imprecision (wide confidence intervals compatible with a benefit in either group).

Summary of findings 3. TLS utilising embryo selection software compared to conventional incubation and assessment for embryo incubation and assessment in assisted reproduction

TLS utilising embryo selection software compared to conventional incubation and assessment for embryo incubation and assessment in assisted reproduction

Patient or population: couples undergoing ART
Setting: fertility clinic
Intervention: TLS utilising embryo selection software
Comparison: conventional incubation and assessment

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)
	Risk with conventional incubation and assessment	Risk with TLS utilising embryo selection software			
Live birth or ongoing pregnancy	475 per 1000	504 per 1000 (455 to 554)	OR 1.12 (0.92 to 1.36)	1617 (3 RCTs)	Very low[a]
Miscarriage	108 per 1000	71 per 1000 (52 to 98)	OR 0.63 (0.45 to 0.89)	1617 (3 RCTs)	Very low[b]
Stillbirth	No events occurred in the only study reporting this outcome.		-	600 (1 RCT)	-

Clinical pregnancy	605 per 1000	593 per 1000 (545 to 640)	OR 0.95 (0.78 to 1.16)	1617 (3 RCTs)	Very low[c]

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

CI: confidence interval; **OR:** odds ratio; **RCT:** randomised controlled trial; **TLS:** time-lapse system

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

[a]We downgraded our assessment of the quality of the evidence for live birth twice for very serious risk of bias (high risk of both performance bias and selection bias in two studies, and of other bias in the third study). In one study, the randomisation of participants was undertaken by the principal investigator, and allocation concealment was not described. In another study, some participants could request the intervention, and this request was granted. In the third study, the day of transfer varied between the two study arms. We also downgraded our assessment of the quality of the evidence once for serious indirectness, as one included study undertook multiple embryo transfers per woman and included women receiving donor oocytes from younger women. Although further downgrading was not possible, there was also serious inconsistency (I^2 = 86%), possibly secondary to differing embryo transfer policies across the studies: one study had blastocyst transfers, one had varied days of transfer, and one had day 3 transfer for the intervention arm and day 5 transfer for the control arm.
[b]We downgraded our assessment of the quality of the evidence for miscarriage twice for very serious risk of bias (as outlined above) and once for serious indirectness secondary to one included study including miscarriages of biochemical pregnancies as well as clinical pregnancies. The authors of the study were unable to separate these miscarriage data.
[c]We downgraded our assessment of the quality of the evidence for clinical pregnancy twice for very serious risk of bias and once for serious indirectness, as one included study undertook multiple embryo transfers per woman and included women receiving donor oocytes from younger women. Although further downgrading was not possible, there was also serious inconsistency (I^2 = 89%), possibly secondary to differing embryo transfer policies across the studies: one study had blastocyst transfers, one had varied days of transfer, and one had day 3 transfer for the intervention arm and day 5 transfer for the control arm.

BACKGROUND

Description of the condition

Embryo incubation is a critical step in all in vitro fertilisation (IVF) procedures. Embryo development within media in culture dishes in an incubator is a dynamic process, moving through the fertilisation stage to cleavage stage and then to the blastocyst stage in some cases. Throughout the incubation period, embryos are usually inspected at specific time points to provide a brief 'snapshot' assessment of the way the embryo is developing (morphological features). Embryologists apply a tiered grading system based on the morphology of the embryo in order to predict the potential for implantation and a successful pregnancy (Cummins 1986; Neuber 2003; Scott 2003; Scott 2003a; Shoukir 1997). A consensus on the minimum data set required for the accurate description of embryo morphology was established by Alpha Scientists in Reproductive Medicine and European Society of Human Reproduction and Embryology (ESHRE) Special Interest Group of Embryology (Alpha & ESHRE SIG 2011). A consensus on timings of observation of fertilised oocytes and embryos was established and deemed critical to the ability to compare results between different laboratories. The recommended checks, in hours, following insemination are:

- a fertilisation check at 17 hours, a syngamy (fusion of gametes) check at 23 hours;
- an early cleavage check at 26 hours post-intracytoplasmic sperm injection (ICSI) or 28 hours post-IVF;
- day 2 embryo assessment at 44 hours;
- day 3 embryo assessment at 68 hours;
- day 4 embryo assessment at 92 hours;
- day 5 embryo assessment at 116 hours.

Traditionally, the checks have been achieved by physically removing embryos from the controlled environment of the incubator to analyse them under a light microscope for assessment of embryo development and quality. This practice exposes the embryos to the potentially suboptimal conditions of the environment outside of the incubator and human handling (Meseguer 2012a). Time-lapse systems (TLSs) have evolved over recent years to increase the frequency of morphological observations whilst minimising the impact of the external environment and human handling on embryo development.

Description of the intervention

A TLS is a device that takes digital images of embryos at set time intervals, for example every 5 to 15 minutes. The system can be installed into an existing embryo incubator or can exist as a combined time-lapse incubation system. The images are compiled using software to create a time-lapse sequence of embryo development. Images can be digitally displayed as a time-lapse sequence on an external monitor to allow embryologists to assess the dynamic morphology of embryos, thus negating the need for the embryologist to remove the embryos from the incubator. Some TLSs also utilise computer-assisted assessment of developmental milestones of embryos, also known as morphokinetic parameters, to offer a semiquantitative process of embryo evaluation (Conaghan 2013). These cell-tracking software algorithms utilise data such as the timing of embryonic development events, and have evolved as a non-invasive, non-subjective way of attempting to improve the selection of embryos with the highest implantation potential. Some clinics have developed their own algorithms to adapt the standardised one that comes with the TLS device (Petersen 2016).

There are a number of commercially available TLSs developed by various manufacturers. Time-lapse systems are available as devices that can be placed within existing conventional incubators, and some exist with an integrated incubator. The integrated TLS combines both the time-lapse cameras and the incubator in one device.

How the intervention might work

There are two potential benefits of a TLS. Firstly, an advantage may lie with the undisturbed nature of the culture conditions, whereby images for embryo assessment can be obtained without removing embryos from the incubator environment for conventional benchtop light microscopy (which usually includes heated microscope stages). This minimises the exposure of embryos to both human handling and changes in air temperature and gas composition, which may lead to improved culture conditions.

A second potential advantage may be the ability of a TLS to accumulate detailed time-lapse images of embryo development at regular time intervals. This includes the timing of cell divisions, intervals between cell cycles, and other development factors (e.g. dynamic pronuclei patterns, presence of multinucleation and fragmentation, and blastomere symmetry). Many of these features that are transient events may be missed by using standard morphological assessment at set time intervals. These detailed time-lapse sequences can be utilised with or without cell-tracking software algorithms as an adjunct to standard morphological assessment, to select the embryo with the highest implantation potential for transfer. This is important because there is a clear correlation between embryo morphology and viability (Finn 2010; Neuber 2006). The ability to select the highest-quality embryo at an optimal stage of development for replacement in an assisted reproductive technology (ART) cycle may lead to a reduction in time to pregnancy and a reduced need for subsequent embryo transfers. It is worth noting that the different models of TLS follow the same basic principles but vary in technical detail such as gas mixture, temperature, group or single culture, and dark- or light-field microscopy.

In order to assess the potential advantage of TLSs (i.e. the stable culture environment or the time-lapse sequence of images which can be assessed with cell-tracking algorithms, or both), studies can be grouped into the following three comparisons.

Trial design 1: TLS with conventional morphological assessment of still TLS images versus conventional incubation and assessment

- These studies control for how the embryos are selected for transfer, but the incubation differs. This will help to establish whether the culture conditions of the TLS potentially impact on favourable outcomes such as pregnancy and live birth.

Trial design 2: TLS utilising embryo selection software versus TLS with conventional morphological assessment of still TLS images

- These studies control for the culture environment, with both arms of the trial being incubated in a TLS, and the way in

which embryos are selected for transfer is tested. This study design will help to establish whether embryo selection software improves the selection of top-quality embryos and increases the pregnancy and live-birth rate.

Trial design 3: TLS utilising embryo selection software versus conventional incubation and assessment

- These studies aim to establish whether a combination of both the stable culture environment and the embryo selection software is superior to conventional embryo incubation and assessment at improving pregnancy and live birth rates.

Why it is important to do this review

New interventions such as TLSs should be evaluated by randomised controlled trials in order to establish their safety, clinical effectiveness, and cost-effectiveness (Campbell 2000; Harper 2012). Countering the potential benefits outlined above, a TLS involves exposing embryos to light during image acquisition, at predetermined intervals. Furthermore, the authorities responsible for the regulation of fertility clinics and research involving human embryos have a responsibility to provide impartial and authoritative information to prospective and current patients on fertility treatments to aid them in making informed decisions about their care (ACART; HFEA). It is therefore vital that up-to-date and thorough systematic reviews that are accessible to patients and healthcare workers are published on the topic. This will enable information on the technology's success rates in terms of live birth or ongoing pregnancy rate, and safety in terms of adverse events, to be accessible and help guide informed decision making.

This is the third update of this Cochrane Review published under the same title initially 2015, Armstrong 2015, and again in 2018 (Armstrong 2018a). This update captures all newly available trial data and corrects an error in Analysis 3.1 in Armstrong 2018a.

We aimed with this updated review to establish whether there is evidence of any overall benefit of culturing embryos in a TLS with or without embryo selection software, over current conventional embryo incubation and assessment.

OBJECTIVES

To determine the effect of a time-lapse system (TLS) compared to conventional embryo incubation and assessment on clinical outcomes in couples undergoing assisted reproductive technology (ART).

METHODS

Criteria for considering studies for this review

Types of studies

Inclusions: any randomised controlled trial (RCT), whether published or not, which in principle could answer questions regarding clinical (postimplantation) outcomes.

Exclusions: quasi-randomised and other concurrently controlled studies were excluded. We excluded trials that randomised oocytes or embryos, as it would not be possible to compare clinical outcomes. We excluded cross-over trials as the design is not valid in this context.

Types of participants

Couples of any age undergoing assisted reproduction where embryo incubation was required.

Types of interventions

- Time-lapse system (TLS) with conventional morphological assessment of still TLS images versus conventional incubation and assessment (trial design 1)
- TLS utilising embryo selection software versus TLS with conventional morphological assessment of still TLS images (trial design 2)
- TLS utilising embryo selection software versus conventional incubation and assessment (trial design 3)

Any type of TLS, using any type of embryo selection software and any type of conventional incubator, was eligible.

Types of outcome measures

Primary outcomes

- Live-birth or ongoing pregnancy rate
- Miscarriage and stillbirth
- Cumulative live birth or ongoing pregnancy rate

Secondary outcomes

- Clinical pregnancy, defined as evidence of a gestational sac, confirmed by ultrasound
- Cumulative clinical pregnancy rate

Search methods for identification of studies

Three review authors (SA, PB, and JM) searched databases (from inception to 7 January 2019) for all published and unpublished RCTs of TLSs, without language restrictions and in consultation with the Cochrane Gynaecology and Fertility Group (CGFG) Information Specialist. We used both electronic searches of bibliographic databases and handsearching as described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011).

Electronic searches

We searched the following electronic databases, trial registers and websites.

- Cochrane Gynaecology and Fertility Group Specialised Register, ProCite platform (searched 7 January 2019) (Appendix 1)
- Cochrane Central Register of Controlled Studies (CENTRAL) (CRSO), web platform (searched 7 January 2019) (Appendix 2)
- MEDLINE In-Process & Other Non-Indexed Citations, Ovid platform (searched from 1946 to 7 January 2019 (Appendix 3)
- Embase, Ovid platform (searched from 1980 to 7 January 2019) (Appendix 4)
- Cumulative Index to Nursing and Allied Health Literature (CINAHL), EBSCO platform (searched from 1961 to 7 January 2019) (Appendix 5)

For MEDLINE, we used the Cochrane Highly Sensitive Search Strategy for identifying randomised controlled trials: sensitivity and precision maximising version (2008 revision), Ovid format (Higgins 2011).

We also searched the following other electronic sources of trials (web platforms, all searched 7 January 2019).

- Trial registers for ongoing and registered trials: World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) (www.apps.who.int/trialsearch/) (Appendix 6) and US National Institutes of Health Ongoing Trials Register ClinicalTrials.gov (www.clinicaltrials.gov)
- Web of Knowledge (wokinfo.com/)
- ProQuest Dissertations and Theses (search.proquest.com/)
- Grey literature through the System for Information on Grey Literature in Europe 'OpenGrey' (www.opengrey.eu/).

Searching other resources

We used the following methods to identify additional relevant RCTs:

- contact with authors of all RCTs identified by other methods;
- contact with manufacturers of TLSs;
- handsearching of selected journals in obstetrics, gynaecology and reproductive medicine, as well as conference proceedings (for abstracts) of the European Society for Human Reproduction and Embryology (ESHRE) and the American Society for Reproductive Medicine (ASRM);
- contacting known experts and personal contacts regarding unpublished materials;
- searching the citation lists of all identified articles for any relevant references.

Data collection and analysis

Selection of studies

We used the software program Covidence to manage the screening of titles and abstracts and to generate the PRISMA flow diagram (Covidence). All review authors took part in independently scanning the titles and abstracts of the articles retrieved by the search. Three review authors (SA, PB, and JM) then obtained the full texts of potentially eligible studies and independently examined these against the inclusion criteria for their eligibility. In the case of doubt between the review authors, a fourth review author (CF) was consulted to establish consensus on whether to include the trial or not. We documented the selection process with a PRISMA flow chart.

Data extraction and management

Three review authors (SA, PB, and JM) independently obtained and extracted data. Any disagreements between review authors were resolved by consulting a fourth review author (CF) to achieve consensus. We extracted data using a data extraction form designed and piloted by the review authors. If studies were reported in multiple publications, we extracted data from the different publications and then combined these into a single data extraction form so that no data were omitted. We included the following characteristics of included studies in the data extraction form:

- methods;
- participants;
- interventions;
- outcomes, including adverse events;
- funding source for studies.

Assessment of risk of bias in included studies

Three review authors (SA, PB, and JM) independently assessed the risk of bias in included studies using the Cochrane 'Risk of bias' assessment tool. We evaluated all included studies for the following: adequacy of sequence generation and allocation concealment; adequacy of blinding of couples, providers, and outcome assessors; completeness of outcome data; risk of selective outcome reporting; and risk of other potential sources of bias (Higgins 2011).

Any disagreements between authors were resolved by consulting a fourth review author (VJ) to achieve consensus. The results of the 'Risk of bias' assessment are presented in the 'Characteristics of included studies' table.

Measures of treatment effect

For dichotomous data (e.g. live birth or not), we calculated Mantel-Haenszel odds ratios (ORs) and 95% confidence intervals (CIs).

Unit of analysis issues

We analysed the data per couple randomised. We excluded studies randomising oocytes or embryos.

Dealing with missing data

If relevant data were missing from an included study, we contacted the original investigators of the trial to request the missing data. All original investigators were contacted. In particular, we obtained clinical pregnancy, live-birth, and stillbirth data from Park 2015; live-birth and stillbirth data from Yang 2018; miscarriage and clinical pregnancy data per woman randomised for Goodman 2016; live-birth and stillbirth data from Kahraman 2013; miscarriage data from Kaser 2017; and updated ongoing pregnancy and miscarriage data from Barberet 2018. If participants were described as 'lost to follow-up' without a specified reason, we assumed the participant did not experience the event or outcome (i.e. did not become pregnant).

Assessment of heterogeneity

We considered whether the clinical and methodological characteristics of the included studies were sufficiently similar for meta-analysis to provide a clinically meaningful summary. We assessed statistical heterogeneity by measuring the I^2 statistic. We assumed that there was substantial heterogeneity when I^2 was calculated as greater than 50% (Higgins 2011).

Assessment of reporting biases

In view of the difficulty of detecting and correcting for publication bias and other reporting biases, the we aimed to minimise their potential impact by ensuring a comprehensive search for eligible studies and by being alert to duplication of data. We assessed within-study reporting bias, which we judged as low risk if all of the study's prespecified primary outcomes were reported as outlined in the study's protocol.

Data synthesis

Where sufficient data were available, we combined the data for the primary outcomes by using a fixed-effect model in the following comparisons.

- TLS with conventional morphological assessment of still TLS images versus conventional incubation and assessment (trial design 1)
- TLS utilising embryo selection software versus TLS with conventional morphological assessment of still TLS images (trial design 2)
- TLS utilising embryo selection software versus conventional incubation and assessment (trial design 3)

Subgroup analysis and investigation of heterogeneity

Where sufficient data were available, we aimed to conduct the following subgroup analyses to determine the potential causes of heterogeneity for the live-birth and clinical pregnancy outcomes:

- donor oocytes (from donors of any age) versus autologous oocytes (from women of any age);
- fresh cycles (where embryos were replaced either at cleavage stage (day 3) or blastocyst (day 5)) versus frozen cycles (where frozen embryos were replaced in an ART cycle).

If we detected substantial heterogeneity, we planned to explore it by employing the random-effects model. We aimed to take any statistical heterogeneity into account when interpreting the results, especially if there was any variation in the direction of effect.

Sensitivity analysis

We planned to undertake sensitivity analyses for the review outcomes to determine whether the results were robust to decisions made during the review process. These analyses would have included consideration of whether the review conclusions would have differed if:

- the summary effect measure had been risk ratio rather than odds ratio;
- eligibility had been restricted to studies with low risk of bias for randomisation and allocation concealment;
- the primary outcome had been live birth only (i.e. not including ongoing pregnancy).

Overall quality of the body of evidence: 'Summary of findings' table

We prepared 'Summary of findings' tables using GRADEpro GDT and Cochrane methods (GRADEpro GDT 2015). These tables evaluate the overall quality of the body of evidence for the main review outcomes (live birth or ongoing pregnancy, miscarriage and stillbirth, and clinical pregnancy) for the review comparisons:

- TLS with conventional morphological assessment of still TLS images versus conventional incubation and assessment (trial design 1);
- TLS utilising embryo selection software versus TLS with conventional morphological assessment of still TLS images (trial design 2); and
- TLS utilising embryo selection software versus conventional incubation and assessment (trial design 3).

We assessed the quality of the evidence using GRADE criteria: risk of bias, consistency of effect, imprecision, indirectness, and publication bias. Two review authors (SA and PB) independently assessed the quality of the evidence as high, moderate, low, or very low, resolving any disagreements by discussion. Judgements were justified, documented, and incorporated into the reporting of results for each outcome.

RESULTS

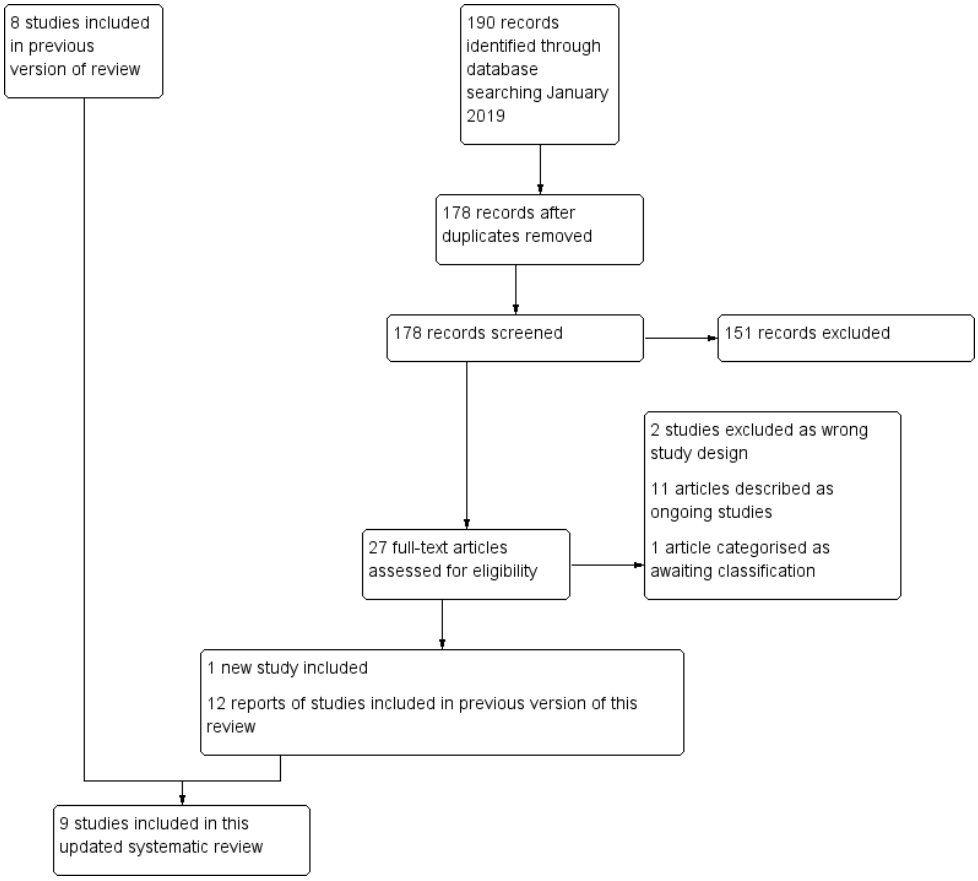
Description of studies

Results of the search

The most recent search took place in January 2019. We imported the 190 retrieved references into Covidence, and after removal of duplicates, all review authors screened 178 studies. We assessed 27 full-text articles for eligibility, of which one was a new RCT eligible for inclusion (Barberet 2018); two were excluded because they did not meet our inclusion criteria for study design (Alhelou 2018; Hardarson 2016); 11 were ongoing (ChiCTR1800017127; ChiCTR-IIR-16008758; ISRCTN17792989; NCT01760278; NCT02222831; NCT02417441; NCT02657811; NCT02852356; NCT02965222; NCT03164551; NCT03445923); and one is awaiting classification (Hulme 2014). The other 12 articles were conference abstracts from existing studies in the review, and have been listed under the main study references.

Taking into account the studies found in previous iterations of the review (described below), the review now has a total of nine included studies, 22 excluded studies, 13 ongoing studies and one study awaiting assessment (Figure 1, Included studies, Excluded studies, Studies awaiting classification; Ongoing studies).

Figure 1. Study flow diagram.



The first iteration of this review included three parallel-design RCTs from a search that retrieved 33 articles in total (Kahraman 2013; Kovacs 2019; Rubio 2014). Two further searches in 2016 and 2017 retrieved 82 and 293 articles, respectively. We retrieved a further four articles through handsearching. We screened the titles and abstracts of 266 articles after removal of duplicates. Of these 25 articles were potentially eligible for inclusion in the review, and we retrieved these in full text. Five new studies met our inclusion criteria (Goodman 2016; Kaser 2017; Park 2015; Wu 2016; Yang 2018). We excluded the remaining 20 studies for the following reasons: three studies were not RCTs; three were systematic reviews; two were letters; nine randomised embryos or oocytes; two were pseudo-randomised; and for one study we were unable to determine the nature of the control group despite our attempts to contact the authors.

Included studies

Study design and setting

We included nine RCTs in this review. The largest study was a multicentre RCT conducted in Spain, which was included in the first iteration of this review (Rubio 2014). The first iteration also included a single-centre RCT conducted in Turkey (Kahraman 2013), and a further multicentre RCT conducted in Hungary for which the completed results are now available (Kovacs 2019). The second iteration of the review added three single-centre studies conducted in the USA (Goodman 2016; Kaser 2017; Wu 2016), one single-centre study conducted in Sweden (Park 2015), and one single-centre study conducted in China (Yang 2018). This third iteration of the review includes completed study data from Yang 2018 and a completed single-centre RCT conducted in France (Barberet 2018)

Participants

The studies included 2955 infertile couples undergoing assisted reproductive technology (ART). Four studies included couples undergoing intracytoplasmic sperm injection (ICSI) alone (Barberet 2018; Kahraman 2013; Park 2015; Rubio 2014). One study included couples undergoing in vitro fertilisation (IVF) (Goodman 2016). The remaining studies included couples undergoing both IVF and ICSI (Kaser 2017; Kovacs 2019; Wu 2016; Yang 2018).

The largest study was Rubio 2014, with 856 participants; the second largest study had 600 participants (Yang 2018), followed by Barberet 2018 with 386 participants, and Park 2015 with 364 participants. The next-largest study had 300 participants (Goodman 2016), followed by Kaser 2017, with 163 participants. Kovacs 2019 had 161 participants, and the remaining two studies were relatively small, with 76 and 49 participants, respectively (Kahraman 2013; Wu 2016).

All studies utilised the autologous oocytes of the women randomised into their study, with the exception of Rubio 2014, which included couples undergoing ART with autologous or donor oocytes. The proportion of couples receiving donor oocytes in this study is unknown. Most donor oocytes in this study were used in fresh cycles, however some donor oocytes were obtained from an oocyte bank and were therefore vitrified.

All studies included women undergoing fresh embryo transfer, hence no cumulative cycle results were available. The majority of studies undertook single embryo transfer (Kahraman 2013; Kaser 2017; Kovacs 2019; Park 2015; Yang 2018). One study describes use of one or two embryos (Barberet 2018), and one study reports replacing between one and three embryos based on published American Society for Reproductive Medicine (ASRM) committee guidance and patient preferences (Goodman 2016). Another study undertook multiple embryo transfer (Rubio 2014), and a further study did not disclose the number of embryos transferred (Wu 2016).

The reported causes of infertility varied between studies. Some studies specifically described their participants as "good prognosis patients" (e.g. Rubio 2014 and Yang 2018). One study specifically described their participants as "poor prognosis patients", but provided no further information (Wu 2016). One study described "tubo-peritoneal factor" as the cause of infertility (Kahraman 2013), and another described male-factor infertility being present in more than 99% of participants in both arms, and female-factor infertility being present in approximately 20% of participants in both arms (Park 2015). Kovacs 2019 described various causes of infertility in participants ("male, tubal, unexplained etc."). One study described "a combination of anovulation, diminished ovarian reserve, endometriosis, male factor, tubal, unknown, and uterine" as causes of infertility (Kaser 2017). Barberet 2018 included male-factor, female-factor, mixed, and idiopathic indications. Goodman 2016 described a range of infertility diagnoses ("unexplained, ovulatory dysfunction, male factor, tubal factor, low ovarian reserve, AMA [advanced maternal age], endometriosis, mixed factors and other").

Interventions

We sought to divide studies into three comparisons depending on the nature of the intervention and the control, in order to truly assess if, and where, the benefit of a TLS lies.

TLS with conventional morphological assessment of still TLS images versus conventional incubation and assessment (trial design 1)

Four studies undertook this comparison (Barberet 2018; Kahraman 2013; Park 2015; Wu 2016). All studies utilised an integrated TLS, and all had two arms. Embryo transfer (ET) was undertaken at blastocyst in Kahraman 2013, day three in Wu 2016, day two in Park 2015, and day 2, day 3, or day 5-6 in Barberet 2018. Correspondence with the authors of one study confirmed that no embryo selection software was utilised in the intervention arm (Kahraman 2013). Embryos were left undisturbed in the TLS in the intervention arm in all three studies. In the control arm, embryos in all studies were assessed by conventional morphology using a benchtop microscope.

TLS utilising embryo selection software versus TLS with conventional morphological assessment of still TLS images (trial design 2)

Two studies undertook this comparison (Goodman 2016; Kaser 2017). One study utilised an integrated TLS (Goodman 2016), and the other utilised a TLS that was placed inside a conventional incubator (Kaser 2017). The embryos in the intervention arms were selected for transfer according to the information obtained from the embryo selection software, however the embryos of the women randomised to the intervention arm in one study were removed from the incubator for conventional benchtop morphology in addition to TLS selection (Kaser 2017). In addition, the embryos in the control arm of this study were assessed with conventional morphological assessment using a benchtop microscope. Time-lapse system images were not utilised for the selection of embryos for replacement in the control arm.

One study had three arms (Kaser 2017). There were two intervention arms: both were TLS utilising embryo selection software, but one arm undertook ET on day 3, and the other undertook ET on day 5. The control arm undertook ET on day 5. The other study had two arms, with ET undertaken on day 3 or day 5 (Goodman 2016).

We conducted in-depth discussions with the authors of Kaser 2017, and decided that trial design 2 was the most appropriate comparison, given that embryo selection software was utilised, and the trial design tested the embryo-selection element of the TLS software.

TLS utilising embryo selection software versus conventional incubation and assessment (trial design 3)

Three studies undertook this comparison (Kovacs 2019; Rubio 2014; Yang 2018). Two of these studies utilised a TLS that was placed inside a conventional incubator (Kovacs 2019; Yang 2018), whilst the third study utilised an integrated TLS (Rubio 2014). In Rubio 2014, ET was undertaken on days 3 and 5 in both arms; in Kovacs 2019, blastocyst transfer was undertaken in both arms. One study undertook ET on day 3 in the intervention arm and day 5 (blastocyst) in the control arm (Yang 2018). We took methodological advice on Yang 2018, and made the decision to keep the study in our review despite the differing days of ET. We gave this study a rating of high risk of bias due to this within-study imbalance.

Outcomes

All nine studies reported clinical pregnancy rates per couple. Miscarriage data were available for all included studies except for Wu 2016. Miscarriage data were confirmed to be loss of

a clinical pregnancy (not biochemical) in six studies (Barberet 2018; Kahraman 2013; Kaser 2017; Kovacs 2019; Park 2015; Yang 2018). In two studies the miscarriage data were a mixture of biochemical and clinical pregnancy losses (Goodman 2016; Rubio 2014). Unfortunately, the authors of these two studies were unable to provide only miscarriage data from clinical pregnancies. In these cases we have taken the pragmatic view to include these data, as according to the authors of these studies the majority of the pregnancy losses were from clinical pregnancies.

Either live birth or ongoing pregnancy was reported in all the studies except Goodman 2016 and Wu 2016. We obtained unpublished live-birth data for three studies following communication with the authors (Kahraman 2013; Park 2015; Yang 2018). For Rubio 2014, we obtained data from a related publication and conference abstract pertaining to the same study (Insua 2015; Insua 2017). We obtained stillbirth data from three studies following communication with the authors (Kahraman 2013; Park 2015; Yang 2018).

Excluded studies

We excluded 22 studies from the review because they did not meet our inclusion criteria for study design. For details see Characteristics of excluded studies.

Risk of bias in included studies

For details of the 'Risk of bias' assessments see Figure 2 and Figure 3.

Figure 2. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.

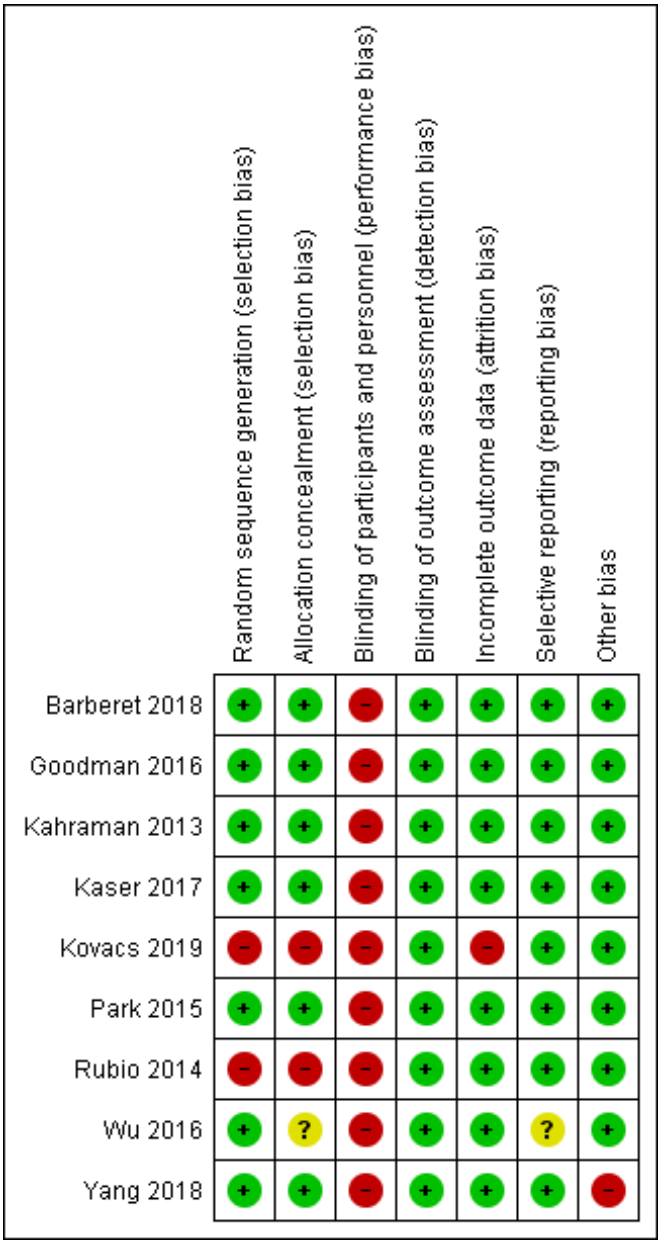
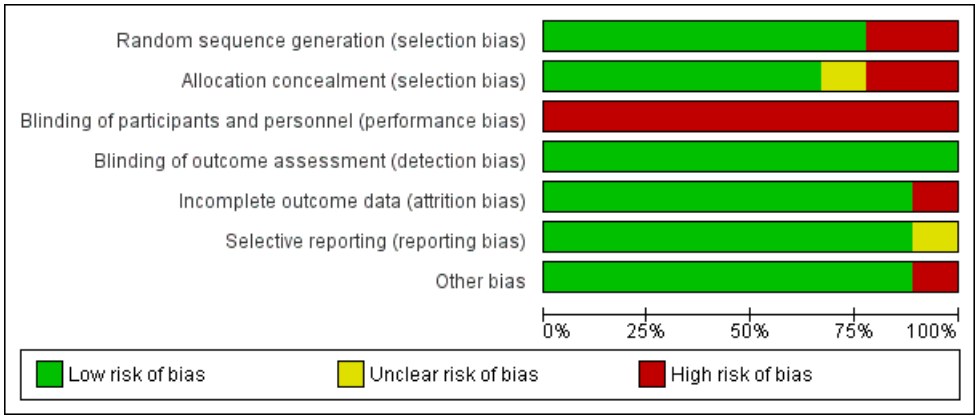


Figure 3. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.



Allocation

Sequence generation

Seven of the nine studies were at low risk of selection bias related to sequence generation. Six used a computer-generated randomisation list (Barberet 2018; Goodman 2016; Kahraman 2013; Kaser 2017; Park 2015; Wu 2016). One study utilised a random number table (Yang 2018).

We deemed two studies to have a high risk of bias for this domain (Rubio 2014; Kovacs 2019). In one study, although adequate random sequence generation was undertaken, some women were able to request the intervention, and in some cases this request was granted (Rubio 2014). The authors of this study assured us that this preferential allocation occurred in a minority of occasions and that the vast majority of participants were truly randomised, therefore we have maintained that this is an RCT. One study undertook paired randomisation whereby two envelopes containing time-lapse or control group assignments were prepared, and the first patient was randomly assigned to one of the groups and the next patient received the other assignment (Kovacs 2019). This was repeated with patient numbers three and four, and so on.

Allocation concealment

Six studies described methods of allocation concealment that resulted in a judgement of low risk of selection bias (Barberet 2018; Goodman 2016; Kahraman 2013; Kaser 2017; Park 2015; Yang 2018). In each of these studies, the randomisation list or numbered, opaque, sealed envelopes were held and administered by personnel not directly involved in the recruitment of participants, or else the allocation was conducted remotely (Barberet 2018).

We deemed two studies to be at high risk of bias for this domain (Kovacs 2019; Rubio 2014). In Kovacs 2019, randomisation was carried out by the principal investigator who was involved in the study. In Rubio 2014, it was reported that in some cases the allocation was non-random.

We judged one study for which there was limited description of randomisation to be at unclear risk of bias for this domain (Wu 2016). We understand that randomisation was undertaken by a member of the team not associated with the treatment cycle, and then subsequently the designation was reported to the embryology staff who processed the participant's oocytes/embryos. However, it was unclear how the randomisation list was stored, at what point the participants were randomised, and whether the person undertaking randomisation was responsible for recruitment.

Blinding

Blinding of participants and personnel (performance bias)

Three studies blinded their couples, and this blinding was not broken unless participants withdrew from the study (Goodman 2016; Kahraman 2013; Park 2015). Clinicians involved in the study were also blinded until after embryo transfer. One study described blinding the embryologist to the Eeva rating for the morphological assessment of embryos (Kaser 2017). The participants and physicians were all blinded to the TLS ratings. In addition, the sonographer was blinded in Goodman 2016, and the statistician was blinded in Park 2015.

Three studies did not blind or maintain blinding of their participating couples (Kovacs 2019; Rubio 2014; Yang 2018). In two of these studies, the clinical staff were also not blinded (Kovacs 2019; Yang 2018). The gynaecologist and statistician were blinded in Rubio 2014. We assessed these three studies as being at high risk of this bias.

Barberet 2018 did not discuss performance bias in detail or report who was blinded, but noted that it was not possible to blind investigators to the allocations. However, in this study embryos were selected for vitrification according to their morphology, which was graded in unblinded embryo assessments.

We deemed one study as having a high risk of performance bias as blinding was not described, and it would have been impossible to blind the embryologist (Wu 2016). We have been unable to contact the authors for further clarification.

None of the included studies blinded the embryologists, but this would have been impossible. We considered a lack of blinding of embryologists as reason for a judgement of high risk of performance bias. This renders all included studies as having a high risk of performance bias. In some studies, the lack of blinding may have influenced the number or day of transfer. In addition, it is impossible to remove the risk of performance bias when the person selecting the embryo for transfer is unblinded.

Blinding of outcome assessors (detection bias)

We judged all nine studies to be at low risk of detection bias because the outcomes (live birth or ongoing pregnancy, clinical pregnancy, miscarriage and stillbirth) are objective, and therefore cannot be influenced by knowledge of the intervention. Two studies described how staff performing the ultrasounds were blinded to the intervention (Goodman 2016; Rubio 2014). The remaining studies did not blind their outcome assessors, however we still deemed these studies as having a low risk of bias due to the reason described above.

Incomplete outcome data

We deemed the following studies to be at low risk of attrition bias:

- Barberet 2018, because outcomes were reported for all participants, using intention-to-treat analysis;
- Goodman 2016, because we were able to obtain the outcome data from the five women excluded after randomisation;
- Kahraman 2013, because the 12 couples who dropped out after randomisation were accounted for, and the reasons were clearly stated;
- Kaser 2017, because all data were presented in their paper as intention-to-treat;
- Park 2015, because there was only one woman excluded from analysis due to having been accidentally randomised twice;
- Wu 2016, because the small number of excluded participants were accounted for according to predetermined grounds for exclusion;
- Rubio 2014, because the 13 couples who were excluded following randomisation were accounted for and were a very small proportion of the total number of couples randomised; and
- Yang 2018, because the 15 couples who were excluded following randomisation were accounted for with clearly stated reasons for exclusion that were predetermined.

We judged one study to be at high risk of attrition bias because a large proportion of the couples recruited were excluded from the trial (22 out of 161 couples randomised) (Kovacs 2019). The reasons for dropout were provided, however not all of the reasons were specified in the predetermined exclusion criteria, and given the high attrition rate, we assessed this study at high risk of attrition bias.

We undertook an intention-to-treat analysis on all dichotomous outcomes, using data from those women excluded postrandomisation where possible.

Selective reporting

We considered eight studies to be at low risk of reporting bias because they reported and published all outcomes they had set out to investigate (Barberet 2018; Goodman 2016; Kahraman 2013; Kaser 2017; Kovacs 2019; Park 2015; Rubio 2014; Yang 2018). This was confirmed on communication with authors and by referencing against information in online trials registers if it was available.

We considered one study to be at unclear risk of reporting bias because access to their protocol was not available and we could not contact the authors to ask whether they had published all prespecified outcomes (Wu 2016).

Other potential sources of bias

We found no potential sources of within-study bias in Barberet 2018, Goodman 2016, Kahraman 2013, Kaser 2017, Kovacs 2019, Park 2015, Rubio 2014, and Wu 2016. We assessed these studies as having a low risk of this form of bias.

We assessed one study, Yang 2018, as having a high risk of within-study bias. This was due to the difference in day of embryo transfer between study arms (day 3 for intervention and day 5 for control). This difference in maturity of the embryo could have had an impact on the likelihood of an ongoing pregnancy.

Effects of interventions

See: **Summary of findings for the main comparison** TLS with conventional morphological assessment of still TLS images compared to conventional incubation and assessment for embryo incubation and assessment in assisted reproduction; **Summary of findings 2** TLS utilising embryo selection software compared to TLS with conventional morphological assessment of still TLS images for embryo incubation and assessment in assisted reproduction; **Summary of findings 3** TLS utilising embryo selection software compared to conventional incubation and assessment for embryo incubation and assessment in assisted reproduction

1. TLS with conventional morphological assessment of still TLS images versus conventional incubation and assessment (trial design 1)

Four studies undertook this comparison (Barberet 2018; Kahraman 2013; Park 2015; Wu 2016), with a total of 875 participants.

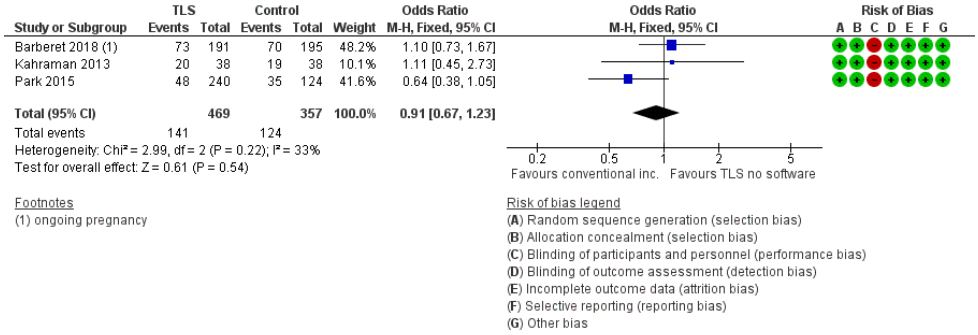
Primary outcomes

1.1 Live birth or ongoing pregnancy

Two studies provided live-birth data following correspondence with their authors (Kahraman 2013; Park 2015; N = 440), and one study provided data on ongoing pregnancy (Barberet 2018; N = 386). There were 141 events reported among the 469 women randomised to the TLS arm, and 124 events among the 357 women randomised to the control arm (conventional incubation and embryo assessment).

It is unclear whether there is any difference between interventions in rates of live birth or ongoing pregnancy (odds ratio (OR) 0.91, 95% confidence interval (CI) 0.67 to 1.23, 3 RCTs, N = 826, I^2 = 33%, low-quality evidence, Analysis 1.1, Figure 4). The evidence suggests that if the rate of live birth or ongoing pregnancy associated with conventional incubation and assessment is 35%, the rate with the use of TLS with conventional morphological assessment of still TLS images would be between 27% and 40%.

Figure 4. Forest plot of comparison: 1 TLS with conventional morphological assessment of still TLS images versus conventional incubation and assessment (trial design 1), outcome: 1.1 Live birth or ongoing pregnancy.



A sensitivity analysis restricting the analysis to studies reporting live birth did not influence this finding substantially.

1.2 - 1.3 Miscarriage and stillbirth

Three studies provided data on miscarriage (Barberet 2018; Kahraman 2013; Park 2015; N = 826), and two studies also provided data on stillbirth (Kahraman 2013; Park 2015; N = 440). The data on stillbirth were made available following communication with the authors of Park 2015.

Out of 469 women randomised to the intervention arm, 33 experienced a miscarriage, whereas out of 357 randomised to the control arm, 15 experienced a miscarriage. It is unclear whether there is any difference between interventions in rates of miscarriage (OR 1.90, 95% CI 0.99 to 3.61, 3 RCTs, N = 826; I^2 = 0%, low-quality evidence, Analysis 1.2). The evidence suggests that if the miscarriage rate with conventional incubation is 4%, the rate associated with the use of TLS with conventional morphological assessment of still TLS images would be between 4% and 14%.

Regarding stillbirth, there were 2 stillbirths out of 38 women randomised to the intervention arm, and 2 stillbirths out of 38 women randomised to the control arm in Kahraman 2013. There were no stillbirths recorded in either arm in Park 2015, meaning that a result is inestimable. In accordance with Cochrane methodological guidance, we have removed Park 2015 from meta-analysis. Results from the single study, Kahraman 2013, suggest that it is unclear whether there is any difference between interventions in rates of stillbirth (OR 1.00, 95% CI 0.13 to 7.49, 1 RCT, N = 76, low-quality evidence, Analysis 1.3).

Cumulative live birth or ongoing pregnancy

No data were provided for this outcome.

Secondary outcomes

1.4 Clinical pregnancy

All four studies provided clinical pregnancy data (Barberet 2018; Kahraman 2013; Park 2015; Wu 2016; N = 875). There were 178 clinical pregnancies among the 493 women randomised to the intervention arm, and 143 clinical pregnancies among the 382 women randomised to the control arm.

It is unclear whether there is any difference between interventions in rates of clinical pregnancy (OR 1.06, 95% CI 0.79 to 1.41, 4 RCTs, N = 875, I^2 = 0%, low-quality evidence, Analysis 1.4).

Cumulative clinical pregnancy

No data were provided for this outcome.

2. TLS utilising embryo selection software versus TLS with conventional morphological assessment of still TLS images (trial design 2)

Two studies undertook this comparison (Goodman 2016; Kaser 2017), with a total of 463 participants. It is worth noting that in Kaser 2017 there were two intervention groups: one involved day 3 embryo transfer and the other day 5 embryo transfer. The two intervention groups are represented as separate entities at meta-analysis, and the single control group has been split to share between the two intervention groups in order to avoid artificially doubling the effect of the control group.

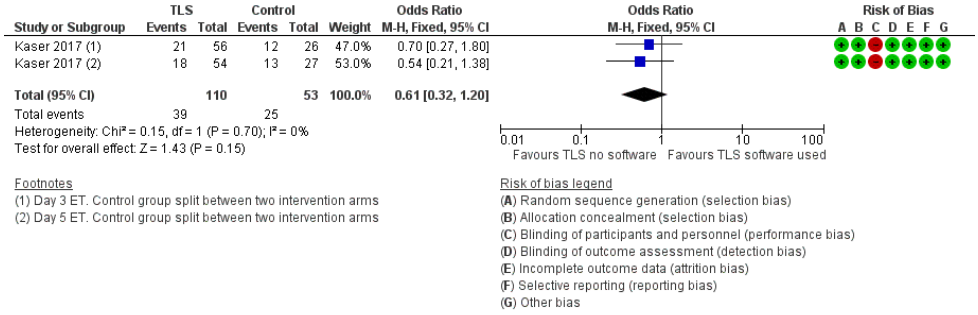
Primary outcomes

2.1 Live birth or ongoing pregnancy

Neither study collected live-birth data. This was confirmed on correspondence with the authors of both studies. One RCT reported ongoing pregnancy (Kaser 2017).

There were 39 ongoing pregnancies among the 110 women randomised to the intervention arm, and 25 ongoing pregnancies among the 53 women randomised to the control arm. It is unclear whether there is any difference between interventions for this outcome (OR 0.61, 95% CI 0.32 to 1.20, 1 RCT, N = 163, very low-quality evidence, Analysis 2.1, Figure 5). The evidence suggests that if the rate of ongoing pregnancy associated with TLS with conventional morphological assessment of still TLS images is 47%, the rate associated with TLS utilising embryo selection software would be between 22% and 52%.

Figure 5. Forest plot of comparison: 2 TLS utilising embryo selection software versus TLS with conventional morphological assessment of still TLS images (trial design 2), outcome: 2.1 Live birth or ongoing pregnancy.



2.2 Miscarriage and stillbirth

Neither study collected data on stillbirth.

We obtained miscarriage data for all randomised women following correspondence with the authors of both studies. For Goodman 2016, the miscarriage data include a combination of biochemical and clinical pregnancy losses. Unfortunately, these data could not be separated for our review. For Kaser 2017, the data include miscarriages from clinical pregnancy losses.

There were 18 miscarriages out of 260 women randomised to the intervention arm, and 11 miscarriages out of 203 women randomised to the control arm. We are uncertain whether TLS utilising embryo selection software influences miscarriage rates (OR 1.39, 95% CI 0.64 to 3.01, 2 RCTs, N = 463, I^2 = 0%, very low-quality evidence, Analysis 2.2). The evidence suggests that if the miscarriage rate associated with assessment of still TLS images is 5%, the rate with embryo selection software would be between 4% and 14%.

Cumulative live birth or ongoing pregnancy

No data were provided for this outcome.

Secondary outcomes

2.3 Clinical pregnancy

Both studies reported this outcome. There were 132 clinical pregnancies out of the 260 women randomised to the intervention group, and 109 pregnancies out of the 203 women randomised to the control group. It is unclear whether there is any difference between interventions in clinical pregnancy rates (OR 0.97, 95% CI 0.67 to 1.42, 2 RCTs, N = 463, I^2 = 0%, very low-quality evidence, Analysis 2.3).

Cumulative clinical pregnancy

No data were provided for this outcome.

3. TLS utilising embryo selection software versus conventional incubation and assessment (trial design 3)

Three studies undertook this comparison (Kovacs 2019; Rubio 2014; Yang 2018), with a total of 1351 participants. There were marked methodological differences between two of these studies, Kovacs 2019; Rubio 2014, and the third study, Yang 2018, with respect to study design as well as internal validity. In contrast to the other two studies, Yang 2018 had differing days of embryo transfer in the intervention and the control arms of the study. Moreover, Yang 2018 was at low risk of selection bias, whereas the other two studies were at high risk of selection bias relating to both sequence generation and allocation concealment. As noted below, there was high heterogeneity when these three studies were combined, which may be attributable to differences in design, differences in risk of bias, or both.

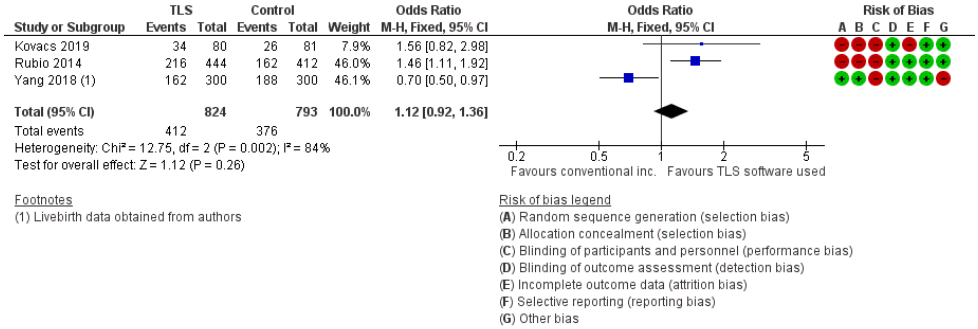
Primary outcomes

3.1 Live birth or ongoing pregnancy

Live-birth data were available for all three studies (Kovacs 2019; Rubio 2014; Yang 2018). For Rubio 2014, we obtained data from a recently published paper and a published conference abstract (the references for these are provided as subreferences under Rubio 2014). Yang 2018 (N = 600) provided data on live birth following email communication. As noted above, the study design of Yang 2018 was very different from that of the other two studies in this comparison owing to the fact that it has differing days of embryo transfer in the intervention and the control arms of the study.

There were 412 events among the 824 women randomised to the intervention arm, and 376 events among the 793 women randomised to the control arm. It is unclear whether there is any difference between interventions in rates of live birth (OR 1.12, 95% CI 0.92 to 1.36, 3 RCTs, N = 1617, I^2 = 84%, very low-quality evidence, Analysis 3.1, Figure 6). There was high statistical heterogeneity for this finding, possibly due to the above mentioned differing study designs. The evidence suggests that if the rate of live birth or ongoing pregnancy associated with conventional incubation is 48%, the rate with TLS utilising embryo selection software would be between 46% and 55%.

Figure 6. Forest plot of comparison: 3 TLS utilising embryo selection software versus conventional incubation and assessment (trial design 3), outcome: 3.1 Live birth or ongoing pregnancy.



3.2 - 3.3 Miscarriage and stillbirth

Two studies defined miscarriage data as loss of clinical pregnancies (Kovacs 2019; Yang 2018). The other study reported a combination of biochemical and clinical pregnancy losses (Rubio 2014). Stillbirth data were made available following email correspondence with Yang 2018. There were no stillbirths in either arm of this study.

There were 60 miscarriages among 824 women randomised to the intervention arm, and 86 miscarriages among 793 women randomised to the control arm. The evidence suggests that TLS utilising embryo selection software may reduce miscarriage rates, but this finding is very uncertain as the evidence is of very low quality (OR 0.63, 95% CI 0.45 to 0.89, 3 RCTs, N = 1617, I^2 = 0%, Analysis 3.2). The evidence suggests that if the miscarriage rate with conventional incubation is 11%, the rate associated with TLS would be between 5% and 10%.

Cumulative live birth or ongoing pregnancy

No data were provided for this outcome.

Secondary outcomes

3.4 Clinical pregnancy

Three studies reported this outcome (Kovacs 2019; Rubio 2014; Yang 2018; N = 1617). There were 489 clinical pregnancies among 824 women randomised to the intervention arm, and 480 clinical pregnancies among 793 women randomised to the control arm. It is unclear whether there is any difference between interventions for this outcome (OR 0.95, 95% CI 0.78 to 1.16, 3 RCTs, N = 1617, I^2 = 89%, Analysis 3.4). This finding is very uncertain due to the high risk of bias in the included studies and the high level of heterogeneity in study design.

Cumulative clinical pregnancy

No data were provided for this outcome.

Subgroup and sensitivity analysis

We did not perform any other planned subgroup or sensitivity analyses as there were insufficient included studies for any specific comparison.

DISCUSSION

Summary of main results

Trial design 1

The comparison 'TLS with conventional morphological assessment of still TLS images versus conventional incubation and assessment' aims to assess the potential advantages of a stable incubator environment. The embryo selection software is not utilised, and the embryos are left undisturbed until transfer. The four relevant studies included participants with a variety of infertility diagnoses. One study described its participants as "poor prognosis", with no further details (Wu 2016). Another study described women with "tubo-peritoneal factor" (Kahraman 2013), and the third study described over 99% male-factor infertility, with 20% female-factor in both arms (Park 2015). One study included women with a variety of diagnoses (Barberet 2018). This variety adds to the broad applicability of results to common clinical practice. Two studies undertook embryo transfer at day 2 or 3 (Park 2015; Wu 2016), whereas one study undertook blastocyst transfer (Kahraman 2013), and the fourth study undertook embryo transfer on a variety of days from day 2 to blastocyst (Barberet 2018). All oocytes were autologous.

The evidence is of low quality, and it is unclear whether there is any difference between interventions in rates of live birth or ongoing pregnancy, miscarriage and stillbirth, or clinical pregnancy.

Trial design 2

The comparison 'TLS utilising embryo selection software versus TLS with conventional morphological assessment of still TLS images' aims to assess the potential advantages of the embryo selection software over conventional morphology. In this comparison, both arms of the study are housed in a TLS, but the embryo selection software is utilised in only one arm. The incubator environment is therefore identical in both arms. Two studies were eligible for this comparison. One study had two intervention arms: embryo transfer on day 3 and embryo transfer on day 5 (Kaser 2017). The control arm had embryo transfer on day 5 only. The other study, Goodman 2016, undertook a combination of embryo transfer on day 3 or 5. It is worth noting that the embryos were left undisturbed in Goodman 2016, however in Kaser 2017, the embryos

in both intervention arms and in the control arm underwent daily conventional morphological assessment, in addition to the application of embryo selection software in the intervention arms. There was a broad variety of infertility diagnoses in both studies, which adds to the overall applicability of results to broad clinical practice.

All findings for this comparison were very uncertain due to the very low quality of the evidence. No data were available on live birth, but one study reported ongoing pregnancy: it is uncertain whether there is any difference between interventions in rates of ongoing pregnancy, miscarriage, or clinical pregnancy. No evidence for stillbirth was available.

Trial design 3

The comparison 'TLS utilising embryo selection software versus conventional incubation and assessment' aims to assess the potential advantages of a combination of the stable incubator environment*and* the embryo selection software versus conventional incubation and assessment. Three studies undertook this comparison. One of these studies utilised a combination of autologous and donor oocytes; the proportion of each is unknown (Rubio 2014). The remaining two studies used autologous oocytes. One study undertook embryo transfer on day 3 in the intervention group and day 5 in the control group (Yang 2018). Another study undertook transfer on day 5 (Kovacs 2019), and in the third study there was a combination of transfer on day 3 and day 5 (Rubio 2014). A variety of infertility diagnoses were recorded in the women in these studies. Two studies described their participants as "good prognosis" (Rubio 2014; Yang 2018).

All findings for this comparison were very uncertain due to the very low quality of the evidence. It is unclear whether there is any difference between interventions in live-birth rates. It is suggested that TLS might reduce miscarriage rates, but it is unclear whether there is any difference between interventions in clinical pregnancy rates. One study examined stillbirth, but as there were no events in either arm, it was not possible to reach any conclusions regarding this outcome.

Overall completeness and applicability of evidence

This updated systematic review on time-lapse systems now includes nine RCTs. Data from 2955 women have gone towards formulating the findings of this review, but unfortunately all the evidence is of low or very low quality.

Approximately 50% of participants were included in trials that assessed TLS utilising embryo selection software versus conventional incubation and assessment (trial design 3). This is mainly due to the largest included trial undertaking this comparison (Rubio 2014). Trial designs 1 and 2 (TLS with conventional morphological assessment of still TLS images versus conventional incubation and assessment, and TLS utilising embryo selection software versus TLS with conventional morphological assessment of still TLS images) include the remaining 33% and 17% of participants, respectively, but there were no women available to inform live-birth findings in trial design 2, meaning there are profound gaps in evidence for TLS in this comparison. In addition, there were no stillbirth data for trial design 2. This may be because stillbirth is so rare that it is not considered to be an important outcome, but it is important that future trials report this outcome, as it is a measure of safety.

Trial designs 1 and 2 included 875 and 463 women, respectively, in comparison to the 1617 women included in trial design 3. Despite the additional information from previous and newly incorporated trials, the results of the review remain unclear. Further trials of each design are required to bolster participant numbers and to interrogate the robustness of the finding of insufficient evidence of differences in live-birth, miscarriage, clinical pregnancy, and stillbirth rates to choose between TLS with or without embryo selection software versus conventional incubation and assessment. The largest trial that informs trial design 3 has a number of biases arising from the non-randomised approach for some participants, the subsequent lack of blinding, the use of donor oocytes in a number of women, and the routine use of multiple embryo transfer.

There was heterogeneity between trials in the diagnosis of infertility, the day of embryo transfer, the use of IVF or ICSI, and the make and model of TLS. All of these factors help to make the results of this review more applicable to clinical practice in the real world, where there is naturally this variation in clinical practices.

All included studies excluded women who underwent frozen embryo transfer, except Kahraman 2013, whose investigators were able to provide data for these women. The investigators of Rubio 2014 were unable to provide data specifically for women who underwent donor oocyte IVF/ICSI. Consequently, in order to subgroup autologous, donor, and frozen oocytes, future studies will need to present their results under these subgroups and state explicitly how many couples underwent these interventions.

Most studies undertook elective single embryo transfer (Kahraman 2013; Kaser 2017; Kovacs 2019; Park 2015; Yang 2018). However, three studies undertook multiple embryo transfers (Barberet 2018; Goodman 2016; Rubio 2014). We were unable to obtain from the authors of Rubio 2014 the exact proportion of couples who received multiple embryo transfer in each arm of the study. Given that this study contributed a large proportion of the data in trial design 3, it is important to recognise that the results presented here may reflect rates of clinical outcomes in keeping with multiple embryo transfer as opposed to single embryo transfer. One study did not disclose the number of embryos transferred per woman (Wu 2016).

Quality of the evidence

The quality of the evidence ranged from very low to low. The main limitations were high risk of bias in the included studies, imprecision, indirectness, and inconsistency. Risk of bias was most commonly associated with performance bias (lack of blinding of participants or those involved in the study) and selection bias (failure to use reliable methods of sequence generation and allocation concealment).

Inconsistency is evident across the comparisons. In particular, the point estimates of meta-analyses of comparison 3 suggest some benefit from TLS in its entirety compared to conventional incubation and assessment, whereas most of the point estimates from comparisons 1 and 2 suggest a reduction in benefit from using TLS without the embryo selection software compared to control. This finding is difficult to explain scientifically given the difference in direction of results in comparisons that assess the stable incubator environment of the TLS, and ability of embryo selection software to help select the best embryo. Despite differences

between the interventions, we would anticipate a consistent direction of effect across the three comparisons.

The inconsistency in the totality of the evidence relates to two studies in comparison 3 that found a benefit for TLS (Kovacs 2019; Rubio 2014). Both these studies were at high risk of selection bias (relating to both sequence generation and allocation concealment), which reduces our confidence in their findings. We rated the evidence for comparison 3 as very low (lower than for comparisons 1 and 2), denoting very little confidence in the effect estimate. With respect to inconsistency within comparison 3, there are two plausible explanations for the high statistical heterogeneity: in contrast to the other two studies, Yang 2018 had differing days of embryo transfer in the intervention and the control arms of the study. Moreover, Yang 2018 was at low risk of selection bias, whereas (as noted above) the other two studies were at high risk of selection bias.

The quality of the evidence for trial design 1 (TLS with conventional morphological assessment of still TLS images versus conventional incubation and assessment) is low, the main limitations being performance bias and imprecision (Summary of findings for the main comparison).

The quality of the evidence for trial design 2 (TLS utilising embryo selection software versus TLS with conventional morphological assessment of still TLS images) is very low, the main limitations being performance bias, indirectness, and imprecision (Summary of findings 2).

The quality of the evidence for trial design 3 (TLS utilising embryo selection software versus conventional incubation and assessment) is also very low, the main limitations being performance bias, selection bias, indirectness, and inconsistency (Summary of findings 3).

Potential biases in the review process

We aimed to identify all eligible studies for inclusion in this review, and contacted authors of the included studies on many occasions in an effort to include as much information as possible. The authors of most studies were forthcoming with further study information, which helped us to accrue a full picture of the study outcomes, as well as providing information needed to assess and establish risk of bias.

Agreements and disagreements with other studies or reviews

There are four published systematic reviews to date using the same inclusion and exclusion criteria on the topic of TLS versus conventional incubation (Chen 2017; Kaser 2014; Polanski 2014; Pribenszky 2017). Two of these are now out of date, with new studies published since their reporting (Kaser 2014; Polanski 2014). Both reviews reported no evidence of a difference between TLS and control.

One systematic review, Kaser 2014, included 13 eligible studies after systematic searching, however the majority of these were retrospective cohort studies, and none of them were RCTs. Kaser 2014 concluded that there is currently limited evidence to support the routine clinical use of TLS for selection of human pre-implantation embryos.

Chen 2017 included six eligible studies, but it missed two further eligible RCTs that are included in this review. Chen 2017 does not include all the potential live-birth data, including data from Kahraman 2013, Kovacs 2019, and Park 2015. It concludes that there is currently "insufficient evidence to support that time-lapse imaging is superior to conventional methods for embryo incubation and selection".

The authors of Pribenszky 2017 undertook a systematic review of TLS utilising TLS embryo selection software. They concluded that TLS using embryo selection software was associated with a significantly higher ongoing pregnancy rate, a significantly lower early pregnancy loss, and a significantly higher live-birth rate in comparison to control. However, we have detected a number of problems with this review that have been published as a letter (Armstrong 2018). The issues outlined are as follows.

- They have combined trials with different intervention and control arms. For example, three of the five included trials are study design 3, but one is study design 1 and one is study design 2.
- They have also included a trial that describes itself as a prospective cohort study, not an RCT. On closer investigation, this trial is pseudo-randomised (randomisation based on patient record number). This is not considered methodologically sound for systematic reviews of RCTs.
- The authors describe applying an intention-to-treat analysis (which is considered the gold standard in fertility research), however the early pregnancy loss, live-birth, and stillbirth data are analysed per woman that became pregnant. This is known to skew the results toward showing a larger intervention effect.
- It appears that full data from the included trials have not been entered into the review. For example, live-birth data are not included from Rubio 2014, despite being published as an abstract in 2015.
- We note that all three authors declared in this review that they work for Vitrolife, a biotechnology company that manufactures and promotes TLS.

AUTHORS' CONCLUSIONS

Implications for practice

Overall, there is insufficient good-quality evidence of differences in rates of live birth or ongoing pregnancy, miscarriage and stillbirth, or clinical pregnancy to choose between time-lapse systems (TLS), with or without embryo selection software, and control.

Women need to be aware, especially in view of the cost of TLS, that there is no good evidence that TLS with or without embryo selection software is more effective than conventional methods of embryo incubation and assessment. They may wish to take part in randomised controlled trials (RCTs) on TLS so as to add to the existing evidence base, and help guide assisted reproductive technology patients of the future.

Implications for research

Randomised controlled trials that randomise couples or women, not embryos or oocytes, to either TLS or conventional incubation should be designed and conducted to add to the currently limited RCT evidence. These studies should be large enough to answer the clinical questions that are important in fertility research, such

as live birth, clinical and ongoing pregnancy, and adverse events. Cumulative clinical pregnancy rates should be reported in future studies in order to determine the impact of a TLS on embryo selection.

Suggested designs of RCTs which seek to differentiate the unique advantages of TLS are as follows.

- Trial design 1) TLS utilising routine morphological assessment of TLS images versus conventional incubation and assessment
- Trial design 2a) TLS utilising embryo selection software versus TLS utilising routine morphological assessment of TLS images
- Trial design 2b) TLS utilising one type of embryo selection software versus TLS utilising a different type of embryo selection software
- Trial design 3) TLS utilising embryo selection software versus conventional incubation and assessment

These study designs will help to differentiate between: the potential advantages of the stable culture environment TLS provides (trial design 1); the potential advantage of embryo selection software (trial design 2); and the potential advantage of TLS in its entirety utilising embryo selection software versus conventional incubation and assessment (trial design 3).

In addition, it would be useful for future trials to include a cost analysis element, which may help patients to balance the costs and benefits of using this technology. It may also be helpful to explore patient satisfaction and quality of life with TLS versus with control. Some clinics are sharing TLS images with patients during the incubation period. It would be useful to explore whether this helps or worsens treatment anxiety.

ACKNOWLEDGEMENTS

Our heartfelt thanks go to:

- Marian Showell for her assistance with search strategies and re-running searches;
- Helen Nagels for her editorial advice and assistance;
- the editorial review team, who dedicated their valued time and effort to this review;
- the women/couples who joined the randomised controlled trials included here, whose altruism has helped to guide evidence;
- Ms Nicola Arroll and Dr Lynsey Cree, who contributed to the first iteration of this review;
- Roger Hart, Andy Vail, and Jeanette MacKenzie for their invaluable peer review.

REFERENCES

References to studies included in this review

Barberet 2018 *{published data only}*

Barberet J, Chammas J, Bruno C, Valot E, Vuillemin C, Jonval L, et al. Randomized controlled trial comparing embryo culture in two incubator systems: G185 K-System versus EmbryoScope. *Fertility and Sterility* 2018;**109**(2):302-9.

Goodman 2016 *{published and unpublished data}*

* Goodman LR, Goldberg J, Falcone T, Austin C, Desai N. Does the addition of time-lapse morphokinetics in the selection of embryos for transfer improve pregnancy rates? A randomized controlled trial. *Fertility and Sterility* 2016;**105**(2):275-85.

Goodman LR, Goldberg JM, Falcone T, Austin C, Desai N. Does use of time-lapse microscopy in the selection of embryos for transfer improve pregnancy rates? A randomized controlled trial. *Fertility and Sterility* 2015;**104**(3):e96.

Kahraman 2013 *{published and unpublished data}*

Kahraman S, Cetinkaya M, Pirkevi C, Yelke H, Kumtepe Y. Comparison of blastocyst development and cycle outcome in patients with eSET using either conventional or time lapse incubators. A prospective study of good prognosis patients. *Journal of Reproductive and Stem Cell Biotechnology* 2013;**3**(2):55-61.

Kaser 2017 *{published data only}*

* Kaser DJ, Bormann CL, Missmer SA, Farland LV, Ginsburg ES, Racowsky C. A pilot randomized controlled trial of Day 3 single embryo transfer with adjunctive time-lapse selection versus Day 5 single embryo transfer with or without adjunctive time-lapse selection. *Human Reproduction* 2017;**32**(8):1598-603.

Kaser DJ, Bormann CL, Missmer SA, Farland LV, Ginsburg ES, Racowsky C. EEVA pregnancy pilot study: a randomized controlled trial of single embryo transfer (SET) on day 3 or day 5 with or without time-lapse imaging (TLI) selection. *Fertility and Sterility* 2016;**106**(3):e312.

Kovacs 2019 *{published and unpublished data}*

* Kovacs P, Matyas S, Forgacs V, Sajgo A, Molnar L, Pribenszky C. Non-invasive embryo evaluation and selection using time-lapse monitoring: results of a randomized controlled study. *European Journal of Obstetrics & Gynecology and Reproductive Biology* 2019;**233**:58-63.

Kovacs P, Matyas S, Forgacs V, Sajgo A, Rarosi F, Pribenszky C. Time-lapse embryo selection for single blastocyst transfer - results of a multicentre, prospective, randomized clinical trial. *Fertility and Sterility* 2013;**100**(3):S90.

Kovacs P, Matyas SZ, Forgacs V, Reichart A, Rarosi F, Bernard A, et al. Can a composite score based on time lapse observation aid embryo selection for single embryo transfer; an interim report. *Human Reproduction* 2013;**28**(Suppl 1):169.

Matyas SZ, Kovacs P, Forgacs V, Sajgo A, Pribenszky CS. Selection of single blastocyst for transfer using time-lapse monitoring during in vitro fertilization in good prognosis patients: a randomized controlled trial. *Human Reproduction* 2015;**30**(Suppl 1):i119.

Park 2015 *{published data only}*

* Park H, Bergh C, Selleskog U, Thurin-Kjellberg A, Lundin K. No benefit of culturing embryos in a closed system compared with a conventional incubator in terms of number of good quality embryos: results from an RCT. *Human Reproduction* 2015;**30**(2):268-75.

Selleskog U, Park H, Bergh C, Lundin K. A prospective randomised controlled trial of the efficacy of using a closed time-lapse system for embryo culture. *Human Reproduction* 2014;**29**(Suppl 1):i61.

Rubio 2014 *{published and unpublished data}*

Ayerdi F, Rubio I, Galan A, Larreategui Z, Vidal C, Meseguer M. Clinical validation of embryo culture and selection by morphokinetic analysis; a randomized controlled trial by time-lapse system. *Human Fertility* 2014;**17**(4):298.

Insua MF, Cobo A, Larreategui Z, Ferrando M, Remohi J, Meseguer M. Obstetric and perinatal outcomes of singleton newborns using time lapse monitoring. *Fertility and Sterility* 2015;**104**(Suppl 3):e212-3.

Insua MF, Cobo AC, Larreategui Z, Ferrando M, Serra V, Meseguer M. Obstetric and perinatal outcomes of pregnancies conceived with embryos cultured in a time-lapse monitoring system. *Fertility and Sterility* 2017;**108**(3):498-504.

Perez S, Rubio I, Aparicio B, Beltran D, Garcia-Laez V, Meseguer M. Prospective validation of a time-lapse based algorithm for embryo selection. *Fertility and Sterility* 2014;**102**(Suppl 3):e322.

* Rubio I, Galan A, Larreategui Z, Ayerdi F, Bellver J, Herrero J, et al. Clinical validation of embryo culture and selection by morphokinetic analysis: a randomized, controlled trial of the EmbryoScope. *Fertility and Sterility* 2014;**10**(14):738.

de los Santos JM, Rubio I, Larreategui Z, Ayerdi F, Remohi J, Meseguer M. Clinical validation of embryo culture and selection by morphokinetic analysis; a randomized controlled trial by time-lapse imaging. *Fertility and Sterility* 2014;**102**(Suppl 3):e87.

Wu 2016 *{published data only}*

Wu Y-G, Lazzaroni-Tealdi E, Wang Q, Zhang L, Barad DH, Kushnir VA, et al. Different effectiveness of closed embryo culture system with time-lapse imaging (Embryoscope TM) in comparison to standard manual embryology in good and poor prognosis patients: a prospectively randomized pilot study. *Reproductive Biology and Endocrinology* 2016;**14**(49):1-11.

Yang 2018 *{published data only}*

* Yang L, Cai S, Zhang S, Kong X, Gu Y, Lu C, et al. Single embryo transfer by Day 3 time-lapse selection versus Day 5 conventional morphological selection: a randomized, open-label, non-inferiority trial. *Human Reproduction* 2018;**33**(5):869-76.

Yang L, Kong X, Zhang S, Dai J, Gong F, Lu G, et al. Single embryo transfer on cleavage-stage(D3) using timelapse selection VS on blastocyst(D5) using traditional morphological selection in patients with good prognosis: a prospective randomized controlled trial. *Human Reproduction* 2017;**32**(Suppl 1):i102-3.

References to studies excluded from this review

Adamson 2016 *{published data only}*

Adamson GD, Abusief ME, Palao L, Witmer J, Palao LM, Gvakharia M. Improved implantation rates of day 3 embryo transfers with the use of an automated time-lapse-enabled test to aid in embryo selection. *Fertility and Sterility* 2016;**105**(2):369-75.

Alhelou 2018 *{unpublished data only}*

Alhelou Y, Mat Adenan NA, Ali J. Embryo culture conditions are significantly improved during uninterrupted incubation: a randomized controlled trial. *Reproductive Biology* 2018;**18**:40-5. [DOI: 10.1016/j.repbio.2017.12.003]

Arnesen 2014 *{published data only}*

Arnesen RE, McEvoy K, Critchlow D, Hunter HR, Lloyd AE, Wilson Y, et al. Comparison of clinical pregnancy rates following day 3 embryo transfer using a time-lapse incubator compared to a flatbed incubator. *Human Fertility* 2014;**17**(4):299.

Belles 2014 *{published data only}*

Belles M, Costa-Borges N, Molina JM, Ballesteros A, Pellicer A, Florensa M, et al. Embryo quality and clinical outcomes using EmbryoscopeTM, MincTM and HeracellTM 150i incubators: preliminary results from a randomized study with donor oocytes. *Human Reproduction* 2014;**29**(Suppl 1):i160.

Cruz 2011 *{published data only}*

Cruz M, Gadea B, Garrido N, Pedersen KS, Martinez M, Perez-Cano I, et al. Embryo quality, blastocyst and ongoing pregnancy rates in oocyte donation patients whose embryos were monitored by time-lapse imaging. *Journal of Assisted Reproduction and Genetics* 2011;**28**(7):569-73.

Freour 2014 *{published data only}*

Freour T, Basile N, Barriere P, Meseguer M. Systematic review on clinical outcomes following selection of human preimplantation embryos with time-lapse monitoring. *Human Reproduction Update* 2015;**21**(1):153-4.

Hardarson 2016 *{published data only}*

Hardarson T, Bungum M, Conaghan J, Meintjes M, Chantilis SJ, Molnar L, et al. Noninferiority, randomized, controlled trial comparing embryo development using media developed for sequential or undisturbed culture in a time-lapse setup. *Fertility and Sterility* 2015;**104**(6):1452-9.

Huang 2014 *{published data only}*

Huang G, Wu LH. Randomized study assessing the impact of Primo Vision time-lapse embryo monitoring system (tlm) on embryo culture and selection. *Human Reproduction* 2014;**29**(Suppl 1):i177.

Ingerslev 2011 *{published data only}*

Ingerslev HJ, Kirkegaard K, Hindkjaer J, Grondahl ML, Kesmodel US, Agerholm I. Cleavage rates of human embryos randomized to culture in conventional incubator and culture in a time lapse instrument. *Human Reproduction* 2011;**26**(Suppl 1):P164.

Kaser 2014 *{published data only}*

Kaser DJ, Racowsky C. Clinical outcomes following selection of human preimplantation embryos with time-lapse monitoring: a systematic review. *Human Reproduction Update* 2014;**20**(5):617-31.

Kirkegaard 2012 *{published data only}*

Kirkegaard K, Hindkjaer JJ, Grondahl ML, Kesmodel US, Ingerslec HJ. A randomized clinical trial comparing embryo culture in a conventional incubator with a time-lapse incubator. *Journal of Assisted Reproduction and Genetics* 2012;**29**(6):565-72.

Kirkegaard 2014 *{published data only}*

Kirkegaard K, Ingerslev HJ. Clinical outcomes following selection of human preimplantation embryos with time-lapse monitoring: a systematic review. *Human Reproduction Update* 2014;**20**(5):617-31.

Kirkegaard 2015 *{published data only}*

Kirkegaard K, Ahlstrom A, Ingerslev HJ, Hardson T. Choosing the best embryo by time lapse versus standard morphology. *Fertility and Sterility* 2015;**103**(2):323-32.

Loewke 2012 *{published data only}*

Loewke K, Moussavi F, Maddah M, Ivani K, Behr B, Suraj V. Development and validation of an automated computer vision algorithm for real-time embryo viability prediction. *Fertility and Sterility* 2012;**98**(Suppl 3):288.

Lowen 2017 *{published data only}*

Lowen PK, Kermack AJ, Wellstead SJ, Calder PC, Houghton FD, Macklon NS. A prospective randomised trial comparing embryo development in the MINC incubator versus the EmbryoScope incubator. *Human Reproduction* 2017;**32**(Suppl 1):i210.

Mara 2010 *{published data only}*

Mara C, Nicols G, Inmaculada P-C, Niels R, Manuel M, Marcos M. Comparative study of embryo quality, blastocyst and ongoing pregnancy rates in oocyte donation patients sharing EmbryoScope and standard incubator. *Fertility and Sterility* 2010;**94**(4 Suppl 1):s78.

Meseguer 2012 *{published data only}*

Meseguer M. Looking at embryo development - clinical results from time-lapse RCT. *Human Reproduction* 2012;**27**(Suppl 2):ii36.

Nakahara 2010 *{published data only}*

Nakahara T, Iwase A, Goto M, Harata T, Suzuki M, Ienaga M, et al. Evaluation of the safety of time-lapse observations for human embryos. *Journal of Assisted Reproduction and Genetics* 2010;**27**(2-3):93-6.

Polanski 2014 *{published data only}*

Polanski LT, Coelho Neto MA, Nastri CO, Navarro PA, Ferriani RA, Raine-Fenning N, et al. Time-lapse embryo imaging for improving reproductive outcomes: systematic review and meta-analysis. *Ultrasound in Obstetrics & Gynecology* 2014;**44**(4):394-401.

Siristatidis 2015 *{published data only}*

Siristatidis C, Komitopoulou MA, Makris A, Sialakouma A, Botzaki M, Mastorakos G, et al. Morphokinetic parameters of early embryo development via time lapse monitoring and their effect on embryo selection and ICSI outcomes: a prospective cohort study. *Journal of Assisted Reproduction and Genetics* 2015;**32**(4):563-70.

Wu 2015 *{published data only}*

Wu Y, Lazzaroni-Tealdi E, Wang Q, Albertini DF, Barad DH, Kushnir VA, et al. Randomized comparison of embryo development in closed time-lapse photography system with traditional standard embryology culture with day-3 embryo transfers. *Fertility and Sterility* 2015;**104**(Suppl 3):e319.

Yang 2014 *{published data only}*

Yang Z, Zhang J, Salem S, Liu X, Kuang Y, Salem R, et al. Selection of competent blastocysts for transfer by combining time-lapse monitoring and array CGH testing for patients undergoing preimplantation genetic screening: a prospective study with sibling oocytes. *BMC Medical Genomics* 2014;**7**(38):1-13.

References to studies awaiting assessment

Hulme 2014 *{published data only}*

Hulme D, Jenner LJ, Campbell A, Fishel S. A randomised controlled comparison of clinical outcome, following time lapse and standard incubation. *Human Fertility. Conference: ACE 9th Biennial Conference* 2014;**17**.

References to ongoing studies

ChiCTR1800017127 *{published data only}*

ChiCTR1800017127. A prospective randomized controlled study of cleavage embryo transplantation using dynamic observation techniques. www.chictr.org.cn/showprojen.aspx?proj=29089 (registered 16 January 2016).

ChiCTR-IIR-16008758 *{published data only}*

ChiCTR-IIR-16008758. A phase IV, prospective, randomized, controlled, multiple-center study to compare the implantation rate of embryos cultured and assessed in GERI? undisturbed incubation system vs conventional open incubator system. www.chictr.org.cn/showprojen.aspx?proj=14790 (registered 1 July 2016).

ISRCTN17792989 *{published data only}*

ISRCTN17792989. A pragmatic, multi-centre, three-arm randomised controlled trial to assess the clinical effectiveness and safety of time lapse imaging in in-vitro fertilisation treatment (also known as the TILT study). www.isrctn.com/ISRCTN17792989 (first received 4 April 2018).

Khan 2018 [pers comm] *{published data only}*

* Khan K. [personal communication]. Conversation with: Sarah Armstrong (University of Sheffield, UK) March 2018.

NCT01760278 *{unpublished data only}*

NCT01760278. Assessment of implantation potential of embryos by time-lapse technology (Embryoscope). www.clinicaltrials.gov/ct2/show/NCT01760278 (first received 4 January 2013).

NCT02222831 *{unpublished data only}*

NCT02222831. Optimizing IVF treatment - the impact of time-lapse culture and preimplantation factor (PIF) on embryo development. www.clinicaltrials.gov/ct2/show/NCT02222831 (first received 21 August 2014).

NCT02417441 *{unpublished data only}*

NCT02417441. TiLE (Time Lapse Eeva) Clinical Trial (TiLE). www.clinicaltrials.gov/ct2/show/NCT02417441 (first received 15 April 2015).

NCT02657811 *{unpublished data only}*

NCT02657811. Time-lapse incubation for embryo culture - morphokinetics and environmental stability. www.clinicaltrials.gov/ct2/show/NCT02657811 (first received 18 January 2016).

NCT02852356 *{unpublished data only}*

NCT02852356. Validation study using a time-lapse morphometry MIRI imaging incubator (TiMMI). www.clinicaltrials.gov/ct2/show/NCT02852356 (first received 2 August 2016).

NCT02965222 *{unpublished data only}*

NCT02965222. A study select top-grade embryo by time-lapse imaging. www.clinicaltrials.gov/ct2/show/NCT02965222 (first received 16 November 2016).

NCT03164551 *{unpublished data only}*

NCT03164551. TICON-Day 3, Time lapse versus conventional method in day 3 embryo culture and assessment (TICON). www.clinicaltrials.gov/ct2/show/NCT03164551 (first received 23 May 2017).

NCT03445923 *{unpublished data only}*

NCT03445923. Can time-lapse parameters be used to predict pregnancy of human embryos?. www.clinicaltrials.gov/ct2/show/NCT03445923 (first received 26 February 2018).

NTR5423 *{published data only}*

* NTR5423. Time-lapse monitoring in IVF and ICSI patients [Embryo SELECtion using TIme-lapse MOnitoring in IVF and ICSI patients]. www.trialregister.nl/trial/5314 (registered 8 September 2015).

Additional references

ACART

ACART. Advisory Committee on Assisted Reproductive Technology: Guiding Principles. www.acart.health.govt.nz/about-us (accessed 30 June 2014).

Alpha & ESHRE SIG 2011

Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. *Human Reproduction* 2011;**26**(6):1270-83. [1472-6483]

Armstrong 2018

Armstrong S, Bhide P, Jordan V, Pacey A, Farquhar C. Time-lapse systems for ART. *Reproductive Biomedicine Online* 2018;**36**(3):288-9.

Campbell 2000

Campbell M, Fitzpatrick R, Haines A, Kinmouth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000;**321**(7262):694-6.

Chen 2017

Chen M, Wei S, Hu J, Yuan J, Liu F. Does time-lapse imaging have favorable results for embryo incubation and selection compared with conventional methods in clinical in vitro fertilization? A meta-analysis and systematic review of randomized controlled trials. *PLOS ONE* 2017;**12**(6):e0178720.

Conaghan 2013

Conaghan J, Chen AA, Willman SP, Ivani K, Chenette PE, Boostanfar R, et al. Improving embryo selection using a computer-automated time-lapse image analysis test plus day 3 morphology: results from a prospective multicenter trial. Fertility and Sterility 2013; Vol. 100, issue 2:412-9.e5. [0015-0282]

Covidence [Computer program]

Veritas Health Innovation. Covidence. Version accessed 10 January 2019. Melbourne, Australia: Veritas Health Innovation.

Cummins 1986

Cummins JM, Breen TM, Harrison KL, Shaw JM, Wilson LM, Hennessey JF. A formula for scoring human embryo growth rates in in vitro fertilization: its value in predicting pregnancy and in comparison with visual estimates of embryo quality. *Journal of In Vitro Fertilization and Embryo Transfer* 1986;**3**(5):284-95.

Finn 2010

Finn A, Scott L, O'Leary T, Davies D, Hill J. Sequential embryo scoring as a predictor of aneuploidy in poor-prognosis patients. *Reproductive BioMedicine Online* 2010;**21**(3):381-90. [1472-6483]

GRADEpro GDT 2015 [Computer program]

McMaster University (developed by Evidence Prime). GRADEpro GDT. Version accessed 9 May 2018. Hamilton (ON): McMaster University (developed by Evidence Prime), 2015.

Harper 2012

Harper J, Magli MC, Lundin K, Barratt CL, Brison D. When and how should new technology be introduced into the IVF laboratory?. *Human Reproduction* 2012;**27**(2):303-13.

HFEA

HFEA. Human Fertilisation and Embryology Authority. www.hfea.gov.uk/ (accessed 30 June 2014).

Higgins 2011

Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Meseguer 2012a

Meseguer M, Rubio I, Cruz M, Basile N, Marcos J, Requena A. Embryo incubation and selection in a time-lapse monitoring system improves pregnancy outcome compared with a standard incubator: a retrospective cohort study. *Fertility and Sterility* 2012;**98**(6):1481-9.e10. [0015-0282]

Neuber 2003

Neuber E, Rinaudo P, Trimarchi JR, Sakkas D. Sequential assessment of individually cultured human embryos as an indicator of subsequent good quality blastocyst development. *Human Reproduction* 2003;**18**(6):1307-12.

Neuber 2006

Neuber E, Mahutte NG, Arici A, Sakkas D. Sequential embryo assessment outperforms investigator-driven morphological assessment at selecting a good quality blastocyst. *Fertility and Sterility* 2006;**85**(3):794-6. [0015-0282]

Petersen 2016

Petersen BM, Boel M, Montag M, Gardner DK. Development of a generally applicable morphokinetic algorithm capable of predicting the implantation potential of embryos transferred on Day 3. *Human Reproduction* 2016;**31**(10):2231–44.

Pribenszky 2017

Pribenszky C, Nilselid A-M, Montag M. Time-lapse culture with morphokinetic embryo selection improves pregnancy and live birth chances and reduces early pregnancy loss: a meta-analysis. *Reproductive BioMedicine Online* 2017;**35**(5):511-20. [DOI: 10.1016/j.rbmo.2017.06.022]

Scott 2003

Scott L. The biological basis of noninvasive strategies for selection of human oocytes and embryos. *Human Reproduction Update* 2003;**9**(3):237-49.

Scott 2003a

Scott L. Pronuclear scoring as a predictor of embryo development. *Reproductive BioMedicine Online* 2003;**6**(2):201-14.

Shoukir 1997

Shoukir Y, Campana A, Farley T, Sakkas D. Early cleavage of in-vitro fertilized human embryos to the 2-cell stage: a novel

indicator of embryo quality and viability. *Human Reproduction* 1997;**12**(7):1531-6.

References to other published versions of this review

Armstrong 2014

Armstrong S, Arroll N, Cree LM, Jordan V, Farquhar C. Time-lapse systems for embryo incubation and assessment in assisted reproduction. *Cochrane Database of Systematic Reviews* 2014, Issue 9. [DOI: 10.1002/14651858.CD011320]

Armstrong 2015

Armstrong S, Arroll N, Cree LM, Jordan V, Farquhar C. Time-lapse systems for embryo incubation and assessment in assisted reproduction. *Cochrane Database of Systematic Reviews* 2015, Issue 2. [DOI: 10.1002/14651858.CD011320.pub2]

Armstrong 2018a

Armstrong S, Bhide P, Jordan V, Pacey A, Farquhar C. Time-lapse systems for embryo incubation and assessment in assisted reproduction. *Cochrane Database of Systematic Reviews* 2018, Issue 5. [DOI: 10.1002/14651858.CD011320.pub3]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Barberet 2018

Methods	Study: completed single-centre RCT of couples with infertility undergoing ICSI Country: France Cause and length of infertility: male factor (76% to 77%), female factor (42% to 46%). Mixed (64% to 74%), idiopathic (3%). Oocytes: autologous oocytes Embryo transfer: 1 or 2 fresh embryos on day 2, day 3 or day 5-6. Informed consent: not mentioned Total study duration: March 2016 to December 2016 Funding sources: not mentioned
Participants	A total of 386 couples with infertility undergoing ICSI with autologous oocytes were randomised: 191 to TLS selection (closed system, Embryoscope incubator) and 195 to conventional selection (benchtop G185 incubator). There were 4 misallocations (1 in G185 group and 3 in the TLI group) and 1 participant not fulfilling the inclusion criteria (only 2 injected oocytes). Data analysed by intention-to-treat as well as per-protocol. Age (years, mean ± SD, time-lapse selection versus conventional selection): 32.1 ± 4.8 versus 32.3 ± 4.6 BMI (kg/m^2, mean ± SD, time-lapse selection versus conventional selection): 23.5 ± 3.8 versus 24 ± 4.3 Ethnicity: not reported Inclusion criteria: • no more than 42 years of age; • undergoing ICSI; • able to provide at least 6 mature oocytes after denudation. Exclusion criteria: • requiring egg donations; • disorders such as hydrosalpinx or obesity (BMI > 32) and uterine diseases and attempts with surgical spermatozoa or performed in a viral context.

Barberet 2018 *(Continued)*

Interventions	TLS with conventional morphological assessment of still TLS images (intervention) Conventional incubation and assessment (control)
Outcomes	Ongoing pregnancy per couple randomised (defined as the presence of a gestational sac with a foetal heartbeat at >/= 12 weeks) (obtained from email communication with author) Miscarriage of clinical pregnancy per couple randomised (updated ongoing-pregnancy and detailed miscarriage rates obtained from authors following email communication) Clinical pregnancy (with at least 1 intrauterine gestational sac visible on the ultrasound examination 4 to 5 weeks after ET)
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "The allocation algorithm, which relied on a minimization approach, was established by the statistician of the coordinating centre before the start of the trial. This allocation was stratified on woman's age (<orR37 years), day of oocyte retrieval (Friday [leading to ET at day 3] or not [leading to ET at day 2]), and rank of attempts (rank 1–2 or 3-4)."
Allocation concealment (selection bias)	Low risk	Quote: "The randomization was performed online by the investigator (embryologist) using the secure Tenalea platform (Formsvision BV), after identification through a personal password and after a final check of the eligibility criteria"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "Due to the nature of the study intervention, it was not possible to blind investigators to the embryo morphology assessments. However, for the analyses the data manager, statistician, and embryologists were blinded to the allocation." "Embryos were selected for vitrification according to their morphology, which was graded in unblinded embryo assessments"
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "Due to the nature of the study intervention, it was not possible to blind investigators to the embryo morphology assessments. However, for the analyses the data manager, statistician, and embryologists were blinded to the allocation."
Incomplete outcome data (attrition bias) All outcomes	Low risk	No losses to follow-up
Selective reporting (reporting bias)	Low risk	Primary outcome (embryo implantation) is as per registered protocol. Pregnancy outcomes not mentioned in protocol.
Other bias	Low risk	No other sources of bias identified.

Goodman 2016

Methods	Study: completed single-centre RCT of couples with infertility undergoing IVF Country: USA

Goodman 2016 *(Continued)*

	Cause and length of infertility: infertility diagnosis included unexplained, ovulatory dysfunction, male factor, tubal factor, low ovarian reserve, AMA, endometriosis, mixed factors, and other. Mean length of infertility in both groups was approximately 31.5 months. Oocytes: autologous oocytes Embryo transfer: between 1 and 3 fresh embryos on day 3 or day 5. The number of embryos transferred was based on published ASRM committee guidance and patient preferences. Informed consent: yes Total study duration: March 2014 to May 2015 (14 months) Funding sources: quote: "no external funding for the study"
Participants	A total of 300 couples with infertility undergoing IVF with autologous oocytes were recruited: 150 randomised to TLS selection (cell-tracking algorithm of TLS utilised) and 150 randomised to conventional selection (TLS with conventional once-daily morphologic embryo screening). 5 couples did not receive the allocated intervention: 2 from the time-lapse selection arm due to lack of fertilisation, and 3 from the conventional selection group, 2 due to no fertilisation and 1 due to no sperm. Age (years, mean ± SD, time-lapse selection versus conventional selection): 33.6 ± 4.0 versus 33.2 ± 3.9 BMI (kg/m^2, mean ± SD, time-lapse selection versus conventional selection): 26.3 ± 6.7 versus 26.9 ± 7.4 Ethnicity: combination of white, black, Asian, Middle Eastern, and other Inclusion criteria: • aged 18 to 43 years; • undergoing autologous IVF cycle between March 2014 and May 2015; • plan for fresh embryo transfer. Exclusion criteria: • did not undergo fresh transfer owing to previously unforeseen reasons; • women with only 1 to 3 zygotes.
Interventions	TLS utilising cell-tracking algorithm (intervention) TLS with conventional assessment of morphological parameters from still TLS images (control)
Outcomes	Clinical pregnancy rate per couple randomised (defined by the presence of foetal cardiac activity on transvaginal ultrasonography at >= 6 weeks gestational age)Miscarriage per couple randomised
Notes	Data on clinical pregnancy from women excluded following randomisation and miscarriage data were obtained following communication with the authors. Live-birth and stillbirth data were requested, but this information was not available.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Patients were randomized 1:1 to conventional embryo selection versus Embryoscope time-lapse morphokinetic selection with the use of a computer-generated random number sequence"
Allocation concealment (selection bias)	Low risk	Quote: "The list was housed in the laboratory, where it was accessible only by research personnel not involved with the recruitment of patients"

Goodman 2016 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "Patients, physicians and staff, and sonographers were blinded to how embryos were selected". However, the embryologist who was responsible for deciding on day of embryo transfer (day 3 or day 5) was unblinded, therefore deemed high risk.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "sonographers were blinded"
Incomplete outcome data (attrition bias) All outcomes	Low risk	We have obtained all relevant data from women who were excluded postrandomisation from the authors.
Selective reporting (reporting bias)	Low risk	We confirmed with the authors that all outcomes the study set out to assess were published.
Other bias	Low risk	No other sources of bias identified.

Kahraman 2013

Methods	Study: completed single-centre RCT of couples with infertility undergoing ICSI Country: Turkey Cause and length of infertility: tubo-peritoneal factor. Length of infertility not reported. Oocytes: autologous oocytes Embryo transfer: single embryo transfer at blastocyst Informed consent: yes Total study duration: December 2011 to June 2012 (6 months) Funding sources: none
Participants	A total of 76 couples with infertility undergoing ICSI with autologous oocytes were recruited: 38 were randomised to TLS and 38 were randomised to conventional incubation. In all, 12 couples withdrew from the study: 7 in the conventional incubation arm and 5 in the TLS arm. Reasons for withdrawal were documented and data for outcomes such as live birth, adverse events, and clinical pregnancy for these couples were included in this review. Age (years, mean ± SD, TLS versus conventional incubation): 28.5 ± 3.32 versus 28.5 ± 3.72; P = 0.83 BMI (kg/m^2, mean ± SD, TLS versus conventional incubation): 23.92 ± 3.79 versus 23.92 ± 4.42; P = 0.77 Ethnicity: not reported Inclusion criteria: • first or second treatment cycle; • age < 35 years, BMI < 28 kg/m^2; • ≥ 8 oocytes retrieved. Exclusion criteria: • recurrent spontaneous abortions; • severe endometriosis;

Kahraman 2013 *(Continued)*

	• PCOS; • hydrosalpinx; • uterine pathology; • severe male factor (< 5 million motile sperm in total ejaculate); • very severe morphological sperm defects (dominantly globozoospermic or macrocephalic samples).
Interventions	TLS with conventional morphological assessment of still TLS images (intervention) Conventional incubation and assessment (control)
Outcomes	Live-birth rates per couple randomised Clinical pregnancy rate per couple randomised (clinical pregnancy was defined as the presence of a gestational sac detected on ultrasound 3 weeks after the first βhCG test, which was performed 14 days after oocyte retrieval) Stillbirth and miscarriage per couple randomised
Notes	Live-birth and stillbirth information was available following communication with the author and was not published.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Computer based randomization list" Quote: "Randomisation was done according to a list generated on random.org"
Allocation concealment (selection bias)	Low risk	Communication with author. Quote: "Randomization list was held by one of the investigators who was not involved clinically with the patients. Also, he was not routinely working in the embryology laboratory. The randomization from random.org was printed out into sequentially numbered lists where the groups were masked and not revealed until the recruitment of each patient"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Communication with author. Quote: "Clinicians were blinded in the study up to the point after the embryo transfer was performed. Also the patients did not know to which group they were allocated. Only the discontinued patients received information about the incubation process once the drop-out decision was made (Due to the need to inform the patients about their early/cancelled transfers)". It was impossible to blind the embryologist, therefore performance bias deemed high risk.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Communication with author. Quote: "Clinicians, those assessing the outcome were not necessarily blinded to the intervention as some of our ART patients prefer to have those controls outside our clinic and report the outcomes to us". The outcomes are objective and are therefore unlikely to be influenced by knowledge of the intervention, therefore detection bias deemed low risk.
Incomplete outcome data (attrition bias) All outcomes	Low risk	A total of 12 couples discontinued the trial following randomisation secondary to adverse events that were not reported as adverse events or analysed within the main publication. However, on communication with the author, the numbers of discontinued participants in each arm were disclosed, alongside reasons for dropouts. Quote: "embryos transferred day 3, 4 and 5 with single blastocyst developed; total freezing because of ovarian hyperstimulation syndrome (OHSS) risk"

Kahraman 2013 *(Continued)*

Selective reporting (reporting bias)	Low risk	Communication with author. Quote: "As reported in our article, we have published all of the outcomes we aimed to assess. Unfortunately, we do not formally prepare a study protocol". On contacting the author, data on live birth and adverse events were made available, although this information was not published.
Other bias	Low risk	None detected.

Kaser 2017

Methods	Study: completed RCT of couples with infertility undergoing a fresh SET Country: USA Cause and length of infertility: a combination of anovulation, diminished ovarian reserve, endometriosis, male factor, tubal, unknown, uterine, and other Oocytes: autologous oocytes Embryo transfer: single embryo transfer Informed consent: yes Total study duration: August 2014 to February 2016 (18 months) Funding sources: Progyny Inc
Participants	A total of 163 couples with infertility undergoing ART with autologous oocytes were recruited: • 56 were randomised to TLS and day 3 ET; • 54 were randomised to TLS and day 5 ET; • 53 were randomised to incubation within the TLS and conventional morphology with day 5 ET (control). In all, 13 couples did not receive the allocated intervention: • 7 in the TLS and day 3 ET arm (1 due to freeze-all for OHSS risk; 2 embryos transferred in 1 woman; in 1 woman the TLS algorithm was not followed; and 4 women elected to have a day 5 ET); • 2 from the TLS and day 5 ET arm (2 women had freeze-all for OHSS risk); • 4 from the control arm (3 women had freeze-all for OHSS risk, and 1 woman had 2 embryos transferred). Age (years, mean ± SD): Day 3 + TLS 34.6 ± 3.1, Day 5 + TLS 33.7 ± 3.4, Day 5 control 34.1 ± 3.1 BMI (kg/m^2, mean ± SD): Day 3 + TLS 26 ± 6.9, Day 5 + TLS 25.5 ± 6.1, Day 5 control 25.5 ± 6.5 Ethnicity: a combination of white, Asian, black, Hispanic, and "other" ethnicities Inclusion criteria: • patients with a planned fresh SET; • aged 18 to 40 years; • can only be randomised if fertilisation occurs. Exclusion criteria: • use of donor oocytes; • more than 3 prior retrievals without an intervening clinical pregnancy; • in vitro maturation;

Kaser 2017 *(Continued)*

	• gestational carrier; • pre-implantation genetic diagnosis or screening; • presence of an uninterrupted hydrosalpinx; • history of intrauterine adhesions; • all embryos frozen due to ovarian hyperstimulation risk prior to randomisation; • less than 4 zygotes and therefore at risk of no blastocyst development.
Interventions	TLS utilising conventional benchtop morphology and embryo selection software (2 intervention arms: day 3 and day 5 embryo transfer) TLS with conventional benchtop morphology (control). Embryo selection software or time-lapse photography was not utilised.
Outcomes	Clinical pregnancy rate per couple randomised Miscarriage rate per couple randomised (data obtained from authors)
Notes	Wrote to authors August 2017 asking for further information. Note differing days of embryo transfer. Control group split between 2 intervention groups for purposes of this review.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Subjects were blocked according to age (<35, 35-37, 38-40 years) and randomised 1:1:1 at the fertilization check by an embryologist using computer-generated, random number sequence cards enclosed in opaque, serially numbered envelopes"
Allocation concealment (selection bias)	Low risk	Quote: "random number sequence cards enclosed in opaque, serially numbered envelopes"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "Embryologists were blinded to the Eeva (time lapse) ratings at the conventional morphology evaluation (i.e. one embryologist performed conventional morphology and a different embryologist reviewed the Eeva ratings, and patients and physicians were blinded to the Eeva ratings until a negative pregnancy test of the primary endpoint was reached". The embryologist was ultimately unblinded to the allocation, therefore high risk of performance bias.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Participants and physicians were blinded to the Eeva ratings. Correspondence with author. Quote: "As patients were randomised to day 3 or day 5 transfer, blinding was not possible between groups 1 vs. group 2/3 (as the patient and physician knew which day the transfer was happening). For patients randomised to groups 2 or 3, both patients and physicians were blinded to study arm (so they knew a day 5 transfer was happening, but not how the embryo was selected for transfer)". The outcomes are objective and are therefore unlikely to be influenced by knowledge of the intervention, therefore detection bias deemed as low risk.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data presented as intention-to-treat and "as treated".
Selective reporting (reporting bias)	Low risk	Communication with authors. Quote: "All outcomes published"

Kaser 2017 *(Continued)*

Other bias	Low risk	None detected.

Kovacs 2019

Methods	Study: completed multicentre RCT of couples with infertility undergoing IVF or ICSI Country: Hungary Cause and length of infertility: various causes (male, tubal, unexplained, etc.) of at least 1 year's duration Oocytes: autologous Embryo transfer: single embryo transfer at blastocyst Informed consent: yes Total study duration: July 2012 to April 2015 (33 months) Funding sources: none
Participants	161 couples with infertility undergoing IVF or ICSI with single embryo transfer at blastocyst. 80 couples were randomised to TLS and 81 were randomised to conventional incubation. 22 couples dropped out of the study after randomisation: 12 dropped out from the TLS arm (2 dual embryo transfer requested; 1 no fertilisation; 7 fewer than 3 good embryos on day 3; 2 elective cryopreservation for OHSS risk), and 10 dropped out from the control arm (1 no fertilisation; 8 fewer than 3 good embryos on day 3; 1 elective cryopreservation for OHSS risk). Age: (years, mean ± SD, TLS versus conventional incubation): 31.2 ± 2.7 versus 32.1 ± 2.5 BMI: (kg/m^2, mean ± SD, TLS versus conventional incubation): 22.3 ± 3.3 versus 22.2 ± 3.0 Ethnicity: Caucasian (understood to be white) Inclusion criteria: • age < 36 years; • baseline FSH < 10 IU/L; • regular 25- to 35-day cycles; • less than 2 previous failed IVF cycles (first or second cycle); • intact uterus; • an indication for IVF; • BMI > 18 to < 30 kg/m^2; • acceptance of single embryo transfer; • normal ovarian reserve; • at least 3 good embryos on day 3. Exclusion criteria: • PCOS; • sperm obtained by surgical extraction; • chromosome abnormality; • presence of hydrosalpinx; • stage III/IV endometriosis; • fewer than 3 good-quality day 3 embryos;

Kovacs 2019 *(Continued)*

- lack of consent.

Interventions	TLS utilising cell-tracking algorithm (intervention) Conventional incubation and assessment (control)
Outcomes	Clinical pregnancy rate per couple Miscarriage per couple randomised Live birth
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	On communication with author, paired randomisation sequence was explained: Quote: "Two envelopes containing time-lapse or control group assignment were prepared. The first patient was randomly assigned to one of the groups and the next patient received the other assignment. This was repeated with patient number 3 and 4 and so on"
Allocation concealment (selection bias)	High risk	Communication with author. Quote: "The randomization is carried out by the principal investigator who is involved in the study"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Communication with author. Quote: "There was no blinding"
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Communication with author. Quote: "There was no blinding". The outcomes are objective and are therefore unlikely to be influenced by knowledge of the intervention, therefore detection bias deemed as low risk.
Incomplete outcome data (attrition bias) All outcomes	High risk	Dropouts following randomisation and not included in intention-to-treat: 161 participants were randomised (80 TLS versus 81 standard monitoring), of which 22 participants dropped out. Reasons for dropouts were provided, however the reasons provided were not all predetermined exclusion criteria, and given the high attrition rate, we deemed this study as at high risk of attrition bias.
Selective reporting (reporting bias)	Low risk	None detected.
Other bias	Low risk	None detected.

Park 2015

Methods	Study: single-centre RCT, couples undergoing ICSI Country: Sweden

Park 2015 *(Continued)*

	Cause and length of infertility: male-factor infertility was present in > 99% of participants in both study arms. Female-factor infertility was present in approximately 20% of participants in both study arms. Duration of infertility was approximately 2.8 years in both study arms. Oocytes: autologous Embryo transfer: single embryo transfer at day 2 Informed consent: yes Total study duration: May 2010 to February 2014 (3 years, 9 months) Funding sources: Sahlgrenska Academy, Sahlgrenska University Hospital, LUA/ALF 70940, Ferring Research Infertility and Gynecology Grant, Hjalmar Svensson Grant, Unisense Fertilitech: Unisense provided the EmbryoScope free of charge during the study.
Participants	364 couples with infertility undergoing their first IVF cycle with ICSI. 1 embryo (in a few cases 2 embryos, N = 12) of good quality, or in some cycles of less good quality (N = 27), was transferred on day 2, and supernumerary good-quality embryos were frozen. 241 couples were randomised to TLS, and 124 were randomised to conventional incubation. 1 couple was excluded from the TLS arm as they had been randomised twice. Age: (years, mean ± SD, TLS versus conventional incubation): 31.8 ± 4.3 versus 31.8 ± 4.1; P = 0.90 BMI: (kg/m^2, mean ± SD, TLS versus conventional incubation): 24.4 ± 3.9 versus 24.3 ± 4.0; P = 0.70 Ethnicity: not reported Inclusion criteria: • ≤ 40 years of age; • undergoing their first IVF cycle using ICSI; • at least 1 oocyte was retrieved. Exclusion criteria: • patients undergoing egg donation.
Interventions	TLS with conventional morphological assessment of still TLS images (intervention) Conventional incubation and assessment (control)
Outcomes	Clinical pregnancy rate per couple randomised Ongoing pregnancy rate defined as presence of the foetal heart at >= 8 weeks' gestation Miscarriage per couple randomised
Notes	Live-birth, stillbirth, and clinical pregnancy data obtained on communication with study authors.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation was undertaken using (quote): "a web-based randomization programme and all the patients' oocytes were allocated to culture in either a conventional incubator or in a closed system, in proportion 1:2"
Allocation concealment (selection bias)	Low risk	Quote: "Randomization was carried out by the embryologist after oocyte retrieval". On communication with the authors, they clarified that the embryolo-

Park 2015 *(Continued)*

		gist undertaking the randomisation may have also undertaken the embryo assessment.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "The patients as well as the treating physician and the person performing the statistical analyses were blinded to which type of procedure was used until the outcome of transfer (pregnant versus not pregnant) was known". It was not possible to blind the embryologists, therefore performance bias deemed at high risk.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "The patients as well as the treating physician and the person performing the statistical analyses were blinded to which type of procedure was used until the outcome of transfer (pregnant versus not pregnant) was known. Embryologists were not possible to blind"
Incomplete outcome data (attrition bias) All outcomes	Low risk	Only 1 woman was excluded from analysis in the intervention arm, as she was randomised twice. No women were excluded from the control arm. No dropouts.
Selective reporting (reporting bias)	Low risk	All predetermined outcomes were reported.
Other bias	Low risk	None detected.

Rubio 2014

Methods	Study: completed multicentre RCT of couples with infertility undergoing ICSI Country: Spain Cause and length of infertility: not reported Oocytes: autologous and donor Embryo transfer: multiple embryo transfer (1.86 per couple, 95% CI 1.8 to 1.9) on day 3 and day 5 Informed consent: not reported Total study duration: February 2012 to July 2013 (17 months) Funding sources: the instrumentation, disposables, and utensils used in this study were fully paid for by IVI. IVI is a minor shareholder in Unisense FertiliTech A/S, but none of the authors have any economic affiliation with Unisense FertiliTech A/S.
Participants	A total of 856 couples with infertility undergoing IVF with autologous and donor oocytes: 444 couples were randomised to TLS and 412 to conventional incubation. In all, 13 couples were excluded from the study: 6 in the TLS arm (reasons: 2 had cancelled oocyte donation, and 4 had their embryos vitrified) and 7 in the conventional incubation arm (reasons: 1 woman had endometrial bleeding; 2 had cancelled oocyte donation; and 4 couples had their embryos vitrified). Age (years, mean ± SD, TLS versus conventional incubation): 34.7 ± 2.7 versus 34.6 ± 2.7 BMI (kg/m^2, mean ± SD, TLS versus conventional incubation): 23.2 ± 3.7 versus 23.04 ± 2.8 Ethnicity: not reported Inclusion criteria: autologous or oocyte donation. Those receiving oocyte donation had 1 of the following diagnoses: failure to achieve pregnancy after at least 3 cycles of ART, genetic female or chromosomal disorders, or low response to controlled ovarian hyperstimulation.

Rubio 2014 *(Continued)*

	Donors were: • aged 18 to 34 years; • BMI 18 to 25 kg/m^2; • had received no endocrine treatment (including gonadotropins and oral contraception) for the last 3 months preceding the study and had a normal uterus and ovaries at transvaginal ultrasound scan (no signs of PCOS). Inclusion criteria for both arms of study: • age 20 to 38 years; • first or second ICSI cycle; • BMI of > 18 and < 25 kg/m^2. Exclusion criteria: • severe male factor (total motile sperm < 1 million); • hydrosalpinx; • presenting uterine diseases after 2D ultrasound evaluation and/or 3D (if in doubt) or hysteroscopy (for acquired or congenital uterine abnormalities); • endocrinopathies (thrombophilia); • recurrent pregnancy losses; • endometriosis; • patients receiving concomitant medications as a treatment for any other condition that might interfere with the results of the study. For autologous oocyte patients: • low-responder patients (fewer than 6 metaphase II per cycle) or those with an FSH basal determination > 12 or an anti-M üllerian hormone concentration of < 1.7 pmol/L (based on authors' own experience) were also excluded.
Interventions	TLS utilising cell-tracking algorithms (intervention) Conventional incubation and assessment (control)
Outcomes	Miscarriage per couple randomised Clinical pregnancy rate per couple randomised Live birth (obtained from Insua 2015 and Insua 2017)
Notes	October 2015: following clarification from authors of comments on this review, it has been made aware to us that the pregnancy data from this study are a combination of biochemical and ongoing pregnancy, therefore the miscarriage data may also include miscarriages from biochemical pregnancies.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Despite adequate random sequence generation, participants were able to request the intervention in some cases, and this was granted. See evidence below: Quote: "Patients were allocated to either TMS (study group) or SI (control group) using a computer generated randomization table which was handled by the embryologist at the laboratory in charge the day before the oocyte retrieval or oocyte donation. The randomization was not perfectly performed as the patient distribution to the two groups would have been expected to be

Rubio 2014 *(Continued)*

		50:50 ratio than the reported 51.9:48.1. The main reason for this deviation was limited patient requests for TMS culture"
Allocation concealment (selection bias)	High risk	In some cases allocation was non-random (see above).
Blinding of participants and personnel (performance bias) All outcomes	High risk	Gynaecologist and statistician were blinded. Participants and embryologist were not blinded. Quote: "The study is considered double blind because 1) the gynaecologist (evaluating the primary effect) did not know to which group the patients had been assigned, and 2) the statistician evaluating the results only knew the incubators by a binary code, not by type" Communication with author. Quote: "The intention was to do triple blinded, but we discovered that some of our patients were informed (because they asked) of the group they were in. Therefore blinding failed in some of our patients. We then decided to describe it as double blind because patients blinding partially failed"
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The gynaecologist evaluating the primary effect was blinded.
Incomplete outcome data (attrition bias) All outcomes	Low risk	A total of 13 participants were excluded from study after randomisation as they suffered adverse events (cancelled oocyte donation, embryos vitrified, and endometrial bleeding). Not included in intention-to-treat, but all excluded participants were accounted for, therefore low risk of attrition bias.
Selective reporting (reporting bias)	Low risk	Reported all outcomes declared on ClinicalTrials.gov On communication with the author: "We are currently collecting data on live birth and stillbirth"
Other bias	Low risk	None detected.

Wu 2016

Methods	Study: completed single-centre RCT of couples with infertility undergoing IVF and ICSI Country: USA Cause and length of infertility: "poor prognosis patients". Length of infertility not reported. Oocytes: autologous oocytes Embryo transfer: day 3 transfer of embryo. Number not disclosed. Informed consent: yes Total study duration: December 2014 to March 2015 (3.5 months) Funding sources: intramural funds from The Center for Human Reproduction and by grants from The Foundation for Reproductive Medicine. Vitrolife, Goteborg, Sweden, contributed a free EmbryoScope for the length of the study. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.
Participants	A total of 49 couples with infertility undergoing IVF or ICSI with autologous oocytes: 24 couples were randomised to TLS and 25 to conventional incubation.

Wu 2016 *(Continued)*

	In all, 18 couples were excluded from the study: 8 in the TLS arm (reasons: 6 had no mature oocytes or no fertilisation after ICSI, and 2 women had their embryos transferred on day 2), and 10 in the conventional incubation arm (reasons: 5 women had no mature oocytes or no fertilisation after ICSI, and 5 women had their embryos transferred on day 2). Age (years, mean ± SD, TLS versus conventional incubation): 38.8 ± 1.0 versus 40.4 ± 1.8 BMI (kg/m^2, mean ± SD, TLS versus conventional incubation): not reported Ethnicity: not reported Inclusion criteria: • couples undergoing autologous IVF (and ICSI) cycles. Exclusion criteria: • not stated.
Interventions	TLS with conventional morphological assessment of still TLS images (intervention) Conventional incubation and assessment (control)
Outcomes	Clinical pregnancy rate per couple randomised (defined as ultrasound confirmation but no gestation was provided)
Notes	Contacted authors August 2017 for further information

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Computer randomization to either TLS or standard embryology was the responsibility of a member of the centre's Statistics Section (SKD) who was completely dissociated from the patient's IVF cycle"
Allocation concealment (selection bias)	Unclear risk	Randomisation was undertaken by a member of the team not associated with the treatment cycle. Quote: "The designation was then reported to the embryology staff which processed the patient's oocytes/embryos"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not described. However, given that it would have been impossible to blind embryologists, performance bias deemed high risk.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Not possible to blind outcome assessors. The outcomes are objective and are therefore unlikely to be influenced by knowledge of the intervention, therefore detection bias deemed low risk.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Excluded participants were accounted for and were considered by trialists to be valid prespecified grounds for exclusion.
Selective reporting (reporting bias)	Unclear risk	No access to protocol
Other bias	Low risk	None detected.

Yang 2018

Methods	Study: completed single-centre RCT of couples with infertility undergoing IVF and ICSI Country: China Cause and length of infertility: quote: "good prognosis patients". Length of infertility not reported. Oocytes: autologous oocytes Embryo transfer: single embryo transfer; day 3 transfer of embryos in intervention group and day 5 transfer in control group Informed consent: obtained from all participants Total study duration: October 2015 to April 2017 (18 months) Funding sources: study funded by Ferring
Participants	A total of 600 couples with infertility undergoing IVF or ICSI with autologous oocytes: 300 couples were randomised to TLS utilising embryo selection software, and 300 couples were randomised to conventional incubation and morphology. In all, 15 couples were excluded from the study for the purpose of modified intention-to-treat analysis: 10 in the TLS arm (6 refused day 3 and time-lapse algorithm; 3 had instrument breakdown; and 1 had an unforeseen medical condition) and 5 in the conventional incubation arm (3 refused day 5 and conventional morphological assessment, and 2 did not receive time-lapse observation). Age (years, mean ± SD, TLS versus conventional incubation): not reported BMI (kg/m^2, mean ± SD, TLS versus conventional incubation): not reported Ethnicity: not reported Inclusion criteria: • Chinese females undergoing first or second fresh autologous IVF (and ICSI) cycles; • </= 36 years; • FSH </= 12 IU/mL on day 3 of cycle; • > 10 oocytes retrieved; • willing to have SET. Exclusion criteria: • underlying uterine conditions including endometriosis, untreated unilateral or bilateral hydrosalpinx; • uterine myoma (multiple, submucous or intramural myoma > 3 cm); • cycle planned for oocyte donation or pre-implantation genetic diagnosis; • recurrent pregnancy loss; • significantly abnormal oocytes; • < 6 normally fertilised embryos (2 polar nuclei); • considered unlikely to complete the study based on the investigator's judgement.
Interventions	TLS utilising embryo selection software (intervention) Conventional incubation and assessment (control)
Outcomes	Live birth per couple randomised (provided following email communication with authors) Miscarriage rate per couple randomised (clinical (gestational sac) pregnancy losses) Clinical pregnancy (defined as presence of gestational sac seen at 4 weeks after embryo transfer)

Yang 2018 *(Continued)*

	Stillbirth (provided following email communication with authors)
Notes	Note differing days of embryo transfer (day 3 for intervention group and day 5 for control). All embryos cultured in TLS to day 3, then control embryos transferred to conventional incubator to day 5. Embryos in control arm evaluated by routine morphological assessment.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Patients were randomised in a 1:1 ratio via online-generated blocks (www.random.org) once they had 2PN (>/=6 normally fertilized oocytes) on Day 1 of the cycle."
Allocation concealment (selection bias)	Low risk	Quote: "The study investigators (YLL and XYK) created the randomization list and study nurses who were unaware of the study protocol enveloped the randomised allocation in a consecutive order. The investigator (YLL) assessed the patient's eligibility and performed the randomization by opening the sealed envelopes."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Communication with authors. Quote: "The study was not blinded because study participants and clinic staff were aware of which group they were following"
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "clinic staff were not blinded" The outcomes are objective and are therefore unlikely to be influenced by knowledge of the intervention, therefore detection bias deemed as low risk.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Low number of dropouts, and reasons for attrition given. Quote: "The subject was excluded from the study post-randomization if she did not undergo fresh transfer due to any unforeseen reason including ovarian hyper-stimulation or uterine disorder."
Selective reporting (reporting bias)	Low risk	All study outcomes were published.
Other bias	High risk	Variation between arms of study in day of transfer (day 3 for intervention and day 5 for control).

AMA: advanced maternal age
ASRM: American Society for Reproductive Medicine
ART: assisted reproductive technology
βhCG: beta human chorionic gonadotropin
BMI: body mass index
CI: confidence interval
ET: embryo transfer
FSH: follicle-stimulating hormone
ICSI: intracytoplasmic sperm injection
IU: international units
IVF: in vitro fertilisation
OHSS: ovarian hyperstimulation syndrome
PCOS: polycystic ovarian syndrome
RCT: randomised controlled trial
SD: standard deviation
SET: single embryo transfer

TLS: time-lapse system
2D: two-dimensional
3D: three-dimensional

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Adamson 2016	Not an RCT
Alhelou 2018	Not an RCT
Arnesen 2014	Pseudo-randomised study - this was established after discussion with the main author, who described allocation to intervention or control based on capacity of either incubator.
Belles 2014	Randomised oocytes
Cruz 2011	Randomised oocytes
Freour 2014	Letter not containing study data
Hardarson 2016	Randomised embryos, and study design not relevant
Huang 2014	Unable to determine the nature of the control arm
Ingerslev 2011	Randomised oocytes
Kaser 2014	Systematic review
Kirkegaard 2012	Randomised oocytes
Kirkegaard 2014	Letter not containing study data
Kirkegaard 2015	Systematic review
Loewke 2012	Not an RCT
Lowen 2017	Randomised embryos
Mara 2010	Randomised oocytes
Meseguer 2012	Not an RCT
Nakahara 2010	Randomised oocytes
Polanski 2014	Systematic review
Siristatidis 2015	Non-randomised study
Wu 2015	Randomised embryos
Yang 2014	Randomised oocytes

RCT: randomised controlled trial

FSH: follicle-stimulating hormone

ICSI: intracytoplasmic sperm injection

IU: international unit

IVF: in vitro fertilisation

mIU: milli-international unit

SET: single embryo transfer

2PN: 2 pronuclei

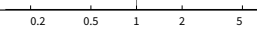
DATA AND ANALYSES

Comparison 1. TLS with conventional morphological assessment of still TLS images versus conventional incubation and assessment (trial design 1)

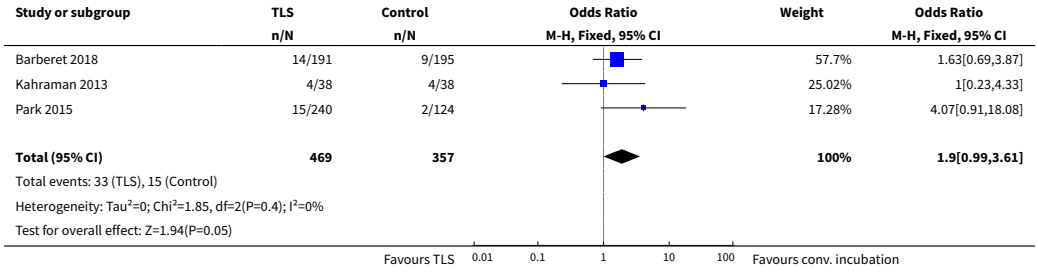
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Live birth or ongoing pregnancy	3	826	Odds Ratio (M-H, Fixed, 95% CI)	0.91 [0.67, 1.23]
2 Miscarriage	3	826	Odds Ratio (M-H, Fixed, 95% CI)	1.90 [0.99, 3.61]
3 Stillbirth	1	76	Odds Ratio (M-H, Fixed, 95% CI)	1.0 [0.13, 7.49]
4 Clinical pregnancy	4	875	Odds Ratio (M-H, Fixed, 95% CI)	1.06 [0.79, 1.41]

Analysis 1.1. Comparison 1 TLS with conventional morphological assessment of still TLS images versus conventional incubation and assessment (trial design 1), Outcome 1 Live birth or ongoing pregnancy.

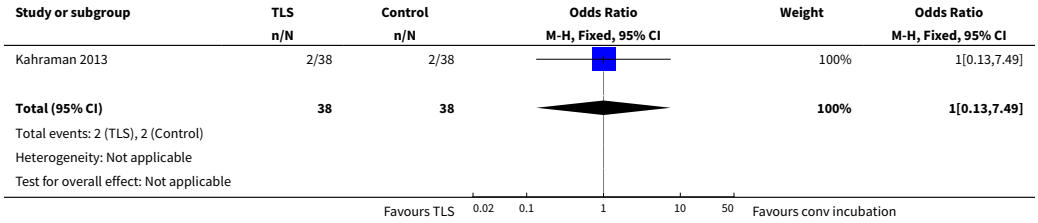
Study or subgroup	TLS n/N	Control n/N	Odds Ratio M-H, Fixed, 95% CI	Weight	Odds Ratio M-H, Fixed, 95% CI
Barberet 2018	73/191	70/195		48.24%	1.1[0.73,1.67]
Kahraman 2013	20/38	19/38		10.14%	1.11[0.45,2.73]
Park 2015	48/240	35/124		41.62%	0.64[0.38,1.05]
Total (95% CI)	**469**	**357**		**100%**	**0.91[0.67,1.23]**
Total events: 141 (TLS), 124 (Control)					
Heterogeneity: Tau²=0; Chi²=2.99, df=2(P=0.22); I²=33.08%					
Test for overall effect: Z=0.61(P=0.54)					



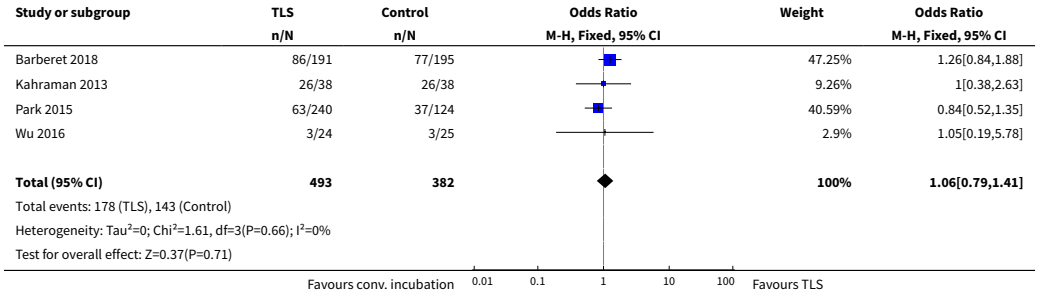
Analysis 1.2. Comparison 1 TLS with conventional morphological assessment of still TLS images versus conventional incubation and assessment (trial design 1), Outcome 2 Miscarriage.



Analysis 1.3. Comparison 1 TLS with conventional morphological assessment of still TLS images versus conventional incubation and assessment (trial design 1), Outcome 3 Stillbirth.



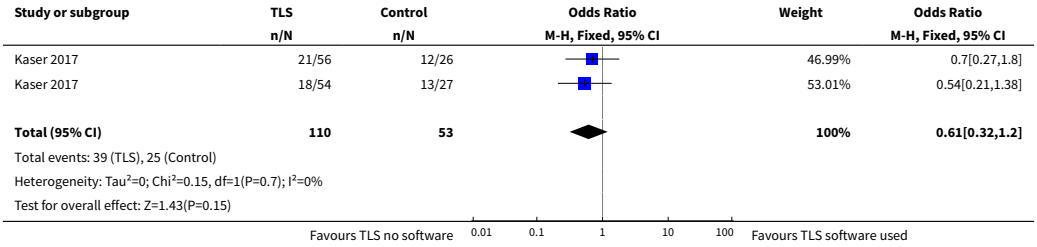
Analysis 1.4. Comparison 1 TLS with conventional morphological assessment of still TLS images versus conventional incubation and assessment (trial design 1), Outcome 4 Clinical pregnancy.



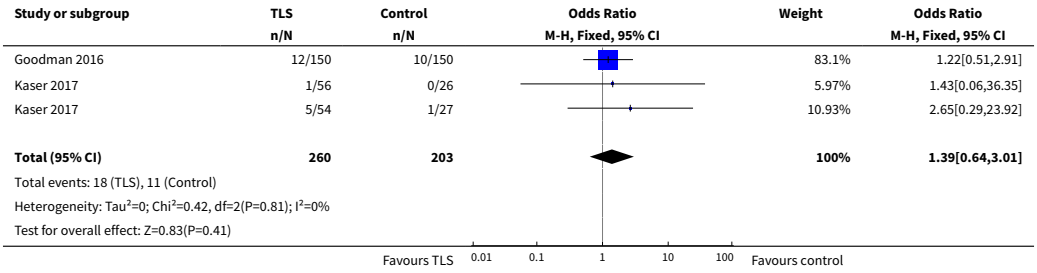
Comparison 2. TLS utilising embryo selection software versus TLS with conventional morphological assessment of still TLS images (trial design 2)

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
1 Live birth or ongoing pregnancy	1	163	Odds Ratio (M-H, Fixed, 95% CI)	0.61 [0.32, 1.20]
2 Miscarriage	2	463	Odds Ratio (M-H, Fixed, 95% CI)	1.39 [0.64, 3.01]
3 Clinical pregnancy	2	463	Odds Ratio (M-H, Fixed, 95% CI)	0.97 [0.67, 1.42]

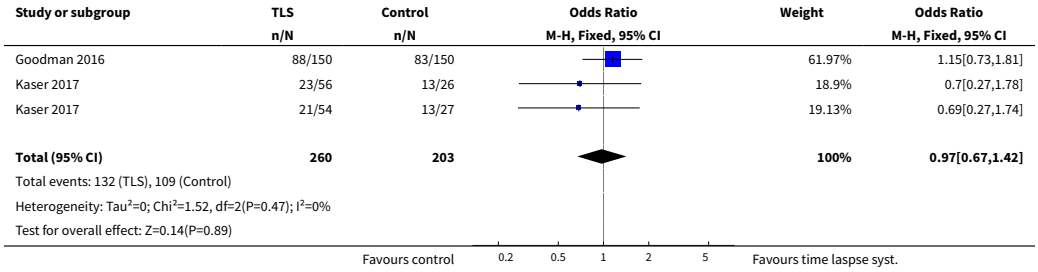
Analysis 2.1. Comparison 2 TLS utilising embryo selection software versus TLS with conventional morphological assessment of still TLS images (trial design 2), Outcome 1 Live birth or ongoing pregnancy.



Analysis 2.2. Comparison 2 TLS utilising embryo selection software versus TLS with conventional morphological assessment of still TLS images (trial design 2), Outcome 2 Miscarriage.



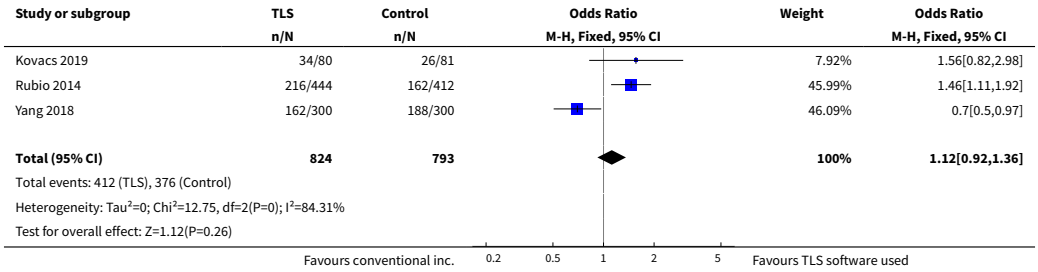
Analysis 2.3. Comparison 2 TLS utilising embryo selection software versus TLS with conventional morphological assessment of still TLS images (trial design 2), Outcome 3 Clinical pregnancy.



Comparison 3. TLS utilising embryo selection software versus conventional incubation and assessment (trial design 3)

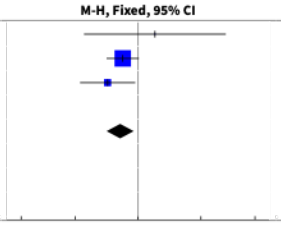
Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
1 Live birth or ongoing pregnancy	3	1617	Odds Ratio (M-H, Fixed, 95% CI)	1.12 [0.92, 1.36]
2 Miscarriage	3	1617	Odds Ratio (M-H, Fixed, 95% CI)	0.63 [0.45, 0.89]
3 Stillbirth	1	600	Odds Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
4 Clinical pregnancy	3	1617	Odds Ratio (M-H, Fixed, 95% CI)	0.95 [0.78, 1.16]

Analysis 3.1. Comparison 3 TLS utilising embryo selection software versus conventional incubation and assessment (trial design 3), Outcome 1 Live birth or ongoing pregnancy.



Analysis 3.2. Comparison 3 TLS utilising embryo selection software versus conventional incubation and assessment (trial design 3), Outcome 2 Miscarriage.

Study or subgroup	TLS n/N	Control n/N	Odds Ratio M-H, Fixed, 95% CI	Weight	Odds Ratio M-H, Fixed, 95% CI
Kovacs 2019	3/80	2/81		2.36%	1.54[0.25,9.47]
Rubio 2014	45/444	59/412		67.98%	0.67[0.45,1.02]
Yang 2018	12/300	25/300		29.66%	0.46[0.23,0.93]
Total (95% CI)	**824**	**793**		**100%**	**0.63[0.45,0.89]**
Total events: 60 (TLS), 86 (Control)					
Heterogeneity: Tau²=0; Chi²=1.81, df=2(P=0.4); I²=0%					
Test for overall effect: Z=2.59(P=0.01)					



Favours TLS — Favours conv. incubation

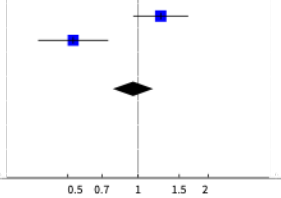
Analysis 3.3. Comparison 3 TLS utilising embryo selection software versus conventional incubation and assessment (trial design 3), Outcome 3 Stillbirth.

Study or subgroup	TLS n/N	Control n/N	Odds Ratio M-H, Fixed, 95% CI	Weight	Odds Ratio M-H, Fixed, 95% CI
Yang 2018	0/300	0/300			Not estimable
Total (95% CI)	**300**	**300**			**Not estimable**
Total events: 0 (TLS), 0 (Control)					
Heterogeneity: Not applicable					
Test for overall effect: Not applicable					

Favours TLS — Favours conv. incubation

Analysis 3.4. Comparison 3 TLS utilising embryo selection software versus conventional incubation and assessment (trial design 3), Outcome 4 Clinical pregnancy.

Study or subgroup	TLS n/N	Control n/N	Odds Ratio M-H, Fixed, 95% CI	Weight	Odds Ratio M-H, Fixed, 95% CI
Kovacs 2019	37/80	28/81		7.62%	1.63[0.86,3.07]
Rubio 2014	272/444	230/412		47.11%	1.25[0.95,1.64]
Yang 2018	180/300	222/300		45.26%	0.53[0.37,0.75]
Total (95% CI)	**824**	**793**		**100%**	**0.95[0.78,1.16]**
Total events: 489 (TLS), 480 (Control)					
Heterogeneity: Tau²=0; Chi²=17.82, df=2(P=0); I²=88.78%					
Test for overall effect: Z=0.48(P=0.63)					



Favours conv. incubation — Favours TLS

CONTRIBUTIONS OF AUTHORS

SA developed the protocol and wrote the first draft of the review. PB, VJ, AP, and CF commented on and made changes to the review. SA, PB, and JM screened the search titles and extracted data from the full-text articles. SA and PB contacted authors for further information. VJ gave her methodological and content opinion on the full review.

Ms Nicola Arroll and Dr Lynsey Cree were both authors of the first iteration of this review, but have not participated in this review update.

DECLARATIONS OF INTEREST

Dr Priya Bhide is a co-investigator for the TILT trial, a randomised controlled trial of time-lapse system versus undisturbed culture versus conventional incubation and assessment, which has recently obtained ethics approval. TILT is funded by the Barts Charity.

There are no other conflicts of interest for any of the review authors.

SOURCES OF SUPPORT

Internal sources

- No sources of support supplied

External sources

- None, Other.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We have altered the review title to reflect both the assessment and culture capability of TLS.

We have altered the wording of the Types of interventions section in the Methods to clarify the comparisons made. We sought to divide studies into the following three comparisons based on the nature of the intervention and the control in order to truly assess if there is a clinical benefit to TLS, and where the benefit of TLS might lie.

TLS with conventional morphological assessment of still TLS images versus conventional incubation and assessment (trial design 1)

TLS utilising embryo selection software versus TLS with conventional morphological assessment of still TLS images (trial design 2)

TLS utilising embryo selection software versus conventional incubation and assessment (trial design 3)

We changed the wording of the outcome 'adverse events' to 'miscarriage and stillbirth'.

We have removed 'alternative imputation strategies' from Sensitivity analysis.

In the 2019 update, we changed the primary outcome to 'live birth or ongoing pregnancy'. The rationale was that there are very few pregnancy losses after 12 weeks' gestation, and the inclusion of the additional data would increase the power of the analysis. We planned

to conduct a sensitivity analysis to investigate the effect of using this composite outcome, however this was not needed because only one study was included in Analysis 2.1.

INDEX TERMS

Medical Subject Headings (MeSH)

*Embryo Culture Techniques; *Reproductive Techniques, Assisted; Embryo Implantation; Embryonic Development [*physiology]; Pregnancy Outcome; Pregnancy Rate; Randomized Controlled Trials as Topic

MeSH check words

Female; Humans; Pregnancy

Paper 3: GM-CSF (granulocyte macrophage colony-stimulating factor) supplementation in culture media for women undergoing assisted reproduction

As outlined in section 6.6, this review was undertaken to provide an up-to-date meta-analysis and appraisal of the evidence on the topic of GM-CSF containing culture media versus culture media not containing GM-CSF. At the time of publication, the only other published systematic review on the subject was from 2012 and was narrative in nature. The primary outcome was livebirth, however only one included study reported on this. The evidence was appraised as being of 'median' quality, but no formal assessment tool for quality was applied and the term median is difficult to interpret in this context.

This Cochrane review represents the first high-quality, systematic review including a meta-analysis on the topic. The review was commissioned in 2019 with the peer reviewed protocol published the same year. The review was conducted between July 2019 and July 2020.

The review is reproduced below in its published format, minus the following sections for the sake of brevity: table of contents; characteristics of studies awaiting assessment; characteristics of ongoing studies; and the appendices, which include the search strategies and the feedback section. These can be read in the full version of the published review [105].

Cochrane Library

Cochrane Database of Systematic Reviews

GM-CSF (granulocyte macrophage colony-stimulating factor) supplementation in culture media for women undergoing assisted reproduction (Review)

Armstrong S, MacKenzie J, Woodward B, Pacey A, Farquhar C

Armstrong S, MacKenzie J, Woodward B, Pacey A, Farquhar C.
GM-CSF (granulocyte macrophage colony-stimulating factor) supplementation in culture media for women undergoing assisted reproduction.
Cochrane Database of Systematic Reviews 2020, Issue 7. Art. No.: CD013497.
DOI: 10.1002/14651858.CD013497.pub2.

www.cochranelibrary.com

[Intervention Review]

GM-CSF (granulocyte macrophage colony-stimulating factor) supplementation in culture media for women undergoing assisted reproduction

Sarah Armstrong[1], Jeanette MacKenzie[2], Bryan Woodward[3], Allan Pacey[1], Cindy Farquhar[4]

[1]Department of Oncology & Metabolism, Academic Unit of Reproductive and Developmental Medicine, The University of Sheffield, Sheffield, UK. [2]Fertility Plus, Women's Health, Auckland District Health Board, Auckland, New Zealand. [3]IVF Consultancy Services, Leicester, UK. [4]Department of Obstetrics and Gynaecology, University of Auckland, Auckland, New Zealand

Contact address: Sarah Armstrong, scarmstrong1@sheffield.ac.uk.

Editorial group: Cochrane Gynaecology and Fertility Group.
Publication status and date: New, published in Issue 7, 2020.

Citation: Armstrong S, MacKenzie J, Woodward B, Pacey A, Farquhar C. GM-CSF (granulocyte macrophage colony-stimulating factor) supplementation in culture media for women undergoing assisted reproduction. *Cochrane Database of Systematic Reviews* 2020, Issue 7. Art. No.: CD013497. DOI: 10.1002/14651858.CD013497.pub2.

ABSTRACT

Background

GM-CSF (granulocyte macrophage colony-stimulating factor) is a growth factor that is used to supplement culture media in an effort to improve clinical outcomes for those undergoing assisted reproduction. It is worth noting that the use of GM-CSF-supplemented culture media often adds a further cost to the price of an in vitro fertilisation (IVF) cycle. The purpose of this review was to assess the available evidence from randomised controlled trials (RCTs) on the effectiveness and safety of GM-CSF-supplemented culture media.

Objectives

To assess the effectiveness and safety of GM-CSF-supplemented human embryo culture media versus culture media not supplemented with GM-CSF, in women or couples undergoing assisted reproduction.

Search methods

We used standard methodology recommended by Cochrane. We searched the Cochrane Gynaecology and Fertility Group Trials Register, CENTRAL, MEDLINE, Embase, CINAHL, LILACS, DARE, OpenGrey, PubMed, Google Scholar, and two trials registers on 15 October 2019, checked references of relevant papers and communicated with experts in the field.

Selection criteria

We included RCTs comparing GM-CSF (including G-CSF (granulocyte colony-stimulating factor))-supplemented embryo culture media versus any other non-GM-CSF-supplemented embryo culture media (control) in women undergoing assisted reproduction.

Data collection and analysis

We used standard methodological procedures recommended by Cochrane. The primary review outcomes were live birth and miscarriage rate. The secondary outcomes were clinical pregnancy, multiple gestation, preterm birth, birth defects, aneuploidy, and stillbirth rates. We assessed the quality of the evidence using GRADE methodology. We undertook one comparison, GM-CSF-supplemented culture media versus culture media not supplemented with GM-CSF, for those undergoing assisted reproduction.

Main results

We included five studies, the data for three of which (1532 participants) were meta-analysed. We are uncertain whether GM-CSF-supplemented culture media makes any difference to the live-birth rate when compared to using conventional culture media not supplemented with GM-CSF (odds ratio (OR) 1.19, 95% confidence interval (CI) 0.93 to 1.52, 2 RCTs, N = 1432, I^2 = 69%, low-quality evidence). The evidence suggests that if the rate of live birth associated with conventional culture media not supplemented with GM-CSF was 22%, the rate with the use of GM-CSF-supplemented culture media would be between 21% and 30%.

We are uncertain whether GM-CSF-supplemented culture media makes any difference to the miscarriage rate when compared to using conventional culture media not supplemented with GM-CSF (OR 0.75, 95% CI 0.41 to 1.36, 2 RCTs, N = 1432, I^2 = 0%, low-quality evidence). This evidence suggests that if the miscarriage rate associated with conventional culture media not supplemented with GM-CSF was 4%, the rate with the use of GM-CSF-supplemented culture media would be between 2% and 5%.

Furthermore, we are uncertain whether GM-CSF-supplemented culture media makes any difference to the following outcomes: clinical pregnancy (OR 1.16, 95% CI 0.93 to 1.45, 3 RCTs, N = 1532 women, I^2 = 67%, low-quality evidence); multiple gestation (OR 1.24, 95% CI 0.73 to 2.10, 2 RCTs, N = 1432, I^2 = 35%, very low-quality evidence); preterm birth (OR 1.20, 95% CI 0.70 to 2.04, 2 RCTs, N = 1432, I^2 = 76%, very low-quality evidence); birth defects (OR 1.33, 95% CI 0.59 to 3.01, I^2 = 0%, 2 RCTs, N = 1432, low-quality evidence); and aneuploidy (OR 0.34, 95% CI 0.03 to 3.26, I^2 = 0%, 2 RCTs, N = 1432, low-quality evidence). We were unable to undertake analysis of stillbirth, as there were no events in either arm of the two studies that assessed this outcome.

Authors' conclusions

Due to the very low to low quality of the evidence, we cannot be certain whether GM-CSF is any more or less effective than culture media not supplemented with GM-CSF for clinical outcomes that reflect effectiveness and safety. It is important that independent information on the available evidence is made accessible to those considering using GM-CSF-supplemented culture media. The claims from marketing information that GM-CSF has a positive effect on pregnancy rates are not supported by the available evidence presented here; further well-designed, properly powered RCTs are needed to lend certainty to the evidence.

PLAIN LANGUAGE SUMMARY

Growth factor-supplemented culture media for women undergoing assisted reproduction

Review question

Does culture media containing the growth factor GM-CSF (granulocyte macrophage colony-stimulating factor) improve the chances of a pregnancy and live-born baby, and reduce the risk of miscarriage, twin or triplet pregnancy, premature birth, birth defects, genetic problems in the baby, and stillbirth?

Background

Assisted reproduction includes processes whereby a woman's eggs and a man's sperm are combined to achieve fertilisation outside of the body. Embryos are placed in a solution called culture medium to support the growing embryo until it can be replaced into the woman's uterus. Culture medium supplemented with GM-CSF is widely available in clinics and is often offered as an 'add-on' to an in vitro fertilisation (IVF) cycle in an effort to improve the success rates of treatment. Using GM-CSF-supplemented culture medium can make IVF more expensive.

Study characteristics

The evidence is current to October 2019. We obtained data from three randomised controlled trials (a type of study in which participants are randomly assigned to one of two or more treatment groups) of 1532 infertile women undergoing IVF or intracytoplasmic sperm injection (ICSI), a specialised form of IVF whereby the sperm is injected into the egg. We compared GM-CSF-supplemented culture media versus culture media not supplemented with GM-CSF for those undergoing assisted reproduction.

What the review found

Low-quality evidence reveals that we are uncertain whether GM-CSF-containing culture media makes any difference to the live-birth rate when compared to using culture media not containing GM-CSF. This suggests that if the rate of live birth associated with culture media not containing GM-CSF is 22%, the rate with the use of GM-CSF-containing culture media would be between 21% and 30%. Low-quality evidence also reveals that we are uncertain whether GM-CSF-containing culture media makes any difference to miscarriage when compared to using culture media not containing GM-CSF. This suggests that if the miscarriage rate associated with culture media not containing GM-CSF is 4%, the rate with the use of GM-CSF-containing culture media would be between 2% and 5%. Low-quality evidence for pregnancy, birth defects, and genetic problems with the baby, and very low-quality evidence for twin or triplet pregnancies, and premature birth, reveals that we are uncertain whether GM-CSF-containing culture media makes any difference to these outcomes when compared to culture media not containing GM-CSF. Two studies looked at stillbirth, but as no stillbirths occurred in either study, we were unable to analyse this outcome.

Overall conclusions

Due to the very low to low quality of the evidence, we cannot be certain whether GM-CSF is any more or less effective or harmful than culture media not supplemented with GM-CSF. It is important that independent information on the available evidence is made accessible to those considering using GM-CSF-supplemented culture media. In the meantime, more large studies are needed to increase the certainty of our conclusions.

SUMMARY OF FINDINGS

Summary of findings 1. GM-CSF-supplemented culture media compared to culture media not supplemented with GM-CSF for women undergoing assisted reproduction

GM-CSF-supplemented culture media compared to culture media not supplemented with GM-CSF for women undergoing assisted reproduction

Patient or population: women undergoing assisted reproduction
Setting: fertility clinics
Intervention: GM-CSF-supplemented culture media
Comparison: culture medium not supplemented with GM-CSF

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)
	Risk with culture media not supplemented with GM-CSF	Risk with GM-CSF-supplemented culture media			
Live birth or ongoing pregnancy	Study population		OR 1.19 (0.93 to 1.52)	1432 (2 RCTs)	⊕⊕⊙⊙ LOW [1 2]
	223 per 1000	254 per 1000 (210 to 303)			
Miscarriage	Study population		OR 0.75 (0.41 to 1.36)	1432 (2 RCTs)	⊕⊕⊙⊙ LOW [3]
	36 per 1000	27 per 1000 (15 to 48)			
Clinical pregnancy	Study population		OR 1.16 (0.93 to 1.45)	1532 (3 RCTs)	⊕⊕⊙⊙ LOW [1 4]
	263 per 1000	293 per 1000 (250 to 342)			
Multiple gestation	Study population		OR 1.24 (0.73 to 2.10)	1432 (2 RCTs)	⊕⊙⊙⊙ VERY LOW [1 3]
	36 per 1000	44 per 1000 (26 to 72)			
Preterm birth	Study population		OR 1.20 (0.70 to 2.04)	1432 (2 RCTs)	⊕⊙⊙⊙ VERY LOW [1 3]
	36 per 1000	43 per 1000 (25 to 70)			
Birth defects	Study population		OR 1.33 (0.59 to 3.01)	1432 (2 RCTs)	⊕⊕⊙⊙ LOW[3]

	14 per 1000	18 per 1000 (8 to 40)			
Aneuploidy	Study population		OR 0.34 (0.03 to 3.26)	1432 (2 RCTs)	⊕⊕⊝⊝ LOW [5]
	3 per 1000	1 per 1000 (0 to 9)			
Stillbirth	Study population		-	1432 (2 RCTs)	-
	See comment[6]	See comment[6]			

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

CI: confidence interval; **OR:** odds ratio; **RCT:** randomised controlled trial

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate quality: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low quality: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low quality: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[1]Downgraded once for inconsistency, as the included studies report differing directions of point estimates: one supports the intervention, and one does not support the intervention.
[2]Downgraded once for imprecision as broad confidence intervals and a low number of included studies, at least one of which is very small.
[3]Downgraded twice for imprecision as very broad confidence intervals and a low number of included studies.
[4]Downgraded once for risk of bias. One included study had an unclear risk of selection bias, performance bias, and detection bias due to limited information available from published abstract.
[5]Downgraded twice for imprecision as included studies had so few reported incidences of aneuploidy that the point estimate is not precise and has very broad confidence intervals.
[6]No stillbirths occurred in either arm of the included studies, therefore the result is inestimable.

BACKGROUND

Description of the condition

Assisted reproduction provides the opportunity to have a family for those unable to become pregnant spontaneously for a variety of reasons, including; infertility; those in single-sex relationships; single women; and those using surrogates. Assisted reproduction is often referred to as a 'cycle', reflecting its stepwise process. It involves a series of procedures from ovarian stimulation and oocyte collection, to mixing the gametes, culturing and assessing the quality of ensuing embryos, and replacing embryos into the uterus of the woman. The success of assisted reproduction is a culmination of all the elements of the cycle, and is in part due to the ability to culture human embryos in vitro using culture media capable of supporting the developing embryo. GM-CSF (granulocyte macrophage colony-stimulating factor)-supplemented culture media was developed in an effort to improve this particular part of the cycle, leading to better outcomes for those undergoing in vitro fertilisation (IVF).

GM-CSF-supplemented culture media can be described as an assisted reproduction 'add-on'. Add-ons are optional extras to an assisted reproduction cycle, which are sometimes novel interventions or therapies that have shown some promise in initial studies, or have been around for many years, but have not yet been proven to be effective through randomised controlled trials (RCTs). GM-CSF-supplemented culture media is one such add-on, often provided at an additional cost to the IVF cycle (Heneghan 2016).

For the purposes of this review, any culture media containing GM-CSF may be compared in a randomised controlled trial (RCT) against any culture media not containing GM-CSF. We addressed the efficacy and safety of GM-CSF-supplemented culture media when compared to culture media not containing GM-CSF. The primary outcomes were live birth and miscarriage.

Description of the intervention

GM-CSF (also known as colony-stimulating factor (CSF)-2) and granulocyte colony-stimulating factor (G-CSF or CSF-3) belong to the CSF family. They are a group of cytokines known for their role in haemopoietic cell proliferation, differentiation, and activation, as well as being an apoptosis suppressor (Rahmati 2015). Their involvement in reproduction was initially investigated in the 1970s in human placenta-conditioned media (Burgess 1977). Amongst the CSF group, GM-CSF is the most widely studied, and its extensive research on assisted reproduction has led to the development of new embryo culture media supplemented with human recombinant GM-CSF. EmbryoGen and BlastGen are examples of commercially available sequential culture media containing GM-CSF.

GM-CSF is a cytokine that is produced by the oestrogen-primed epithelial cells in the female reproductive tract (Robertson 1992). It is maximally expressed at the luminal and glandular epithelial cells of the endometrium in the secretory phase, and in the lining of the fallopian tube during the late proliferative and early mid-secretory phases of the menstrual cycle (Giacomini 1995; Zhao 1994). Later during implantation, GM-CSF is produced by the chorionic villi cells and the maternal decidua (Jokhi 1994). In response to local inflammatory stimuli, GM-CSF acts by stimulating and activating mature monocytes, granulocytes, macrophages, and dendritic cells which promote chemotactic, phagocytic, and cytotoxic actions as well as antigen-presenting properties needed in the immunomodulation of early pregnancy and embryogenesis (Baldwin 1992; Robertson 2007).

How the intervention might work

The control of the immunological environment during early pregnancy involves a series of autocrine and paracrine signalling between the maternal fetal interface (Robertson 1994; Robertson 2007; Wegmann 1992). Several studies have suggested an association between recurrent pregnancy loss and infertility and the dysregulation of growth factors and cytokines (Hambartsoumian 1998; Torry 2007; Vuorela 2000). In studies of genetically GM-CSF-deficient mice, there was a reduced inner cell mass observed which resulted in delayed blastocyst formation, increased fetal resorption in late gestation, decreased fetal size, and greater postnatal mortality (Robertson 1999). Other murine studies have also supported that GM-CSF is crucial in optimal fetal growth and survival, as animal models lacking GM-CSF expression experience more pregnancy losses and impaired long-term survival of the newborn animals (Savion 2002; Seymour 1997).

The initial studies of growth factor supplementation of culture media are limited mostly to animal models, but have largely revealed improved blastocyst development rates, Lighten 1998; Sjöblom 1992; Sjöblom 1999; Spanos 2000; Yu 2012, and increased implantation and birth rates (Block 2003; Lim 2006; Roudebush 2004; Sjöblom 2005). The use of growth factor supplementation in human culture media has been limited, as it is costly to produce, and there are concerns about adverse effects (Richter 2008). Most growth factors are anti-apoptotic, that is they inhibit programmed cell death. If not controlled, adverse effects may occur, as apoptosis is a crucial phenomenon in embryogenesis. Inhibition of apoptosis may lead to abnormal embryo development such as the well-documented 'large offspring syndrome' that occurs in mice models (Lazzari 2002; Young 2001).

Early studies on human embryos have revealed that those cultured in GM-CSF-supplemented culture media had more viable inner cell masses and reduced apoptosis. This could potentially contribute to improved fetal viability (Sjöblom 1999; Sjöblom 2002). Supplementation of culture media with GM-CSF is reported to be safe for human embryos; there are no increases or changes in ploidy rates or embryonic chromosomes (Agerholm 2010). Furthermore, initial RCTs in women revealed an improvement in the clinical pregnancy and live-birth rates of those women randomised to culture of their embryos in GM-CSF-supplemented culture media (Mignini 2013; Sfontouris 2013; Tevkin 2014; Ziebe 2013). There were no major and minor birth abnormalities (Mignini 2013; Sfontouris 2013; Tevkin 2014).

Why it is important to do this review

GM-CSF-supplemented culture media is widely commercially available and is offered to women undergoing assisted reproduction worldwide. It is often considered an 'add-on', or supplementary therapy, given alongside standard IVF in an effort to improve success rates (Heneghan 2016). There is currently no up-to-date systematic review of RCTs on this topic, and the one published systematic review relied on non-randomised studies and studies where oocytes rather than women were randomised (Siristatidis 2013). The available RCTs were small with differing

results and did not provide certainty as to what should be done in practice. Use of GM-CSF can carry an additional cost to women undergoing IVF. It was therefore important to distil the available RCT evidence in a meaningful way to provide information on the effectiveness and safety of this intervention for women, clinicians, and embryologists, and regulatory and advisory bodies such as the Human Fertilisation and Embryology Authority (HFEA).

OBJECTIVES

To assess the effectiveness and safety of GM-CSF-supplemented human embryo culture media versus culture media not supplemented with GM-CSF, in women or couples undergoing assisted reproduction.

METHODS

Criteria for considering studies for this review

Types of studies

We included all published and unpublished RCTs. We included cross-over studies for completeness, but only pooled data from the first phase in meta-analyses because this study design is not valid in the context of infertility trials (Vail 2003). We excluded quasi- and pseudo-randomised trials. There was no limitation on language, publication date, or publication status.

Types of participants

Women undergoing IVF or intracytoplasmic sperm injection (ICSI) for any cause of infertility, using autologous or donor oocytes. Women undergoing IVF or ICSI with a background of recurrent miscarriage or recurrent implantation failure were also included.

Types of interventions

We included all studies that compared GM-CSF (including G-CSF)-supplemented embryo culture media versus any other non-GM-CSF-supplemented embryo culture media (control).

Types of outcome measures

Primary outcomes

1. Live birth per woman randomised, defined as a live baby born after 20 weeks' gestation. We used ongoing pregnancy, defined as clinical pregnancy of 12 or more weeks' gestation, as a surrogate for live birth in cases where studies did not report live birth.
2. Miscarriage per woman randomised. The definition used was miscarriage of clinical pregnancy.

Secondary outcomes

1. Clinical pregnancy per woman randomised, defined as presence on ultrasound scan of one or more gestational sacs, or definitive signs of clinical pregnancy. This included ectopic pregnancy. Note that multiple gestational sacs were counted as one clinical pregnancy.
2. Multiple gestation per woman randomised.
3. Preterm birth per woman randomised (defined as birth before 37 weeks' gestation).
4. Birth defects (defined as any structural anomaly present at birth that may interfere with function depending upon the organ or structure involved).
5. Aneuploidy (defined as any genetic disorder diagnosed during pregnancy or at the time of birth).
6. Stillbirth (defined as a baby born with no signs of life after 20 completed weeks of pregnancy).

Search methods for identification of studies

We searched for relevant studies with no language or date restriction in consultation with the Cochrane Gynaecology and Fertility Group Information Specialist.

Electronic searches

We designed search strategies for the following databases:

1. Cochrane Gynaecology and Fertility Group Specialised Register of Controlled Trials; ProCite platform, searched 15 October 2019 (Appendix 1);
2. Cochrane Central Register of Controlled Trials (CENTRAL); Ovid platform, searched 15 October 2019 (Issue 9; 2019) (Appendix 2);
3. MEDLINE; Ovid platform, searched from 1946 to 15 October 2019 (Appendix 3);
4. Embase; Ovid platform, searched from 1980 to 15 October 2019 (Appendix 4);
5. CINAHL (Cumulative Index to Nursing and Allied Health Literature), Ebsco platform, searched from 1961 to 15 October 2019 (Appendix 5);
6. LILACS (Latin American and Caribbean Health Science Information database) (lilacs.bvsalud.org/en/), Web platform, searched 15 October 2019 (Appendix 6).

The MEDLINE search was combined with the Cochrane Highly Sensitive Search Strategy for identifying randomised trials in Section 4.3.1 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2019). The Embase and CINAHL search strategies are combined with trial filters developed by the Scottish Intercollegiate Guidelines Network (SIGN; www.sign.ac.uk/methodology/filters.html#random).

Other electronic sources of trials (Web platforms, searched 15 October 2019) included:

1. trial registers for ongoing and registered trials: US National Institutes of Health Ongoing Trials Register ClinicalTrials.gov (www.clinicaltrials.gov) and the World Health Organization International Clinical Trials Registry Platform (WHO ICTRP) (apps.who.int/trialsearch/Default.aspx) (Appendix 7);
2. DARE (Database of Abstracts of Reviews of Effects) on the Cochrane Library (onlinelibrary.wiley.com/o/cochrane/cochrane_cldare_articles_fs.htm) (Appendix 8);
3. Web of Knowledge (wokinfo.com) (Appendix 9);
4. OpenGrey (www.opengrey.eu/) for unpublished literature from Europe (Appendix 10);
5. PubMed and Google Scholar (for recent trials not yet indexed in the major databases) (Appendix 11 and Appendix 12).

Searching other resources

We handsearched reference lists of included and excluded studies retrieved by the search, and communicated with experts in the field to inquire after any additional studies.

We did not perform a separate search for adverse effects of GM-CSF-supplemented culture media. We considered adverse effects described in the studies only.

Data collection and analysis

Selection of studies

All review authors independently undertook assessment of eligibility of all studies identified by the search using Covidence (Covidence). We retrieved the full-text publications of potentially eligible studies. Three review authors (SA, JM, and AP) screened the full texts to identify studies for inclusion, and recorded reasons for exclusion of the excluded studies in the 'Characteristics of excluded studies' table. Any disagreements were resolved by discussion or consultation with another review author.

Data extraction and management

Two review authors (SA and JM) independently extracted data on study characteristics and primary and secondary outcomes from the included studies using a data extraction form designed and piloted by the review authors. We included the following characteristics of included studies in the data extraction form:

1. methods;
2. participants;
3. interventions;
4. outcomes, including adverse events;
5. funding source for studies.

Any disagreements or discrepancies were resolved by discussion. Where there were multiple publications for a study, we used the main trial report as the reference and obtained additional details from secondary papers, which appear as subreferences. We corresponded with study investigators for further information on study methods and results as required. This correspondence is documented in the 'Characteristics of included studies' table and in Appendix 13.

Assessment of risk of bias in included studies

Two review authors (SA and JM) independently assessed the included studies for methodological quality and undertook data extraction according to the Cochrane 'Risk of bias' assessment tool (Higgins 2011). We assessed selection bias (random sequence generation and allocation concealment), attrition bias (incomplete outcome data), reporting bias (selective reporting), performance bias (blinding of participants and personnel), detection bias (blinding of outcome assessors), and other biases (other problems that could put a trial at high risk of bias). Our judgements are presented and described in the 'Risk of bias' table in Characteristics of included studies. Any disagreements were resolved by discussion.

Measures of treatment effect

We summarised the effects and adverse events related to the intervention as odds ratios (ORs) using a fixed-effect model. We presented 95% confidence intervals (CIs) for all outcomes to evaluate the precision of the estimate. We considered the clinical relevance of the results from the meta-analysis of each comparison, taking into account the precision of the estimate. When adding data from individual studies to comparisons, we considered whether the rates of events in both the intervention and control arm reflect current practice. For example, we explored major discrepancies in direction and magnitude of effect in the Results section, and these are reflected in our 'Risk of bias' assessment.

Unit of analysis issues

The denominator for all outcomes was the number of women randomised. We did not use per-cycle data.

We counted multiple births (e.g. twins or triplets) as one live-birth event.

Dealing with missing data

We analysed the data on an intention-to-treat basis and attempted to obtain missing data from the primary investigators (Appendix 13). We assumed that participants who dropped out after randomisation (e.g. because of cycle cancellation), or who were lost to follow-up or withdrew, did not achieve clinical pregnancy or live birth. We made no other assumptions.

Assessment of heterogeneity

We considered whether the clinical and methodological characteristics of the included studies were sufficiently similar for meta-analysis to provide a clinically meaningful summary. We assessed statistical heterogeneity using the I^2 statistic, considering an I^2 statistic greater than 50% to indicate substantial heterogeneity (Higgins 2019). Where there was significant heterogeneity, we undertook planned subgroup analyses to explore this in more detail.

Assessment of reporting biases

We reduced the potential impact of publication and reporting bias by performing a comprehensive search for eligible studies and looking for duplication of data. We decided to construct a funnel plot to explore the possibility of small-study effects (a tendency for estimates of the intervention effect to be more beneficial in smaller studies) if there were 10 or more studies included in an analysis. When possible, we used published protocols and prospective trial registration web pages for included studies to investigate selective reporting (i.e. comparisons of outcomes listed in the study protocol versus outcomes reported in papers).

Data synthesis

We performed meta-analyses where data were available from multiple studies investigating the same treatment, and the outcome was measured in a standard way between the studies. We used a fixed-effect model. We undertook meta-analysis according to the methods recommended in the *Cochrane Handbook for Systematic Reviews of Interventions* for the following comparison (Higgins 2019).

1. Studies that include GM-CSF supplementation in human embryo culture media versus any other non-GM-CSF-supplemented human embryo culture media.

Subgroup analysis and investigation of heterogeneity

We conducted subgroup analyses for all outcomes when data were available to determine the separate effect between the following subgroups.

1. Studies including only women with recurrent implantation failure, defined as the failure to achieve a clinical pregnancy after transfer of at least four good-quality embryos in a minimum of three fresh or frozen cycles (Coughlan 2014), versus studies not including women with recurrent miscarriage.
2. Studies using single-step culture media versus studies using sequential culture media.
3. Studies including only women with donor oocytes versus studies using autologous oocytes.
4. Studies including only women with recurrent miscarriage (loss of three or more consecutive pregnancies before 20 weeks' gestation) versus studies not including women with recurrent miscarriage.
5. Studies replacing embryos at cleavage stage (day 2 or 3) versus studies replacing embryos at blastocyst stage (day 5).

Sensitivity analysis

We conducted sensitivity analyses for the primary outcomes to determine whether the conclusions were robust to arbitrary decisions made regarding eligibility and analysis. These analyses included consideration of whether the review conclusions would have differed if:

1. eligibility was restricted to studies without high risk of bias (we defined low risk of bias studies as those with low risk of bias in at least the following two domains: random sequence generation and allocation concealment);
2. a random-effects model had been adopted;
3. the summary effect measure was risk ratio rather than OR.

Overall quality of the body of evidence: 'Summary of findings' table

We prepared a 'Summary of findings' table to evaluate the overall quality of the body of evidence for the main review outcomes (live birth, miscarriage, clinical pregnancy, multiple gestation, preterm birth, birth defects, aneuploidy, stillbirth) using GRADE criteria (study limitations (i.e. risk of bias), consistency of effect, imprecision, indirectness, and publication bias) (Summary of findings 1) (GRADEpro GDT). We justified and documented judgements about the quality of the evidence (high, moderate, low, and very low) and incorporated this information into the reporting of the results for each outcome. The 'Summary of findings' table compared GM-CSF-supplemented embryo culture media versus any other non-GM-CSF-supplemented embryo culture media (control).

RESULTS

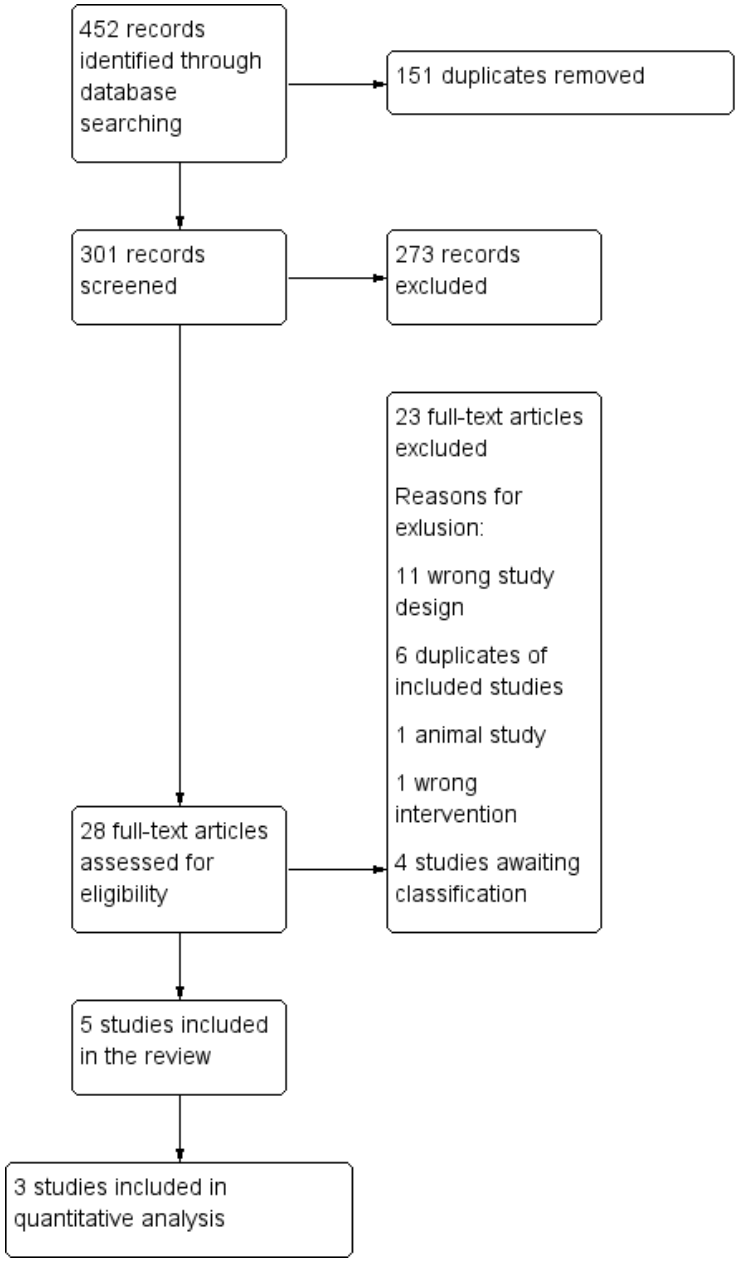
Description of studies

Results of the search

The following databases were systematically searched by Marian Showell, the Information Specialist at Cochrane Gynaecology and Fertility, on 15 October 2019: Cochrane Gynaecology and Fertility Specialised Register, CENTRAL, MEDLINE, Embase, and CINAHL. In addition, the 2019 European Society of Human Reproduction and Embryology (ESHRE) and American Society for Reproductive Medicine (ASRM) conference abstracts were handsearched by review author SA on 28 October 2019. In addition, a Google search using the terms 'GM-CSF, culture media, RCT, and live birth' was undertaken on 28 October 2019.

The search returned 452 records, 151 of which were duplicates. This left 301 titles and abstracts for screening, which was undertaken by all co-authors using the online software Covidence (Covidence). Each record was screened by two review authors at every stage. We considered 28 papers to be eligible for full-text screening. We excluded 23 full texts for the following reasons: 11 were the wrong study design; seven were trial registry information only, without data; three were duplicates of included studies; one was an animal study; and one was the wrong intervention. We considered five studies to be eligible for inclusion in the review (Rose 2020; Sbracia 2014; Zafardoust 2017; Zavvar 2016; Ziebe 2013), of which three could be used in meta-analysis (Rose 2020; Sbracia 2014; Ziebe 2013). The two studies that were not included in meta-analysis were conference abstracts, and the data could not be extracted reliably without further information from the study authors (Zafardoust 2017; Zavvar 2016); unfortunately we were unable to contact the authors of these studies to obtain the needed clarification. The PRISMA figure illustrates the flow of studies through the review (Figure 1).

Figure 1. PRISMA study flow diagram.



Included studies

Five studies were eligible for inclusion in the review. Two of these were conference abstracts that could not be included in meta-analysis because data could not be reliably extracted (Zafardoust 2017; Zavvar 2016). The remaining three studies included two fully published and peer-reviewed papers, Rose 2020; Ziebe 2013, and one conference abstract (Sbracia 2014).

The largest study was undertaken in Europe; the trial was co-ordinated from the Netherlands, and participants were recruited from 14 fertility clinics in Sweden and Denmark (Ziebe 2013). Ziebe 2013 included a total of 1332 participants, of whom 654 were randomised to the intervention arm and 678 were randomised to the control arm. The study was sponsored, co-ordinated, and authored by the worldwide market-leading manufacturer of GM-CSF-supplemented culture media. Women in the intervention arm had all of their embryos cultured in GM-CSF-supplemented culture medium at a concentration of 2 ng/mL from fertilisation through to embryo transfer. Women randomised to the control arm of the study had all of their embryos cultured in an IVF culture medium that did not contain GM-CSF from fertilisation through to embryo transfer. Both IVF and ICSI were undertaken, and a maximum of two embryos were transferred on day 3 in a fresh embryo transfer cycle. The inclusion criteria for the study were as follows: women aged 25 to 39 years, women who had a regular menstrual cycle of 21 to 35 days, women treated with a standard gonadotropin-releasing hormone (GnRH) agonist or antagonist protocol, and women with three or more follicles with a diameter of 14 mm on the day of human chorionic gonadotropin (hCG) administration, including a leading follicle of 17 mm. The exclusion criteria were: previous participation in the study; use of assisted hatching; use of non-ejaculated sperm; medical conditions or genetic disorders prohibiting IVF/ICSI or interfering with the interpretation of results; use of investigational drugs within 30 days before oocyte retrieval; severe chronic disease of relevance for reproduction; and oocyte donation.

Rose 2020 was a smaller, single-centre study undertaken in a fertility clinic in Australia. A total of 100 women were randomised, 50 to the intervention arm and 50 to the control arm of the study. There were no dropouts. Rose 2020 was a cross-over RCT, but the published data were from the first phase of the trial prior to cross-over. The study was sponsored by the worldwide market-leading manufacturer of GM-CSF-supplemented culture media. The same company also funded two co-authors of the study for statistical support. The women in this study underwent fresh embryo transfer following IVF or ICSI. The women in the intervention arm had all of their embryos cultured in GM-CSF-supplemented culture medium from fertilisation through to embryo transfer on day 5. The concentration of GM-CSF in the intervention culture medium was 2 ng/mL, and the medium was changed on day 3 following observation, scoring, and washing, to the next phase of sequential fresh culture medium with the same concentration of GM-CSF. The control culture medium did not contain GM-CSF, and similarly, day 3 embryos were observed, scored, washed and then transferred to a fresh sequential culture medium. All trial participants had a day 5 embryo transfer, apart from one woman in the control arm and two in the intervention arm who underwent day 3 embryo transfer. Participants underwent single-embryo transfer, except four women in the control arm and six women in the intervention arm, who underwent double-embryo transfer.

The inclusion criteria were: patients must have previously had consecutive transfer of two or more embryos without a positive pregnancy outcome OR have had a history of at least one previous pregnancy loss OR a previous history of poor embryo development (< 20% of embryos developing on the time at day 3 or no blastocysts above grade 2 on day 5). Other additional inclusion parameters included a maternal age between 25 and 41 years, the use of a standard GnRH agonist or antagonist protocol, and three or more follicles of > 14 mm as seen by transvaginal ultrasound before the day of hCG administration. Exclusion criteria included: a need for surgical sperm retrieval (except in cases of previous vasectomy), the use of another investigational drug within 30 days of oocyte retrieval, and/or the presence of a severe chronic disease that could impact the IVF cycle or reproductive outcomes.

Sbracia 2014 was another small, single-centre RCT, undertaken in a fertility clinic in Italy. The study was written as an abstract for an international conference. The authors reported that there was no funding for the study. A total of 100 women were randomised, 50 to the intervention arm and 50 to the control arm of the study. The women in the intervention arm had all of their embryos cultured in GM-CSF-supplemented culture media at a concentration of 2 ng/mL from fertilisation through to embryo transfer. Women in the control arm of the study had all of their embryos cultured in a medium not containing GM-CSF from fertilisation to embryo transfer. The brand name of the control culture medium was not disclosed in the paper. Fresh embryo transfer of up to a maximum of three embryos following ICSI was undertaken in all cycles in both the intervention and control arms of the study. The inclusion criteria were: women with recurrent implantation failure, three or more consecutive failed IVF cycles with a total of at least 8 good embryos replaced in the uterus, and women aged 40 or less. The exclusion criteria were: women aged over 40, chromosomal defects in the couple, metabolic diseases (diabetes, etc.), and other genetic diseases (thalassaemia, cystic fibrosis, etc.).

Both Zafardoust 2017 and Zavvar 2016 were eligible for inclusion, however data could not be reliably extracted for meta-analysis. Zafardoust 2017 was a conference abstract which outlined that it was a single-centre RCT undertaken at a fertility clinic in Iran. The study included couples undergoing frozen embryo transfer following an ICSI cycle with their own gametes. Couples were randomised to either have their frozen embryos thawed and cultured in a test medium containing 2 ng/mL of GM-CSF, or a control medium not containing GM-CSF. Couples were eligible for inclusion in the study if the female partner was < 40 years old, had at least four good-quality embryos after thawing (grade A), and had not had more than one previous embryo transfer. Couples were excluded from entering the study if they needed ICSI cycles requiring pre-implantation genetic diagnosis, if the female partner had an anatomic disorder of the uterus, or one or more hydrosalpinges. The abstract outlines that 90 women were randomised, and 10 were excluded from the final analysis due to various reasons, however the original numbers of women randomised to each group are not disclosed. The outcome of interest reported by the study was clinical pregnancy, which is reported as two percentages, alongside a P value. However, it was not clear which percentage belonged to which arm of the study, therefore it was impossible to extract any meaningful data for meta-analysis. Review author SA attempted to contact two of the authors of this study by email on three separate occasions for clarification of these issues, but unfortunately no response was forthcoming.

Zavvar 2016 was also a conference abstract, which outlined a single-centre RCT undertaken in a fertility clinic in Iran. Zavvar 2016 sought to compare the outcomes of women undergoing ICSI who were randomised to receive an embryo culture medium containing 2 ng/mL GM-CSF or to a culture medium not containing GM-CSF. The inclusion criterion was women who produced only immature oocytes in spite of stimulation with gonadotropins. No exclusion criteria were described. The day of embryo transfer and length of time embryos were exposed to the intervention or control media were not described. The outcome of interest, the clinical pregnancy rate, was reported as percentages alongside a P value. However, it was not possible to identify which result was associated with which arm of the study, therefore we could not reliably include data from this study. Review author SA attempted to contact the authors of this study on two separate occasions by email, but unfortunately no response was received.

Excluded studies

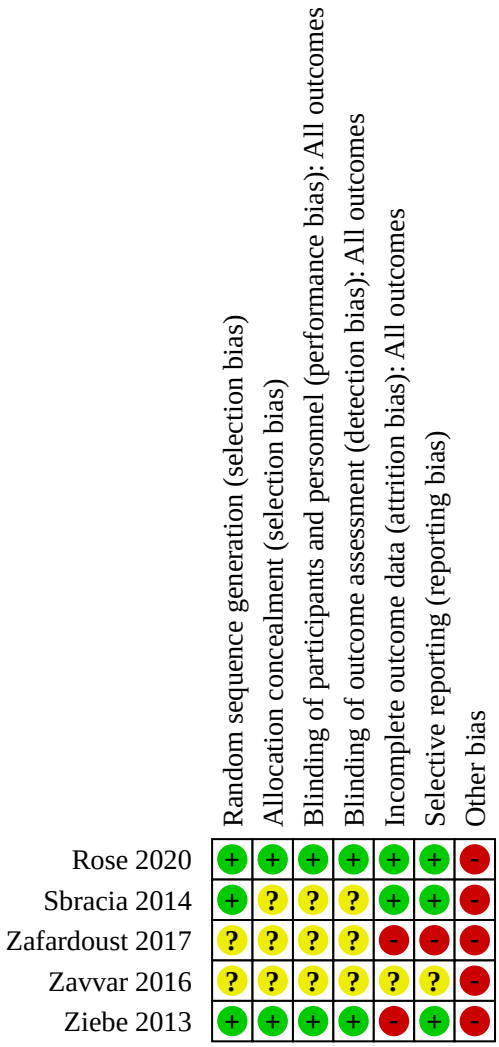
We excluded 23 studies following full-text screening. Eleven studies were the wrong study design, and were excluded for the following reasons: Agerholm 2010 was a phase I safety study and did not replace embryos; Fawzy 2019, Shapiro 2003, Sjoblom 1998, Sjoblom 1999, Sjoblom 1999a, and Sjoblom 2001 randomised oocytes opposed to women or couples; Kinoshita 2019 was a retrospective study; Min 2017 and Sfontouris 2013 were observational studies; and Siristatidis 2013 was a systematic review. Scarpellini 2011 was excluded because it did not study the intervention we were interested in, and Siqueira 2016 was excluded because it was an animal study. Six studies were duplicate references of included studies: Rose 2020, Sbracia 2014, Zafardoust 2017, and Zavvar 2016. Four studies are awaiting classification because the nature of the study and whether women or oocytes was randomised was unclear (ISRCTN94726536; NCT01689428; NCT01689454; NCT02651285).

Risk of bias in included studies

Allocation

We assessed the risk of selection bias for the three studies included in meta-analysis to be low (Figure 2) (Rose 2020; Sbracia 2014; Ziebe 2013), as random sequence generation was described in detail and considered to be adequate to achieve a truly random sequence. Rose 2020 described how 50 cards with 'control' and 50 with 'BlastGen' written on them were placed in sealed envelopes by a person unrelated to the trial. They were shuffled several times, and the envelopes were then numbered and opened in consecutive order by the embryologist when an eligible participant was scheduled for egg retrieval. Sbracia 2014 described how participants were randomised using a computer-generated number sequence; however, allocation concealment was not described, therefore we deemed this study to be at unclear risk. Ziebe 2013 described how they used a computer-generated randomisation list in blocks of four for each individual clinic in order to maintain balance between the treatment groups at each site. Allocation concealment was described in detail and considered to be at low risk on the basis that each study site received a list of study-specific consecutive patient ID numbers and a corresponding number of identical-looking randomised bottles of test and control media that were individually labelled with the corresponding study-specific ID numbers. On site, the lowest number on the list was always allocated to any new patient recruited at the time of informed consent signature. Consequently, the clinician, embryologist, and participant were all blinded to the allocation.

Figure 2. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.



Zafardoust 2017 and Zavvar 2016 could not be included in meta-analysis, but were considered to be at unclear risk of selection bias because no description of randomisation or allocation concealment was provided.

Blinding

We considered Rose 2020 and Ziebe 2013 to be at low risk for performance and detection bias. Rose 2020 described how clinicians, sonographers, statisticians, and participants were completely blinded to the intervention. Ziebe 2013 described how participants and investigators, including clinicians and embryologists, were blinded to treatment allocation. Following email correspondence, we established that the clinicians performing the ultrasound scans were blinded to the treatment allocation at all times.

Sbracia 2014 did not provide any description of blinding of participants, personnel, or outcome assessors, and was therefore assessed as at unclear risk of bias of performance and detection bias. We judged Zafardoust 2017 and Zavvar 2016 to be at unclear risk of performance and detection bias because there was no description of who, if anyone, was blinded.

Incomplete outcome data

We assessed Rose 2020 and Sbracia 2014 as at low risk of attrition bias. Both studies reported no dropouts. We considered Ziebe 2013 to be at high risk of attrition bias because despite all dropouts being accounted for, the reasons given were not included within the predefined exclusion criteria. For example, no oocytes retrieved, no semen sample, no fertilisation, no embryo transfer, and "non-includable after randomisation" were given as reasons for exclusion after randomisation, however none of these were listed as exclusion criteria. We contacted the authors to obtain accurate intention-to-treat (ITT) data for both arms of the study, which they were able to provide.

We considered Zafardoust 2017 to be at high risk of attrition bias because 10 women were not included in the final analysis, with no reasons provided. We considered this to be a high rate of attrition in a small study. Unfortunately we were unable to use data from this study as the data could not be reliably extracted. We considered Zavvar 2016 to be at unclear risk of attrition bias as dropouts were not described. We could not include data from this study because it was unclear how many participants were included in the analysis.

Selective reporting

We rated Rose 2020, Sbracia 2014, and Ziebe 2013 as being at low risk of reporting bias because the study authors confirmed via email correspondence that they had reported all outcomes as per their prospective clinical trials registrations (NCT02305420, NCT01718210, and NCT00565747, respectively).

We rated Zafardoust 2017 as being at high risk of reporting bias. The available abstract did not report data on miscarriage, multiple pregnancy, and beta human chorionic gonadotropin (BHCG) levels, which are secondary outcomes noted on the prospective clinical trials register (Zafardoust 2017). We attempted to contact the study authors to establish if further trial data were available, but received no response. We rated Zavvar 2016 as being at unclear risk of reporting bias. We had no access to a protocol or an online clinical trial registry.

Other potential sources of bias

We have been in extensive contact with the authors of Ziebe 2013 via email to clarify various numbers from their published study. The co-authors of this study have been very forthcoming in answering all of our queries and have offered clear explanations of how various numbers are reached in their paper. However, we have assessed Ziebe 2013 as at high risk of bias for this domain because the numbers published in the paper differ from those published in this review, that is we discovered through correspondence that some participants were inaccurately described as miscarriages opposed to biochemical pregnancy losses. We also asked for individual participant data in relation to those babies that suffered aneuploidy or birth defects, or both. On reviewing the data, we discovered that some women underwent termination of pregnancy in light of aneuploidy or birth defects, which had not been included in their aneuploidy or birth defect data. We also discovered that one baby had been classified as having a birth defects, when in fact it was reported as having immature lungs secondary to prematurity. In addition, the reporting of multiple pregnancies in the paper is very confusing. We clarified all of these issues through correspondence, which is summarised in Appendix 13. One co-author of this review has written a letter, which has been published, outlining the concerns regarding the statistical analysis presented in the paper (Farquhar 2015). Examples of concerns include the adjustment of sample size and the increase of concentration of human serum albumin following interim analysis, and the reporting of 'ongoing implantation rate' as number of transferred embryos opposed to per woman.

Effects of interventions

See: **Summary of findings 1** GM-CSF-supplemented culture media compared to culture media not supplemented with GM-CSF for women undergoing assisted reproduction

GM-CSF-supplemented culture media versus culture media not supplemented with GM-CSF for women undergoing assisted reproduction

A total of five studies undertook this comparison. Three of these studies (1532 participants) reported data that could be included in meta-analysis (Rose 2020; Sbracia 2014; Ziebe 2013).

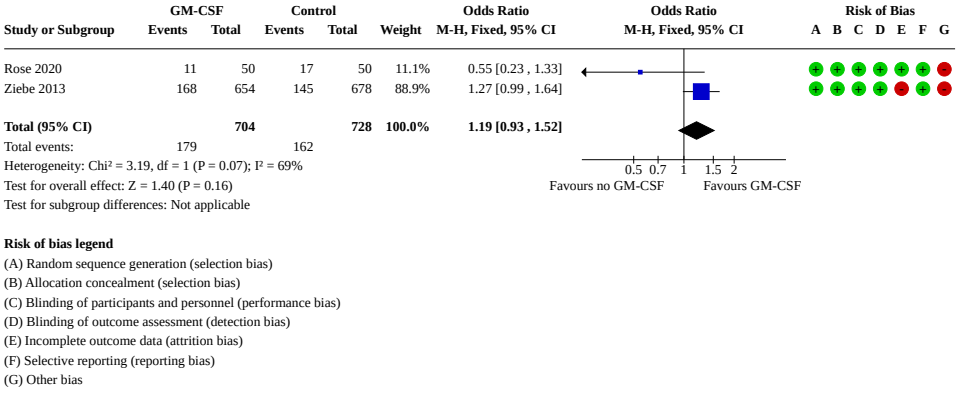
Primary outcomes

1.1 Live birth

Two studies (N = 1432) provided live-birth data (Rose 2020; Ziebe 2013). We obtained ITT live-birth data following correspondence with the authors of Ziebe 2013 (see Appendix 13). There were 179 live births reported amongst the 704 women randomised to the GM-CSF arm, and 162 live births amongst the 728 women randomised to the control arm. No studies reported ongoing pregnancy as a proxy to live birth.

We are uncertain whether GM-CSF-supplemented culture media makes any difference to the live-birth rate when compared to using conventional culture media not supplemented with GM-CSF (odds ratio (OR) 1.19, 95% confidence interval (CI) 0.93 to 1.52, 2 RCTs, N = 1432, I^2 = 69%, low-quality evidence) (Analysis 1.1; Figure 3).

Figure 3. Forest plot of comparison: 1 GM-CSF-supplemented culture medium versus culture medium not supplemented with GM-CSF, outcome: 1.1 Live birth.



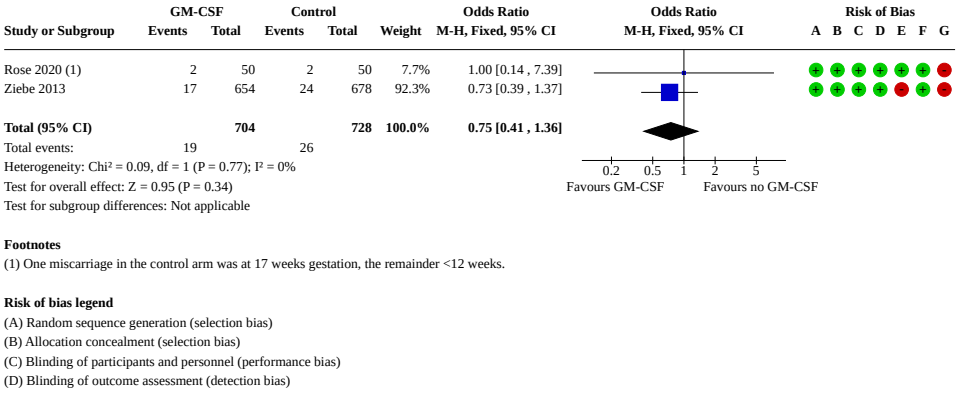
The evidence suggests that if the rate of live birth associated with conventional culture media not supplemented with GM-CSF was 22%, the rate with the use of GM-CSF-supplemented culture media would be between 21% and 30%.

1.2 Miscarriage

Two RCTs (N = 1432) provided miscarriage data (Rose 2020; Ziebe 2013). The authors of both studies were able to clarify that the miscarriages were of clinical pregnancies. Based on correspondence, we were able to remove terminations of pregnancy that had been classified as miscarriage in these two studies (Appendix 13). Terminations of pregnancy as a result of aneuploidy or birth defect are accounted for in Analysis 1.6 and Analysis 1.7. All miscarriages were first-trimester losses, apart from one in the control arm of Rose 2020, which was a midtrimester loss at 17 weeks' gestation. There were 19 miscarriages amongst the 704 women randomised to the GM-CSF arm, and 26 miscarriages amongst the 728 women randomised to the control arm.

It is unclear whether use of GM-CSF-supplemented culture media makes any difference to miscarriage rate when compared to conventional culture media not supplemented with GM-CSF (OR 0.75, 95% CI 0.41 to 1.36, 2 RCTs, N = 1432, I^2 = 0%, low-quality evidence) (Analysis 1.2; Figure 4).

Figure 4. Forest plot of comparison: 1 GM-CSF-supplemented culture medium versus culture medium not supplemented with GM-CSF, outcome: 1.2 Miscarriage.



The evidence suggests that if the miscarriage rate associated with conventional culture media not supplemented with GM-CSF was 4%, the rate with the use of GM-CSF-supplemented culture media would be between 2% and 5%. It is worth noting that these figures are per woman randomised, hence the apparently very low miscarriage rates. They do not include miscarriages that occurred before the diagnosis of a clinical pregnancy on ultrasound scan, otherwise known as biochemical pregnancy losses.

Secondary outcomes

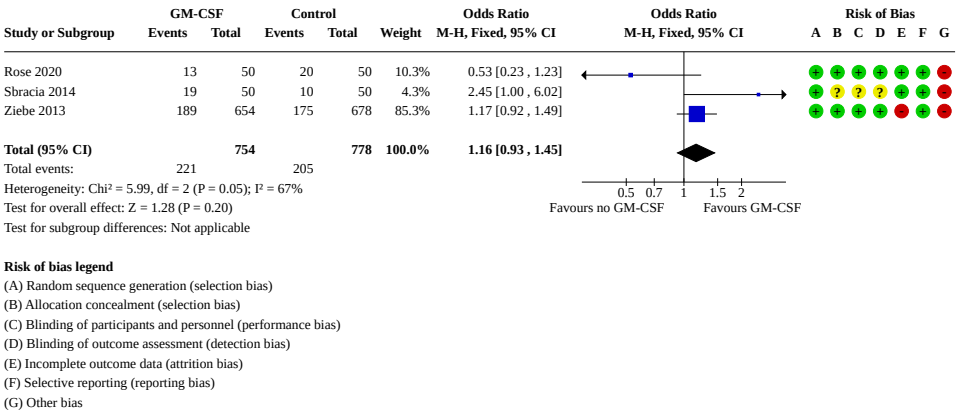
1.3 Clinical pregnancy

Three studies (N = 1532) reported clinical pregnancy rates (Rose 2020; Sbracia 2014; Ziebe 2013). Both Rose 2020 and Ziebe 2013 describe how an ultrasound scan was performed at seven weeks' gestation in order to diagnose clinical pregnancy. Information on the methods of Sbracia 2014 was limited, as the study is only available as a conference abstract, and we received no response to our emails to the authors of the study. The authors of Sbracia 2014 describe pregnancy rate as their primary outcome, however there are no further details as to what stage pregnancy was diagnosed, and whether they were clinical pregnancies diagnosed with ultrasound. The authors report an "implantation rate", which we have taken to mean a biochemical pregnancy rate. Consequently, for the purposes of this review, we have assumed the pregnancy rate in Sbracia 2014 to be clinical.

There were 221 clinical pregnancies amongst the 754 women randomised to the GM-CSF arm, and 205 clinical pregnancies amongst the 778 women randomised to the control arm.

We are uncertain whether GM-CSF-supplemented culture media makes any difference to the clinical pregnancy rate when compared to using a conventional culture medium not supplemented with GM-CSF (OR 1.16, 95% CI 0.93 to 1.45, 3 RCTs, N = 1532 women, I^2 = 67%, low-quality evidence) (Analysis 1.3; Figure 5).

Figure 5. Forest plot of comparison: 1 GM-CSF-supplemented culture medium versus culture medium not supplemented with GM-CSF, outcome: 1.3 Clinical pregnancy.



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

The evidence suggests that if the clinical pregnancy rate associated with conventional culture media not supplemented with GM-CSF was 26%, the rate with the use of GM-CSF-supplemented culture media would be between 25% and 34%.

1.4 Multiple gestation

Two studies (N = 1432) reported multiple gestation rate (Rose 2020; Ziebe 2013). The multiple gestation rate was clarified following correspondence with authors of both studies (Appendix 13). The authors of Ziebe 2013 also detail the incidence of monozygotic and dizygotic twins, but in this review we did not differentiate between types of twins. The authors of Rose 2020 report single-embryo transfer as standard, but explain that four women in the control arm and six women in the intervention arm received double-embryo transfer (Appendix 13). The authors of Ziebe 2013 describe how a maximum of two embryos were replaced per woman with a mean embryo transfer rate of 1.51 for the control arm and 1.49 for the GM-CSF arm. There was one triplet pregnancy in the intervention arm of the study by Ziebe 2013. The remaining multiple gestations reported here were all twins.

There were 31 women with a multiple gestation amongst the 704 women randomised to the GM-CSF arm, and 205 women with a multiple pregnancy amongst the 728 women randomised to the control arm.

We are uncertain whether GM-CSF-supplemented culture media makes any difference to the multiple pregnancy rate when compared to use of a conventional culture medium not supplemented with GM-CSF (OR 1.24, 95% CI 0.73 to 2.10, 2 RCTs, N = 1432, I^2 = 35%, very low-quality evidence) (Analysis 1.4).

The evidence suggests that if the multiple gestation rate associated with conventional culture media not supplemented with GM-CSF was 4%, the rate with the use of GM-CSF-supplemented culture media would be between 3% and 7%.

1.5 Preterm birth

Two studies (N = 1432) reported the preterm birth rate, defined as the birth of a baby (or babies in the case of multiple pregnancy) under 37 weeks' gestation, per woman randomised (Rose 2020; Ziebe 2013). Preterm birth was detailed in the published study by Ziebe 2013. For singletons, the preterm birth data were easily extractable. For women with multiple gestations, the authors of Ziebe 2013 report gestational age at birth with a standard deviation, therefore we clarified these data with the study authors to establish the number of preterm births (Appendix 13). We sought preterm birth data through correspondence with the authors of Rose 2020 (Appendix 13). We counted twins and triplets that were born before 37 weeks as one event for this outcome, as we undertook ITT analysis.

There were 30 women with a preterm birth amongst the 704 women randomised to the GM-CSF arm, and 26 women with a preterm birth amongst the 728 women randomised to the control arm.

We are uncertain whether GM-CSF-supplemented culture media makes any difference to the preterm birth rate when compared to using a conventional culture medium not supplemented by GM-CSF (OR 1.20, 95% CI 0.70 to 2.04, 2 RCTs, N = 1432, I^2 = 76%, very low-quality evidence) (Analysis 1.5).

The evidence suggests that if the preterm birth rate associated with conventional culture media not supplemented with GM-CSF was 4%, the rate with the use of GM-CSF-supplemented culture media would be between 3% and 7%.

1.6 Birth defects

The authors of two studies (N = 1432) were able to provide details on birth defects following correspondence (Appendix 13) (Rose 2020; Ziebe 2013). The authors of Rose 2020 explained that there was one baby with multiple birth defects, which was detected antenatally (this participant was classified as experiencing a miscarriage in the published study, but we have clarified that this was a termination of pregnancy, and it has therefore been removed from the miscarriage group in this review). The authors of Ziebe 2013 provided details on 22 infants who were born with defects (three participants underwent termination of pregnancy for birth defects). We did not count any infants as having both birth defects and aneuploidy, but rather divided them into one of the two groups. We are not aware of twins within the birth defects group, and have assumed all data to be per woman randomised.

Thirteen women had a baby with a birth defect amongst the 704 women randomised to the GM-CSF arm, and 10 women had a baby with a birth defect amongst the 728 women randomised to the control arm.

We are uncertain whether GM-CSF-supplemented culture media makes any difference to the rate of birth defects when compared to using a conventional culture medium not supplemented by GM-CSF (OR 1.33, 95% CI 0.59 to 3.01, I^2 = 0%, 2 RCTs, N = 1432, low-quality evidence) (Analysis 1.6).

The evidence suggests that if the birth defect rate associated with conventional culture media not supplemented with GM-CSF was 1%, the rate with the use of GM-CSF-supplemented culture media would be between 1% and 4%.

1.7 Aneuploidy

Two studies (N = 1432) provided data regarding aneuploidy after we inquired about this outcome (Rose 2020; Ziebe 2013). The authors of Rose 2020 described how one baby had a trisomy, which was detected antenatally, and Ziebe 2013 reported that one baby had a trisomy detected antenatally.

No women had a baby with aneuploidy amongst the 704 women randomised to the GM-CSF arm, and two women had a baby with aneuploidy amongst the 728 women randomised to the control arm. We are uncertain whether GM-CSF-supplemented culture media makes any difference to the rate of aneuploidy when compared to using a conventional culture medium not supplemented by GM-CSF (OR 0.34, 95% CI 0.03 to 3.26, I^2 = 0%, 2 RCTs, N = 1432, low-quality evidence) (Analysis 1.7).

The evidence suggests that if the aneuploidy rate associated with conventional culture media not supplemented with GM-CSF was 0.3%, the rate with the use of GM-CSF-supplemented culture media would be between 0% and 0.9%.

1.8 Stillbirth

Two studies (N = 1432) reported stillbirth (Rose 2020; Ziebe 2013). Following correspondence, the authors of Rose 2020 provided data on stillbirth that were not published (Appendix 13). There were no stillbirths reported in either arm of the study, hence the OR was not estimable (Analysis 1.8). The average rate of stillbirth ranges from approximately 4 per 1000 total births in high-income countries to approximately 28 per 1000 total births in low-income countries such as sub-Saharan Africa (Lawn 2016), therefore the stillbirth rate in this review is better than average for the high income countries.

Planned additional analyses

We did not need to undertake a funnel plot to explore the possibility of small-study effects as there were only three included studies in the quantitative analysis.

Subgroup analyses

1) Studies including only women with recurrent implantation failure. Two studies were defined as including women with "poor prognosis" as a result of previous recurrent implantation failure. The definition of recurrent implantation failure (the failure to achieve a clinical pregnancy after transfer of at least four good-quality embryos in a minimum of three fresh or frozen cycles (Coughlan 2014)) was met by only one of these studies (Sbracia 2014). When examining this study alone, the only outcome it informs is clinical pregnancy. The low-quality evidence suggests that GM-CSF-supplemented culture media may slightly improve pregnancy rates when compared to culture media not supplemented by GM-CSF (OR 2.45, 95% CI 1.00 to 6.02, 1 RCT, N = 100, low-quality evidence).

2) Single-step versus sequential culture media. A single-step culture medium supplemented with GM-CSF would involve culturing the embryos in one medium following fertilisation all the way through to blastocyst embryo replacement if required. The included study that cultured embryos through to blastocyst utilised a sequential culture medium supplemented with GM-CSF (Rose 2020), therefore we were unable to undertake this subgroup analysis. Correspondence with Cooper Surgical revealed that a single-step culture medium supplemented by GM-CSF is yet to

obtain its CE mark (certification mark that indicates conformity with health, safety, and environmental protection standards for products sold within the European Economic Area), and for this reason is not yet available in Europe from this company.

3) Donor versus autologous oocytes. No included studies utilised donor oocytes, so a subgroup analysis was not possible.

4) Studies including only women with recurrent miscarriage. No included studies involved only women who had experienced recurrent miscarriage, so a subgroup analysis was not possible.

5) Studies replacing embryos at cleavage stage versus blastocyst stage. We know that Rose 2020 was the only study that definitely replaced all embryos at day 5, thereby satisfying the criterion of blastocyst stage transfer. The authors of Ziebe 2013 describe how they undertook all day 3 embryo transfers, which classifies this study as cleavage stage transfer. The authors of Sbracia 2014 did not describe whether they undertook cleavage stage or blastocyst stage transfer. However, they do report using EmbryoGen as the intervention culture media, which is a culture medium licensed to culture embryos to day 3, therefore we have assumed for the sake of subgroup analysis that Sbracia 2014 is classified as a cleavage stage transfer study.

Two studies, one cleavage stage transfer, Ziebe 2013, and one blastocyst stage transfer, Rose 2020, reported on the outcome of live birth. The subgroup analysis for both cleavage stage and blastocyst stage transfer did not alter the finding from the pooled meta-analysis. In other words, for both subgroup analyses and the main pooled meta-analysis, we are uncertain whether GM-CSF-supplemented culture media makes any difference to the live-birth rate when compared to using conventional culture media not supplemented with GM-CSF. The quality of the evidence of the subgroup analyses is low given that only one study informs each analysis.

The same two studies that reported on live birth also reported on miscarriage (Rose 2020; Ziebe 2013), one cleavage and one blastocyst stage transfer. The subgroup analyses did not change the outcome of the main meta-analysis. In other words, we are uncertain whether GM-CSF-supplemented culture media makes any difference to the miscarriage rate when compared to using conventional culture media not supplemented with GM-CSF. The quality of the evidence of the subgroup analyses is low given that only one study informs each analysis.

The outcome clinical pregnancy is informed by three included studies. Two studies undertook cleavage stage transfers (Sbracia 2014; Ziebe 2013), and one undertook blastocyst stage transfer (Rose 2020). The subgroup analyses did not alter the results of the main pooled meta-analysis. Low-quality evidence remains of uncertainty as to whether GM-CSF-supplemented culture media makes any difference to the clinical pregnancy rate when compared to using a conventional culture medium not supplemented with GM-CSF.

The outcome multiple gestation is informed by two included studies, one of which undertook cleavage stage transfers, Ziebe 2013, and the other blastocyst stage transfers, Rose 2020. The subgroup analyses did not alter the results of the main pooled meta-analysis. Low-quality evidence remains of uncertainty as to whether GM-CSF-supplemented culture media makes any difference to the multiple gestation rate when compared to using a conventional culture medium not supplemented with GM-CSF. We downgraded the quality of the evidence from moderate for the main meta-analysis, to low, given that only one study informs each subgroup analysis.

Sensitivity analyses

We decided to undertake three sensitivity analyses for the primary outcomes to determine whether our conclusions were robust to arbitrary decisions made regarding the eligibility and analysis. The analyses were as follows.

1) If eligibility was restricted to studies without high risk of bias (studies at low risk of bias were defined as those with low risk of bias in at least the following two domains: random sequence generation and allocation concealment). For live birth and miscarriage, our two primary outcomes, both studies included in the meta-analysis were low risk according to our definition, therefore we did not need to undertake this sensitivity analysis.

2) If a random-effects model had been adopted. We applied the random-effects model to both of our primary outcomes, and it did not alter the conclusions of the review.

3) If the summary effect measure was risk ratio rather than OR. We altered the summary effect measure to risk ratio for both live birth and miscarriage, however it did not alter the conclusions of the review.

DISCUSSION

Summary of main results

Despite GM-CSF-supplemented culture media being commercially available for a number of years, there were very few RCTs from around the world with data that could be included in this review. The three trials included in meta-analysis, Rose 2020; Sbracia 2014; Ziebe 2013, involved a total 1532 women, of whom the vast majority (1332 women) were from Ziebe 2013, a trial designed, conducted, and written by Cooper Surgical, one of the global market leaders in culture media supplemented with GM-CSF. Having said this, there was transparency in communication with the authors of Ziebe 2013; this review contains ITT data and further details on methods as a result of the authors' willingness to share information (Appendix 13).

For the primary and secondary outcomes assessed, including live birth, miscarriage, clinical pregnancy, multiple gestation, preterm birth, birth defects, and aneuploidy, due to very low- to low-quality evidence we cannot be certain whether GM-CSF is any more or less effective than culture media not supplemented with GM-CSF for clinical outcomes that reflect effectiveness and safety (Summary of findings 1). We were unable to undertake analysis of stillbirth, as there were no events in either study arm; however, the lack of events in either arm supports the hypothesis that there is no advantage or disadvantage to using culture media supplemented by GM-CSF versus culture media not supplemented by GM-CSF.

Overall completeness and applicability of evidence

The evidence for this review was dominated by one large, multicentre European RCT, which was designed and conducted by industry. However, despite concerns about the equipoise of the trial designers, co-ordinators, and data analysts, the study appears

to be well executed overall. There is without doubt a number of flaws in the design of the study, in particular the statistical analysis of the results presented in the paper (Farquhar 2015); however, we sought ITT data from the authors, which means the data are as transparent as possible (Appendix 13). The three studies included in the quantitative analysis have a number of similarities as described below, but most notable is the concentration of GM-CSF within the intervention culture media, which is the same across all studies. This possibly reflects the dominance of one particular company who supplied the intervention culture media for all of the included studies. On balance, the available data represent women or couples attending for assisted reproduction with their own gametes, on a single fresh embryo transfer cycle. Two hundred of the 1532 included women were considered to be 'poor prognosis', with recurrent implantation failure or a history of poor embryo development. Further studies including frozen cycles, donor oocytes, and cumulative embryo transfer data from one cycle are required to make the evidence more broadly applicable to the types of women or couples attending for assisted reproduction.

The studies included in the quantitative analysis were conducted in high-income countries, as defined by the World Bank (World Bank). Two included studies were undertaken in the upper-middle-income country of Iran (Zafardoust 2017; Zavvar 2016), however no data could be reliably extracted from these studies, therefore the data available for this review were based solely on those residing in high-income countries. There was a mixture of IVF and ICSI across the three studies included in quantitative analysis. All three studies undertook fresh embryo transfers, and no cumulative embryo transfer data were available from subsequent frozen embryo transfer cycles.

The intervention culture media contained the same concentration of GM-CSF in all three studies, which makes the intervention arm more homogenous in terms of what participants received than the control culture media, which were described in less detail. Rose 2020 describes a control of "standard embryo culture media", whilst the authors of Sbracia 2014 do not provide any information on the control medium. The authors of Ziebe 2013 describe using "EmbryoAssist without cytokine", a culture media manufactured by Cooper Surgical. The potential variation in the control culture media makes the result more generalisable to 'real-world' practice, where individual clinics use a variety of culture media 'as standard'.

The inclusion of studies with variations in day of embryo transfer, poor- and good-prognosis patients, variation in underlying medical conditions of participants, and numbers of embryo transferred helps make the results of this review generalisable.

The two studies that could not be included in the quantitative analysis, Zafardoust 2017; Zavvar 2016, were in some ways different to the studies included in the quantitative analysis. For example, Zafardoust 2017 utilised frozen embryos, and Zavvar 2016 included women in whom immature oocytes were retrieved in spite of stimulation with gonadotropins. Both studies included the same concentration of GM-CSF in the intervention culture media (2 ng/mL), which were supplied by Cooper Surgical, the supplier of all of the intervention culture media across all included studies. Both Zafardoust 2017 and Zavvar 2016 undertook ICSI. It may be possible to include data from these two studies in the future, if we are able to make contact with the authors to clarify issues regarding data and methods. Both studies reported that there was no significant difference in the pregnancy rate between the GM-CSF and conventional culture media arms of their studies.

Quality of the evidence

Overall, the quality of the evidence using the GRADE approach is low or very low for all outcomes (Summary of findings 1). Live birth, clinical pregnancy, multiple gestation, and preterm birth were all downgraded once for inconsistency. There were differing point estimates in the included studies, with one supporting GM-CSF-supplemented culture media, and one supporting the control culture medium. The point estimates have broad confidence intervals, and in many cases the I^2 result is high, representing a high degree of heterogeneity between trials.

Live birth, miscarriage, multiple gestation, preterm birth, birth defects, and aneuploidy were all downgraded for imprecision. We downgraded live birth once and the remaining outcomes twice. These outcomes had point estimates with broad or very broad confidence intervals, from a low number of included studies, at least one of which was very small in terms of number of participants. We downgraded clinical pregnancy once for risk of bias, as all studies that inform this outcome have high risk of other bias, and one study was at unclear risk of selection, performance, and detection bias. The unclear risk of bias in this study was due mainly to a lack of information from the available published abstract.

We were unable to undertake meta-analysis of the results for stillbirth because there we no occurrences of stillbirth in either arm of the study in the one trial that reported this outcome.

Regarding risk of bias, both Rose 2020 and Ziebe 2013 were considered to be overall low risk of bias. Rose 2020 was low risk in all domains except for other bias, which was rated high risk because although the study appears to be well run, the sample size of 100 participants is small, and is unlikely to be powered to detect meaningful difference between the groups in terms of clinical outcomes. We considered Ziebe 2013 to be at low risk of selection, performance, detection, and reporting bias, but at high risk of attrition and other bias. Sbracia 2014 was rated as having high risk of other bias and unclear risk of selection, performance, and detection bias.

Potential biases in the review process

We aimed to identify all eligible studies for inclusion in this review, and contacted authors of all five included studies on many occasions in an effort to include as much information as possible. The authors of two studies were forthcoming with further study information, which helped us to acquire a full picture of the study outcomes, as well as providing information needed to assess and establish risk of bias (Appendix 13).

Agreements and disagreements with other studies or reviews

We found one published systematic review examining GM-CSF-supplemented culture media for women undergoing assisted reproduction (Siristatidis 2013). This review undertook a search of studies published between 1966 and 2012. Siristatidis 2013 included all study designs except case series and case reports. The primary outcome was live birth per woman/couple. Secondary outcomes were clinical pregnancy per woman/couple (defined as

evidence of fetal heart on ultrasound at seven weeks), miscarriage rate (defined as the number of miscarriages divided by the number of clinical pregnancies), fertilisation rate (rate of oocytes fertilised per oocytes retrieved), and laboratory parameters, such as progression of embryos to blastocyst stage, blastocyst performance and hatching, and chromosomal abnormalities of the embryos. The search yielded 152 records, 112 of which were discarded. Six of the remaining 41 studies were considered eligible for inclusion in the review, four of which were RCTs and two prospective observational studies. The review by Siristatidis 2013 has one RCT in common with our review (Ziebe 2013). The other three RCTs included in Siristatidis 2013 were excluded here because they randomised oocytes (Shapiro 2003; Sjoblom 2001), or no embryos were replaced in women (Agerholm 2010).

The authors of Siristatidis 2013 concluded that most of the included studies trend towards favouring the supplementation of culture media with GM-CSF in terms of good-quality embryos reaching the blastocyst stage, improved hatching initiation and number of cells in the blastocyst, and reduction in cell death. However, no statistically significant differences were found in implantation and pregnancy rates in all but one trial, which reported favourable outcomes in terms of implantation and live birth. The authors of Siristatidis 2013 go on to propose properly conducted and adequately powered RCTs to further validate and extrapolate the current findings. The quality of included studies is deemed by the Siristatidis 2013 authors to be "average".

Our review adds two further RCTs, and has the advantage of conducting a meta-analysis and undertaking a 'Risk of bias' assessment, as well as applying GRADE to the findings of the review. The fact that three of the included RCTs in Siristatidis 2013 randomised oocytes or did not replace embryos into the women means that interpreting the data and applying it to real-world clinical situations is almost impossible, as the trial design is not adequate for phase III clinical trials assessing clinical outcomes.

We received email correspondence from a reader of the protocol for this review who expressed concerns that we were planning to include studies that used GM-CSF or G-CSF as the intervention medium (Appendix 14). The author expressed concerns about this decision as they felt those media were not homogenous enough to be included in one group. This review only found studies that included GM-CSF as the intervention medium and we explained that we would look into this issue for the re-run of the review in the coming years.

AUTHORS' CONCLUSIONS

Implications for practice

As is the case with most in vitro fertilisation (IVF) add-ons, granulocyte macrophage colony-stimulating factor (GM-CSF)-supplemented culture media is already widely in use in many fertility clinics across the world, well before the arrival of this systematic review of randomised trials. Despite GM-CSF being available commercially for a number of years, there are still only five randomised trials, only three of which have data to extract for meta-analysis, and two of which are extremely small and not adequately powered to detect a meaningful clinical differences between groups. It is also notable that two of the three studies included in quantitative analysis were sponsored by industry; however, the risk of bias in these studies appears to be low (Figure 2).

The findings of this review reveal that overall there is very low- to low-quality evidence to suggest that GM-CSF supplemented culture media is no more or less effective or harmful than culture media not supplemented with GM-CSF for the following outcomes: live birth, miscarriage, clinical pregnancy, multiple gestation, preterm birth, birth defects, and aneuploidy. The evidence for stillbirth was also low quality, but we were unable to undertake meta-analysis as there were no events in either arm of the studies.

Given these findings, clinicians, embryologists and women/couples considering using GM-CSF-supplemented culture media during an assisted reproduction cycle should be aware that the available evidence neither supports, nor opposes its use. It would appear important that independent information on the available evidence is made accessible to those considering using GM-CSF-supplemented culture media, particularly in the face of strong marketing information. GM-CSF-supplemented culture media is marketed as being recommended for those who have experienced recurrent clinical and biochemical pregnancy loss; recurrent implantation failure; unexplained infertility; and advanced maternal age (Cooper Surgical 2020a). However, the available evidence from randomised controlled trials (RCTs) only includes two small studies (N = 200), which included women aged 40 or over, or women who had experienced recurrent miscarriage or implantation failure (Rose 2020; Sbracia 2014). The claim that GM-CSF has a positive effect on implantation and pregnancy rates is simply not supported by the available evidence presented here (Cooper Surgical 2020a).

Implications for research

The question of whether GM-CSF-supplemented culture media has any advantage over culture media not supplemented by GM-CSF urgently requires further high-quality, independent, adequately powered RCTs to add to this systematic review. The ideal RCT would pre-register a protocol, recruit women/couples attending for assisted reproduction, of all ages and backgrounds, with all types of infertility, undergoing IVF or intracytoplasmic sperm injection (ICSI). It would be powered highly enough to be able to detect whether particular subgroups would benefit from this type of culture media, for example those with previous failed IVF attempts. Women, clinicians, embryologists, and those assessing outcomes would be blinded to the intervention. After randomisation, women would remain in the group to which they had been allocated for the sake of data analysis, regardless of whether they had received the allocated intervention, or whether they dropped out, in order to maintain the effects of randomisation. Data would be analysed using intention-to-treat, maintaining the denominator as all women who were randomised to that arm of the study, for all outcomes, including pregnancy, miscarriage, and live birth. The study would be registered on a trials registry beforehand, and the protocol published, including which statistical analyses were planned. The study would ideally be designed and conducted, and data analysed by independent researchers, who hold equipoise regarding the efficacy of GM-CSF-supplemented culture media.

ACKNOWLEDGEMENTS

We would like to thank Jasmine Lee for her contribution to the protocol draft.

We thank Rik van Eekelen, Demián Glujovsky, Debbie Blake, Mohamed Youssef, Eleonora Uphoff, Co-ordinating Editor Madelon van Wely, and the editors who gave their time to peer review the protocol and review.

Thank you to Marian Showell, Information Specialist, for her time and expertise with the search of the literature.

Special thanks to Elena Kostova, Managing Editor, for her help in getting this review through all stages to publication.

Thank you to the authors of studies who provided additional information and data at our request, and to those women who participated in the trials which have contributed to this review.

REFERENCES

References to studies included in this review

Rose 2020 *{published and unpublished data}*

EmbryoGen/Blastgen for Couples With Implantation Problems or Previous Miscarriage (BlastGen) [NCT02305420]. ClinicalTrials.gov 2014.

Hull ML, Barry ML, Cameron LP, Knight EJ, Norman RJ, Rose RD. A randomised control pilot study comparing Embryogen®/ BlastGen™ with standard media culture in poor prognosis couples undertaking IVF. In: Fertility Society of Australia Annual Conference. 1 edition. Vol. 1. Melbourne: Fertility Society of Australia, 2018:85.

* Rose RD, Barry MF, Dunstan EV, Yuen SM, Cameron LP, Knight EJ, Norman RJ, Hull ML. The BlastGen study: A randomised controlled trial of GM-CSF supplemented blastocyst media. *Reproductive BioMedicine Online* 2020;**www.doi.org/10.1016/j.rbmo.2020.01.011**.

Sbracia 2014 *{published data only}*

* Sbracia M, Morgia F, Torti M, Amodei M, Corizza R, Scarpellini F. Use of GM-CSF supplemented IVF medium in patients with recurrent implantation failure: A controlled trial. In: Human Reproduction. Vol. 29 (Suppl. 1). Oxford University Press, 2014.

Use of GM-CSF Supplemented IVF Medium in Patients With Recurrent Implantation Failure [NCT01718210]. ClinicalTrials.gov 2012.

Zafardoust 2017 *{published data only}*

IRCT2014021611430N4. Iranian Registry of Clinical Trials 2014.

Zafardoost S. Evalutation of efficacy of GM-CSF in embryo culture for freezed embryos in embryo development and pregnancy outcomes. Iranian Registry of Clinical Trials 17/03/2014;**www.irct.ir/trial/11703**.

* Zafardoust S, Mohammadzadeh A, Fatemi F, Jouhari S, Karimi A, Shahbakhsh A, Barchinezhad M, Fathi Z, Sadeghi MR. Evaluation of efficacy of GM-CSF in embryo culture for frozen embryos on embryo development and pregnancy outcomes. In: Human Reproduction. Vol. 32 (Suppl. 1). Netherlands Oxford University Press, 2017.

Zavvar 2016 *{published data only}*

Zavvar TS, Roustaei H, Torshizi MM, FadaviIslam M, Sabouri E. The effect of embryogen medium on oocyte in vitro maturation, fertilization rate and embryo development. In: Journal of Reproduction and Infertility. Vol. 17 (2 Suppl.2). Netherlands Avicenna Research Institute, 2016:99-100.

Ziebe 2013 *{published data only}*

Ziebe S, Loft A, Povlsen BB, Erb K, Agerholm I, Aasted M, Gabrielsen A, Hnida C, Zobel DP, Munding B, Bendz SH, Robertson SA. A randomized clinical trial to evaluate the effect of granulocyte-macrophage colony-stimulating factor (GM-CSF) in embryo culture medium for in vitro fertilization. *Fertility & Sterility* 2013;**99**(6):1600-1609.

References to studies excluded from this review

Agerholm 2010 *{published data only}*

Agerholm I, Loft A, Hald F, Lemmen JG, Munding B, Sorensen PD, Ziebe S. Culture of human oocytes with granulocyte-macrophage colony-stimulating factor has no effect on embryonic chromosomal constitution. *Reproductive Biomedicine Online* 2010;**20**(4):477-84.

Fawzy 2019 *{published data only}*

Fawzy M, Emad M, Elsuity MA, Mahran A, Abdelrahman MY, Fetih AN, Abdelghafar H, Sabry M, Nour M, Rasheed SM. Cytokines hold promise for human embryo culture in vitro: results of a randomized clinical trial. *Fertility and Sterility* 2019;**112**(5):849-57.

Kinoshita 2019 *{published data only}*

Kinoshita M, Fujita M, Takahashi K. The use of granulocyte macrophage colony-stimulating factor (GM-CSF) containing media for embryo transfer increases the clinical pregnancy and ongoing pregnancy rate. *Human Reproduction* 2019;**34**(Suppl 1):368.

Min 2017 *{published data only}*

Min SH, Ja-Seong K, Hyeyoung L, Hwa-Sook M. Treatment with GM-CSF improves pregnancy rates in poor responders. *Human Reproduction* 2017;**32**(Suppl 1).

Scarpellini 2011 *{published data only}*

Scarpellini F, Sbracia M. The use of G-CSF for implantation failure in IVF: a clinical trial. *Fertility and Sterility* 2011;**96**(Suppl 3).

Sfontouris 2013 *{published data only}*

Sfontouris IA, Lainas GT, Anagnostara K, Kolibianakis EM, Lainas TG. Effect of granulocyte-macrophage colony-stimulating factor (GM-CSF) on pregnancy rates in patients with multiple unsuccessful IVF attempts. *Human Reproduction* 2013;**28**(Suppl 1).

Shapiro 2003 *{published data only}*

Shapiro B, Richter K, Daneshmand S, Quinn P, Behr B. Granulocyte-macrophage colony-stimulating factor enhances human embryo development to the blastocyst stage: a randomized study. *Fertility & Sterility* 2003;**79**(Suppl 2):S15-16.

Siqueira 2016 *{published data only}*

Siqueira LGB, Tribulo P, Denicol AC, Ortega MS, Negron-Perez VM, Kannampuzha-Francis J, Hansen PJ. Consequences of in vitro production of embryos with or without colonystimulating factor 2 in culture medium on morphometric features of the bovine conceptus at day 86 of gestation. *Reproduction, Fertility and Development* 2016;**28**:218.

Siristatidis 2013 *{published data only}*10.1155/2013/704967

Siristatidis C, Vogiatzi P, Salamalekis G, Creatsa M, Vrachnis N, Glujovsky D, Iliodromiti Z, Chrelias C. Granulocyte macrophage colony stimulating factor supplementation in culture media

for subfertile women undergoing assisted reproduction technologies: A systematic review. *International Journal of Endocrinology* 2013;**2013:704967**.

Sjoblom 1998 *{published data only}*

Sjoblom C, Robertson SA, Wikland M. Granulocyte-macrophage colony-stimulating factor promotes the rate and extent of human blastocyst development in vitro. *Human Reproduction* 1998;**13**.

Sjoblom 1999 *{published data only}*

Sjoblom C, Robertson SA, Wikland M. Granulocyte -macrophage colony stimulating factor (GM-CSF) promotes development of human pre-implantation embryos through the alpha-chain of the GM-CSF receptor. In: 11th world congress on in vitro fertilization and human reproductive genetics. 1999.

Sjoblom 1999a *{published data only}*

Sjoblom C, Wikland M, Robertson SA. Granulocyte-macrophage colony-stimulating factor promotes human blastocyst development in vitro. *Human Reproduction* 1999;**14**(12):3069-76.

Sjoblom 2001 *{published data only}*

Sjoblom C, Wikland M, Robertson S. Granulocyte-macrophage-colony stimulating factor (GM-CSF) promotes inner cell mass blastomere viability in human pre-implantation embryos. *Fertility & Sterility* 2001;**76**(3 Suppl 1):S70.

References to studies awaiting assessment

ISRCTN94726536 *{published data only}*

ISRCTN94726536. Clinical efficiency of a ready-to-use medium for human embryo culture supplemented with growth factors in patients with previous in vitro fertilization failures. isrctn.com/ISRCTN94726536 first received 12 January 2016.

NCT01689428 *{published data only}*

NCT01689428. Ism1 Versus EmbryoGen Media for Embryo Culture in Previous Pregnancy Loss Cases. clinicaltrials.gov/ct2/show/NCT01689428 first received 21 September 2012.

NCT01689454 *{published data only}*

NCT01689454. ISM1 Versus EmbrioGen Media for Embryo Culture With Previous Failure Implantation in ART Treatments. clinicaltrials.gov/ct2/show/NCT01689454 first received 21 September 2012.

NCT02651285 *{published data only}*

NCT02651285. Use of G-CSF Supplemented IVF Medium in Patients Undergoing IVF. clinicaltrials.gov/ct2/show/NCT02651285 first received 8 January 2016.

Additional references

Agerholm 2010

Agerholm I, Loft A, Hald F, Lemmen JG, Munding B, Sorensen PD, et al. Culture of human oocytes with granulocyte-macrophage colony-stimulating factor has no effect on embryonic chromosomal constitution. *Reproductive Biomedicine Online* 2010;**20**(4):477-84.

Baldwin 1992

Baldwin GC. The biology of granulocyte-macrophage colony-stimulating factor: effects on hematopoietic and nonhematopoietic cells. *Developmental Biology* 1992;**151**(2):352-67. [0012-1606]

Biggers 2000

Biggers JD. Ethical issues and the commercialization of embryo culture media. *Reproductive Biomedicine Online* 2000;**1**:74-6.

Block 2003

Block J, Drost M, Monson RL, Rutledge JJ, Rivera RM, Paula-Lopes FF, et al. Use of insulin-like growth factor-I during embryo culture and treatment of recipients with gonadotropin-releasing hormone to increase pregnancy rates following the transfer of in vitro-produced embryos to heat-stressed, lactating cows. *Journal of Animal Science* 2003;**81**(6):1590-602.

Burgess 1977

Burgess AW, Wilson EMA, Metcalf D. Stimulation by human placental conditioned medium of hemopoietic colony formation by human marrow cells. Blood 1977;**49**(4):573-83.

Chronopoulou 2015

Chronopoulou E, Harper JC. IVF culture media: past, present and future. Human Reproduction Update 2015;**21**(1):39-55.

Cooper Surgical 2020

Cooper Surgical. Bringing knowledge to life. www.fertility.coopersurgical.com/about-us/ (accessed 11 March 2020).

Cooper Surgical 2020a

Cooper Surgical. Culture media with GM-CSF. Unique culture media for poor prognosis patients. www.fertility.coopersurgical.com/products/culture-media-with-gm-csf/ (accessed 13 March 2020).

Coughlan 2014

Coughlan C, Ledger W, Wang Q, Liu F, Demirol A, Gurgan T, et al. Recurrent implantation failure: definition and management. *Reproductive BioMedicine Online* 2014;**28**(1):14-38.

Covidence [Computer program]

Veritas Health Innovation Covidence. Version accessed 1 Jan 2020. Melbourne, Australia: Veritas Health Innovation.Available at covidence.org.

Dyrlund 2014

Dyrlund TF, Kirkegaard K, Poulsen ET, Sanggaard KW, Hindkjaer JJ, Kjems J, et al. Unconditioned commercial embryo culture media contain a large variety of non-declared proteins: a comprehensive proteomics analysis. Human Reproduction (Oxford, England) 2014;**29**(11):2421-30.

Farquhar 2015

Farquhar C, Vail A, Mourad S. [personal communication to editor]. Letter to: editor of Fertility & Sterility 01 December 2015.

Giacomini 1995

Giacomini G, Tabibzadeh SS, Satyaswaroop PG, Bonsi L, Vitale L, Bagnara GP, et al. Epithelial cells are the major source of biologically active granulocyte macrophage colony-stimulating factor in human endometrium. Human Reproduction (Oxford, England) 1995;**10**(12):3259-63.

GRADEpro GDT [Computer program]

McMaster University (developed by Evidence Prime) GRADEpro GDT. Version accessed 1 Feb 2020. Hamilton (ON): McMaster University (developed by Evidence Prime).Available at gradepro.org.

Hambartsoumian 1998

Hambartsoumian E. Endometrial leukemia inhibitory factor (LIF) as a possible cause of unexplained infertility and multiple failures of implantation. American Journal of Reproductive Immunology 1998;**39**(2):137-43.

Heneghan 2016

Heneghan C, Spencer EA, Bobrovitz N, Collins DR, Nunan D, Pluddemann A, et al. Lack of evidence for interventions offered in UK fertility centres. *BMJ* 2016;**355**:i6295.

Higgins 2011

Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Higgins 2019

Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editor(s). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook.

Jokhi 1994

Jokhi PP, King A, Loke YW. Production of granulocyte-macrophage colony-stimulating factor by human trophoblast cells and by decidual large granular lymphocytes. Human Reproduction (Oxford, England) 1994;**9**(9):1660-9.

Lawn 2016

Lawn JE, Blencowe J, Waiswa P, Amouzou A, Mathers C, Hogan D, et al. Stillbirths: rates, risk factors, and acceleration towards 2030. *Lancet* 2016;**387**(10018):587-603.

Lazzari 2002

Lazzari G, Wrenzycki C, Herrmann D, Duchi R, Kruip T, Niemann H, et al. Cellular and molecular deviations in bovine in vitro-produced embryos are related to the large offspring syndrome. Biology of Reproduction 2002;**67**(3):767-75.

Lighten 1998

Lighten AD, Moore GE, Winston RM, Hardy K. Routine addition of human insulin-like growth factor-I ligand could benefit clinical in-vitro fertilization culture. Human Reproduction (Oxford, England) 1998;**13**(11):3144-50.

Lim 2006

Lim JJ, Lee DR, Song HS, Kim KS, Yoon TK, Gye MC, et al. Heparin-binding epidermal growth factor (HB-EGF) may improve embryonic development and implantation by increasing vitronectin receptor (integrin alphanubeta3) expression in peri-implantation mouse embryos. Journal of Assisted Reproduction and Genetics 2006;**23**(3):111-9.

Mignini 2013

Mignini Renzini M, Dal Canto M, Coticchio G, Comi R, Brigante C, Caliari I, et al. Clinical efficiency and perinatal outcome of ART cycles following embryo culture in the presence of GM-CSF in patients with miscarriage or early pregnancy loss history. Human Reproduction (Oxford, England) 2013;**28**(Suppl 1):i139-49.

Morbeck 2014

Morbeck DE, Krisher RL, Herrick JR, Baumann NA, Matern D, Moyer T. Composition of commercial media used for human embryo culture. Fertility and Sterility 2014;**102**(3):759-66.e9. [0015-0282]

Rahmati 2015

Rahmati M, Petitbarat M, Dubanchet S, Bensussan A, Chaouat G, Ledee N. Colony stimulating factors 1, 2, 3 and early pregnancy steps: from bench to bedside. Journal of Reproductive Immunology 2015;**109**:1-6. [0165-0378]

Richter 2008

Richter KS. The importance of growth factors for preimplantation embryo development and in-vitro culture. Current Opinion in Obstetrics & Gynecology 2008;**20**(3):292-304.

Robertson 1992

Robertson SA, Mayrhofer G, Seamark RF. Uterine epithelial cells synthesize granulocyte-macrophage colony-stimulating factor and interleukin-6 in pregnant and nonpregnant mice. Biology of Reproduction 1992;**46**(6):1069-79.

Robertson 1994

Robertson SA, Seamark RF, Guilbert LJ, Wegmann TG. The role of cytokines in gestation. Critical Reviews in Immunology 1994;**14**(3-4):239-92.

Robertson 1999

Robertson SA, Roberts CT, Farr KL, Dunn AR, Seamark RF. Fertility impairment in granulocyte-macrophage colony-stimulating factor-deficient mice. Biology of Reproduction 1999;**60**(2):251-61.

Robertson 2007

Robertson SA. GM-CSF regulation of embryo development and pregnancy. Cytokine & Growth Factor Reviews 2007;**18**(3-4):287-98.

Roudebush 2004

Roudebush WE, Massey JB, Kort HI, Elsner CW, Toledo AA, Mitchell-Leef D, et al. Exposure of preimplantation embryos to platelet-activating factor increases birth rate. Journal of Assisted Reproduction and Genetics 2004;**21**(8):297-300.

Savion 2002

Savion S, Zeldich E, Orenstein H, Shepshelovich J, Torchinsky A, Carp H, et al. Cytokine expression in the uterus of mice with pregnancy loss: effect of maternal immunopotentiation with GM-CSF. Reproduction 2002;**123**(3):399-409.

Seymour 1997

Seymour JF, Lieschke GJ, Grail D, Quilici C, Hodgson G, Dunn AR. Mice lacking both granulocyte colony-stimulating factor (CSF) and granulocyte-macrophage CSF have impaired reproductive capacity, perturbed neonatal granulopoiesis, lung disease, amyloidosis, and reduced long-term survival. Blood 1997;**90**(8):3037-49.

Sfontouris 2013

Sfontouris I, Lainas G, Anagnostara K, Kolibianakis E, Lainas T. Effect of granulocyte-macrophage colony-stimulating factor (GM-CSF) on pregnancy rates in patients with multiple unsuccessful IVF attempts. *Human Reproduction (Oxford, England)* 2013;**28**(Suppl 1):160-2.

Sfontouris 2016

Sfontouris IA, Martins WP, Nastri CO, Viana IG, Navarro PA, Raine-Fenning N, et al. Blastocyst culture using single versus sequential media in clinical IVF: a systematic review and meta-analysis of randomized controlled trials. *Journal of Assisted Reproduction and Genetics* 2016;**33**(10):1261-72.

Siristatidis 2013

Siristatidis C, Vogiatzi P, Salamalekis G, Creatsa M, Vrachnis N, Glujovsky D, et al. Granulocyte macrophage colony stimulating factor supplementation in culture media for subfertile women undergoing assisted reproduction technologies: a systematic review. International Journal of Endocrinology 2013;**2013**:704967.

Sjöblom 1992

Sjöblom C, Wikland M, Robertson SA. Granulocyte-macrophage-colony stimulating factor (GM-CSF) promotes inner cell mass blastomere viability in human pre-implantation embryos. Fertility and Sterility 1992;**76**(3):S70. [0015-0282]

Sjöblom 1999

Sjöblom C, Wikland M, Robertson SA. Granulocyte-macrophage colony-stimulating factor promotes human blastocyst development in vitro. Human Reproduction (Oxford, England) 1999;**14**(12):3069-76.

Sjöblom 2002

Sjöblom C, Wikland M, Robertson SA. Granulocyte-macrophage colony-stimulating factor (GM-CSF) acts independently of the beta common subunit of the GM-CSF receptor to prevent inner cell mass apoptosis in human embryos. Biology of Reproduction 2002;**67**(6):1817-23.

Sjöblom 2005

Sjöblom C, Roberts CT, Wikland M, Robertson SA. Granulocyte-macrophage colony-stimulating factor alleviates adverse consequences of embryo culture on fetal growth trajectory and placental morphogenesis. Endocrinology 2005;**146**(5):2142-53.

Spanos 2000

Spanos S, Becker DL, Winston RM, Hardy K. Anti-apoptotic action of insulin-like growth factor-I during human preimplantation embryo development. Biology of Reproduction 2000;**63**(5):1413-20.

Steptoe 1978

Steptoe PC, Edwards RG. Birth after the reimplantation of a human embryo. *Lancet* 1978;**2**:366.

Tevkin 2014

Tevkin S, Lokshin V, Shishimorova M, Polumiskov V. The frequency of clinical pregnancy and implantation rate after cultivation of embryos in a medium with granulocyte macrophage colony-stimulating factor (GM-CSF) in patients with preceding failed attempts of ART. Gynecological Endocrinology 2014;**30**(Suppl 1):9-12.

Torry 2007

Torry DS, Leavenworth J, Chang M, Maheshwari V, Groesch K, Ball ER, et al. Angiogenesis in implantation. Journal of Assisted Reproduction and Genetics 2007;**24**(7):303-15.

Vail 2003

Vail A, Gardener E. Common statistical errors in the design and analysis of subfertility trials. *Human Reproduction (Oxford, England)* 2003;**18**:1000-4.

Vuorela 2000

Vuorela P, Carpen O, Tulppala M, Halmesmaki E. VEGF, its receptors and the tie receptors in recurrent miscarriage. Molecular Human Reproduction 2000;**6**(3):276-82.

Wegmann 1992

Wegmann TG, Guilbert LJ. Immune signalling at the maternal – fetal interface and trophoblast differentiation. *Developmental & Comparative Immunology* 1992;**16**(6):425-30. [0145-305X]

World Bank

The World Bank. Countries and economies. www.data.worldbank.org/country (accessed 11 March 2020).

Young 2001

Young LE, Fernandes K, McEvoy TG, Butterwith SC, Gutierrez CG, Carolan C, et al. Epigenetic change in IGF2R is associated with fetal overgrowth after sheep embryo culture. Nature Genetics 2001;**27**(2):153-4.

Yu 2012

Yu Y, Yan J, Li M, Yan L, Zhao Y, Lian Y, et al. Effects of combined epidermal growth factor, brain-derived neurotrophic factor and insulin-like growth factor-1 on human oocyte maturation and early fertilized and cloned embryo development. Human Reproduction (Oxford, England) 2012;**27**(7):2146-59.

Zhao 1994

Zhao Y, Chegini N. Human fallopian tube expresses granulocyte-macrophage colony stimulating factor (GM-CSF) and GM-CSF alpha and beta receptors and contain immunoreactive GM-CSF protein. Journal of Clinical Endocrinology and Metabolism 1994;**79**(2):662-5.

Ziebe 2013

Ziebe S, Loft A, Povlsen BB, Erb K, Agerholm I, Aasted M, et al. A randomized clinical trial to evaluate the effect of granulocyte-macrophage colony-stimulating factor (GM-CSF) in embryo culture medium for in vitro fertilization. Fertility and Sterility 2013;**99**(6):1600-9.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Rose 2020

Study characteristics	
Methods	**Study design:** single-centre randomised controlled trial (RCT). The study was cross-over in design, but only data from the first phase of the study, prior to cross-over, are included in the review. **Country:** Australia **Setting:** fertility clinic, Adelaide **Sponsorship source**: EmbryoGen/BlastGen media was supplied by ORIGIO, a Cooper Surgical company. A grant from Cooper Surgical also funded MLH and EJK for statistical support. Fertility SA and The Robinson Institute, University of Adelaide provided in kind support from staff and affiliates providing time and expertise. **Trials registry number:** NCT02305420 (prospectively registered)
Participants	**Inclusion criteria**: the inclusion criteria selected for a poor-prognosis patient population. Patients must have previously had consecutive transfer of 2 or more embryos without a positive pregnancy outcome OR have had a history of at least 1 previous pregnancy loss OR a previous history of poor embryo development (< 20% of embryos developing on the time at day 3 or no blastocysts above grade 2 on day 5). Other additional inclusion parameters included a maternal age between 25 and 41 years, the use of a standard gonadotrophin-releasing hormone (GnRH) agonist or antagonist protocol, and 3 or more follicles of > 14 mm as seen by transvaginal ultrasound before the day of human chorionic gonadotropin administration. **Exclusion criteria**: exclusion criteria included a need for surgical sperm retrieval (except in cases of previous vasectomy), the use of another investigational drug within 30 days of oocyte retrieval, and/or the presence of a severe chronic disease that could impact the in-vitro fertilisation (IVF) cycle or reproductive outcomes. **Number of participants randomised to each arm of the study:** 50 to intervention arm, 50 to control arm. **Dropouts:** none **Group differences:** participants were matched for age, body mass index (BMI), and smoking status. A greater number of participants in the intervention group underwent intracytoplasmic sperm injection (ICSI) opposed to IVF versus the control group (37/50 versus 29/50). **Fresh or frozen cycle?:** fresh cycle **IVF or ICSI?:** IVF and ICSI **Length of time exposed to intervention medium:** 5 days **Trade name and concentration of granulocyte macrophage colony-stimulating factor (GM-CSF) in intervention medium**: EmbryoGen days 0 to 3 followed by BlastGen days 3 to 5 **Trade name of control medium:** "Standard embryo culture medium" **Day of embryo transfer:** day 5 (apart from 1 in the control group and 2 in the intervention group who transferred 1 embryo on day 3 due to delayed embryo development)

Rose 2020 *(Continued)*

	Number of embryos transferred: single-embryo transfer. "Couples who insisted on transferring two embryos, once embryo quality was known, were included in the study but were recorded as having a trial variation (four in the standard group and six in the intervention group)."
Interventions	**Intervention**: GM-CSF-supplemented culture medium **Control**: Sydney IVF medium
Outcomes	• Live-birth rate • Miscarriage rate • Clinical pregnancy rate • Multiple gestation rate • Preterm birth rate • Birth defect rate • Aneuploidy rate • Stillbirth rate
Notes	Email correspondence with all additional data outlined in Appendix 13.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"50 cards with control and 50 with BlastGen written on them were placed in opaque sealed envelopes by a person unrelated to the trial. They were shuffled several times and the envelopes were marked from 1-100 in consecutive order."
Allocation concealment (selection bias)	Low risk	"When the eligible patient was scheduled for an egg retrieval the embryologist took the next envelope by sequence and the random assignment was made."
Blinding of participants and personnel (performance bias) All outcomes	Low risk	"Clinicians, sonographers, statisticians, and patients were completely blinded"
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Sonographers were blinded.
Incomplete outcome data (attrition bias) All outcomes	Low risk	All dropouts accounted for.
Selective reporting (reporting bias)	Low risk	"We reported all outcomes as per our prospective clinical trials registration (ClinicalTrials.gov Identifier: NCT02305420)."
Other bias	High risk	There is a potential conflict of interest regarding the funding and administration of this study, as it was granted by the manufacturer of the intervention culture medium. However, the study authors have been very forthcoming with information regarding the trial, and there has been transparency. We assessed the study as at high risk of other bias due to the small sample size of 100 participants.

Sbracia 2014

Study characteristics	
Methods	**Study design:** single-centre RCT **Country:** Italy **Setting:** fertility clinic **Sponsorship source**: no funding **Trials registry number**: NCT01718210 (prospectively registered)
Participants	**Inclusion criteria:** women with recurrent implantation failure, 3 or more consecutive failed IVF cycles with a total of at least 8 good embryos replaced in the uterus, women aged 40 or less **Exclusion criteria**: over 40 years of age, chromosomal defects in the couple, metabolic diseases (diabetes, etc.), other genetic diseases (thalassaemia, cystic fibrosis, etc.) **Number of participants randomised to each arm of the study:** 50 to the intervention arm and 50 to the control arm **Dropouts:** none **Group differences:** not disclosed **Fresh or frozen cycle?:** fresh cycle **IVF or ICSI?**: ICSI **Length of time exposed to intervention medium**: from fertilisation to embryo transfer **Trade name and concentration of GM-CSF in intervention medium**: EmbryoGen **Trade name of control medium:** not disclosed **Day of embryo transfer:** not disclosed **Number of embryos transferred:** a maximum of 3 embryos were transferred
Interventions	**Intervention**: GM-CSF-supplemented culture medium **Control:** standard IVF medium
Outcomes	Clinical pregnancy rate
Notes	Emailed 29 October 2019 and 9 January 2020 for further information, but received no response

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	"Patients were randomised by a computer generated number sequence"
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not described
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	No description of who, if anyone, was blinded

Sbracia 2014 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No description of blinding of outcome assessors
Incomplete outcome data (attrition bias) All outcomes	Low risk	Intention-to-treat data reported, and no dropouts described.
Selective reporting (reporting bias)	Low risk	The 2 outcomes reported in the study were outlined on the trials registry ClinicalTrials.gov Identifier NCT01718210.
Other bias	High risk	Small sample size of 100 participants

Zafardoust 2017

Study characteristics	
Methods	**Study design:** single-centre RCT **Country:** Iran **Setting:** fertility clinic **Sponsorship source**: infertility and recurrent abortion treatment centre of Avicenna **Trials registry number**: www.irct.ir/trial/11703 (registered during recruitment)
Participants	**Inclusion criteria:** couples undergoing treatment with their own gametes, women aged < 40 years old, women with at least 4 good-quality embryos after thawing (grade A), women who had not had more than 1 previous embryo transfer **Exclusion criteria**: cycles requiring pre-implantation genetic diagnosis, women with an anatomic disorder of the uterus or hydrosalpinx/hydrosalpinges **Number of participants randomised to each arm of the study:** number randomised to each arm before dropout was not disclosed **Group differences:** the average age between the 2 groups differed **Fresh or frozen cycle?**: frozen cycles **IVF or ICSI?**: ICSI **Length of time exposed to intervention medium**: from embryo thawing through to embryo transfer **Trade name and concentration of GM-CSF in intervention medium**: BlastGen 2 ng/mL **Trade name of control medium:** Vitrolife, USA **Day of embryo transfer:** not disclosed **Number of embryos transferred**: not disclosed
Interventions	**Intervention:** GM-CSF-supplemented culture medium **Control:** standard IVF medium
Outcomes	Clinical pregnancy rate

Zafardoust 2017 *(Continued)*

Notes	Emailed on 29 October 2019, 26 November 2019, and 9 January 2020 for further information, but received no response

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No description of how randomisation was undertaken
Allocation concealment (selection bias)	Unclear risk	No description of how allocation concealment was undertaken
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	No description of who, if anyone, was blinded
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Blinding of outcome assessors not described
Incomplete outcome data (attrition bias) All outcomes	High risk	10 women were not included in the final analysis, which is a high rate of attrition.
Selective reporting (reporting bias)	High risk	No data on miscarriage, multiple pregnancy, and beta human chorionic gonadotrophin (BHCG) levels, which are secondary outcomes noted on trials register
Other bias	High risk	Small sample size of 90 participants

Zavvar 2016

Study characteristics	
Methods	**Study design:** single-centre RCT **Country:** Iran **Setting:** fertility clinic **Sponsorship source:** not disclosed **Trials registry number:** not disclosed
Participants	**Inclusion criteria**: women who produce only immature oocytes in spite of stimulation with gonadotropins **Exclusion criteria**: not disclosed **Number of participants randomised to each arm of the study:** 31 women were randomised to receive the intervention medium, and 45 were randomised to receive the control medium **Group differences:** not disclosed **Fresh or frozen cycle?:** not disclosed

Zavvar 2016 *(Continued)*

	IVF or ICSI?: ICSI **Length of time exposed to intervention medium:** not disclosed **Trade name and concentration of GM-CSF in intervention medium**: MediCult 2 ng/mL **Trade name of control medium**: not disclosed **Day of embryo transfer:** not disclosed **Number of embryos transferred:** not disclosed
Interventions	**Intervention:** GM-CSF-supplemented culture medium **Control:** standard IVF medium
Outcomes	Clinical pregnancy rate
Notes	First author contacted by email for further information on 26 November 2019 and 9 January 2020, but no response received.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No description of how random sequence was generated
Allocation concealment (selection bias)	Unclear risk	No description of how allocation was concealed
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Blinding was not mentioned.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Blinding of outcome assessors was not described.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	We assessed the study as at unclear risk, as we do not have access to number of dropouts or indeed how many participants were included in the analysis.
Selective reporting (reporting bias)	Unclear risk	No access to a protocol or a trial registry entry
Other bias	High risk	Small sample size of 76 participants

Ziebe 2013

Study characteristics	
Methods	**Study design:** multicentre RCT (14 different clinics) **Country:** study co-ordinated in the Netherlands. Participants were recruited from clinics in Sweden and Denmark.

Ziebe 2013 *(Continued)*

	Setting: fertility clinics **Sponsorship source**: supported by ORIGIO, Maløv, Denmark **Trials registry number:** NCT00565747 (prospectively registered)
Participants	**Inclusion criteria**: women aged 25 to 39 years, women who had a regular menstrual cycle of 21 to 35 days, women treated with a standard gonadotrophin-releasing hormone(GnRH) agonist or antagonist protocol, women with 3 or more follicles with a diameter of >/= 14 mm on the day of human chorionic gonadotrophin (hCG) administration, including a leading follicle of >/= 17 mm **Exclusion criteria:** previous participation in the study, use of assisted hatching, use of non-ejaculated sperm, medical conditions or genetic disorders prohibiting in vitro fertilisation(IVF)/ intracytoplasmic sperm injection (ICSI) or interfering with interpretation of results, use of investigational drugs within 30 days before oocyte retrieval, severe chronic disease of relevance for reproduction, and oocyte donation **Number of participants randomised to each arm of the study:** 654 women randomised to receive GM-CSF-supplemented culture medium, and 678 women randomised to the control culture medium **Group differences:** no significant between-group differences in baseline characteristics **Fresh or frozen cycle?:** fresh **IVF or ICSI?:** IVF and ICSI **Length of time exposed to intervention medium:** from fertilisation through to embryo transfer **Trade name and concentration of GM-CSF in intervention medium**: ORIGIO 2 ng/mL **Trade name of control medium**: EmbryoAssist **Day of embryo transfer:** day 3 **Maxiumum number of embryos transferred:** 2
Interventions	**Intervention:** GM-CSF-supplemented culture medium **Control:** standard IVF medium
Outcomes	• Live-birth rate • Miscarriage rate • Clinical pregnancy rate • Multiple gestation rate • Preterm birth rate • Birth defect rate • Aneuploidy rate • Stillbirth rate
Notes	Email correspondence with all additional data outlined in Appendix 13.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Randomisation (blocks of four) was computer-generated individually for each clinic to maintain balance between the treatment groups at each site."
Allocation concealment (selection bias)	Low risk	Following email correspondence: "Randomization was performed by ORIGIO a/s, and based on a randomisation list per clinic generated automatically using www.randomization.com. Each study site received a list of study specific consecutive patient ID numbers (e.g. clinic 1: 01001, 01002, 01003, 01004 etc.)

Ziebe 2013 *(Continued)*

		and a corresponding number of identically looking randomized bottles of test and control media individually labelled with the corresponding study specific patient ID numbers. On site the lowest number on the list was always allocated to any new patient recruited, at the time of informed consent signature. Therefore, both the clinician, embryologist and patient, was blinded to the treatment allocation. The master randomisation list was held by ORIGIO a/s. All data analysis was performed externally by a Clinical Research Organization (CRO). During the interim analysis the statistician was blinded at all times to the media, which were presented as Medium A and Medium B. For the final statistical analyses, the codes for blinding were broken after database lock and patient classification."
Blinding of participants and personnel (performance bias) All outcomes	Low risk	"Participants and investigators were blinded to treatment allocation. Study media were packaged unidentifiably and labelled only with the randomization number. For each new patient recruited, the lowest available randomization number was used."
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Following email correspondence: "The clinicians performing the ultrasound scans were blinded to the treatment allocation at all times."
Incomplete outcome data (attrition bias) All outcomes	High risk	All dropouts accounted for, however the reasons given were not included in the predefined exclusion criteria.
Selective reporting (reporting bias)	Low risk	The authors confirmed through email correspondence that all outcomes were published.
Other bias	High risk	There is a potential conflict of interest regarding the funding and administration of this study as it was granted by the manufacturer of the intervention culture medium. However, the study authors have been very forthcoming with information regarding the trial, and there has been transparency. We assessed this trial as at high risk of other bias because the data provided in private email correspondence differ from the published data.

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Agerholm 2010	Wrong study design
Fawzy 2019	Wrong study design
Kinoshita 2019	Wrong study design
Min 2017	Wrong study design
Scarpellini 2011	Wrong intervention
Sfontouris 2013	Wrong study design
Shapiro 2003	Wrong study design
Siqueira 2016	Animal study

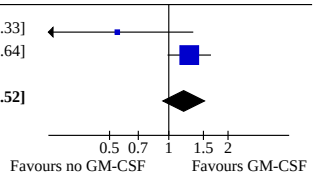
Comparison 1. GM-CSF-supplemented culture medium versus culture medium not supplemented with GM-CSF

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.1 Live birth	2	1432	Odds Ratio (M-H, Fixed, 95% CI)	1.19 [0.93, 1.52]
1.2 Miscarriage	2	1432	Odds Ratio (M-H, Fixed, 95% CI)	0.75 [0.41, 1.36]
1.3 Clinical pregnancy	3	1532	Odds Ratio (M-H, Fixed, 95% CI)	1.16 [0.93, 1.45]
1.4 Multiple gestation	2	1432	Odds Ratio (M-H, Fixed, 95% CI)	1.24 [0.73, 2.10]
1.5 Preterm birth	2	1432	Odds Ratio (M-H, Fixed, 95% CI)	1.20 [0.70, 2.04]
1.6 Birth defects	2	1432	Odds Ratio (M-H, Fixed, 95% CI)	1.33 [0.59, 3.01]
1.7 Aneuploidy	2	1432	Odds Ratio (M-H, Fixed, 95% CI)	0.34 [0.03, 3.26]
1.8 Stillbirth	2	1432	Odds Ratio (M-H, Fixed, 95% CI)	Not estimable

Analysis 1.1. Comparison 1: GM-CSF-supplemented culture medium versus culture medium not supplemented with GM-CSF, Outcome 1: Live birth

Study or Subgroup	GM-CSF Events	GM-CSF Total	Control Events	Control Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Rose 2020	11	50	17	50	11.1%	0.55 [0.23 , 1.33]
Ziebe 2013	168	654	145	678	88.9%	1.27 [0.99 , 1.64]
Total (95% CI)		**704**		**728**	**100.0%**	**1.19 [0.93 , 1.52]**
Total events:	179		162			

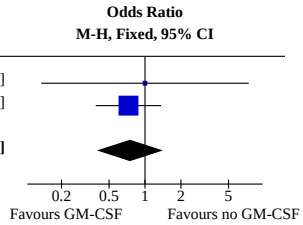
Heterogeneity: Chi² = 3.19, df = 1 (P = 0.07); I² = 69%
Test for overall effect: Z = 1.40 (P = 0.16)
Test for subgroup differences: Not applicable



Analysis 1.2. Comparison 1: GM-CSF-supplemented culture medium versus culture medium not supplemented with GM-CSF, Outcome 2: Miscarriage

Study or Subgroup	GM-CSF Events	GM-CSF Total	Control Events	Control Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Rose 2020 (1)	2	50	2	50	7.7%	1.00 [0.14 , 7.39]
Ziebe 2013	17	654	24	678	92.3%	0.73 [0.39 , 1.37]
Total (95% CI)		**704**		**728**	**100.0%**	**0.75 [0.41 , 1.36]**
Total events:	19		26			

Heterogeneity: Chi² = 0.09, df = 1 (P = 0.77); I² = 0%
Test for overall effect: Z = 0.95 (P = 0.34)
Test for subgroup differences: Not applicable



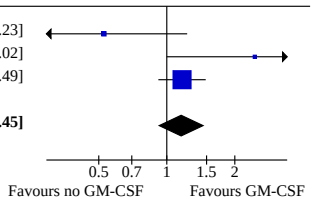
Footnotes

(1) One miscarriage in the control arm was at 17 weeks gestation, the remainder <12 weeks.

Analysis 1.3. Comparison 1: GM-CSF-supplemented culture medium versus culture medium not supplemented with GM-CSF, Outcome 3: Clinical pregnancy

Study or Subgroup	GM-CSF Events	GM-CSF Total	Control Events	Control Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Rose 2020	13	50	20	50	10.3%	0.53 [0.23 , 1.23]
Sbracia 2014	19	50	10	50	4.3%	2.45 [1.00 , 6.02]
Ziebe 2013	189	654	175	678	85.3%	1.17 [0.92 , 1.49]
Total (95% CI)		**754**		**778**	**100.0%**	**1.16 [0.93 , 1.45]**
Total events:	221		205			

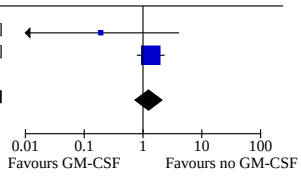
Heterogeneity: Chi² = 5.99, df = 2 (P = 0.05); I² = 67%
Test for overall effect: Z = 1.28 (P = 0.20)
Test for subgroup differences: Not applicable



Analysis 1.4. Comparison 1: GM-CSF-supplemented culture medium versus culture medium not supplemented with GM-CSF, Outcome 4: Multiple gestation

Study or Subgroup	GM-CSF Events	GM-CSF Total	Control Events	Control Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Rose 2020	0	50	2	50	9.9%	0.19 [0.01 , 4.10]
Ziebe 2013	31	654	24	678	90.1%	1.36 [0.79 , 2.34]
Total (95% CI)		**704**		**728**	**100.0%**	**1.24 [0.73 , 2.10]**
Total events:	31		26			

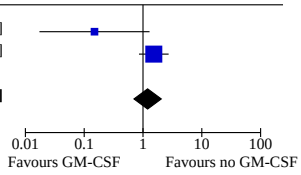
Heterogeneity: Chi² = 1.53, df = 1 (P = 0.22); I² = 35%
Test for overall effect: Z = 0.80 (P = 0.42)
Test for subgroup differences: Not applicable



Analysis 1.5. Comparison 1: GM-CSF-supplemented culture medium versus culture medium not supplemented with GM-CSF, Outcome 5: Preterm birth

Study or Subgroup	GM-CSF Events	GM-CSF Total	Control Events	Control Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Rose 2020	1	50	6	50	23.9%	0.15 [0.02 , 1.29]
Ziebe 2013	29	654	20	678	76.1%	1.53 [0.85 , 2.73]
Total (95% CI)		**704**		**728**	**100.0%**	**1.20 [0.70 , 2.04]**
Total events:	30		26			

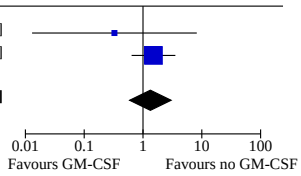
Heterogeneity: Chi² = 4.25, df = 1 (P = 0.04); I² = 76%
Test for overall effect: Z = 0.66 (P = 0.51)
Test for subgroup differences: Not applicable



Analysis 1.6. Comparison 1: GM-CSF-supplemented culture medium versus culture medium not supplemented with GM-CSF, Outcome 6: Birth defects

Study or Subgroup	GM-CSF Events	GM-CSF Total	Control Events	Control Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Rose 2020	0	50	1	50	14.6%	0.33 [0.01 , 8.21]
Ziebe 2013	13	654	9	678	85.4%	1.51 [0.64 , 3.55]
Total (95% CI)		**704**		**728**	**100.0%**	**1.33 [0.59 , 3.01]**
Total events:	13		10			

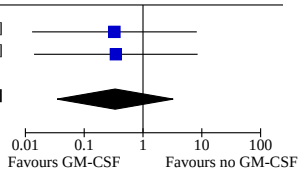
Heterogeneity: Chi² = 0.81, df = 1 (P = 0.37); I² = 0%
Test for overall effect: Z = 0.70 (P = 0.49)
Test for subgroup differences: Not applicable



Analysis 1.7. Comparison 1: GM-CSF-supplemented culture medium versus culture medium not supplemented with GM-CSF, Outcome 7: Aneuploidy

Study or Subgroup	GM-CSF Events	GM-CSF Total	Control Events	Control Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Rose 2020	0	50	1	50	50.2%	0.33 [0.01 , 8.21]
Ziebe 2013	0	654	1	678	49.8%	0.35 [0.01 , 8.49]
Total (95% CI)		**704**		**728**	**100.0%**	**0.34 [0.03 , 3.26]**
Total events:	0		2			

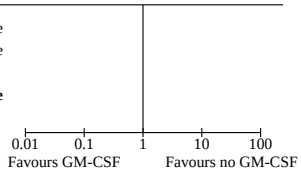
Heterogeneity: Chi² = 0.00, df = 1 (P = 0.98); I² = 0%
Test for overall effect: Z = 0.94 (P = 0.35)
Test for subgroup differences: Not applicable



Analysis 1.8. Comparison 1: GM-CSF-supplemented culture medium versus culture medium not supplemented with GM-CSF, Outcome 8: Stillbirth

Study or Subgroup	GM-CSF Events	GM-CSF Total	Control Events	Control Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Rose 2020	0	50	0	50		Not estimable
Ziebe 2013	0	654	0	678		Not estimable
Total (95% CI)		**704**		**728**		**Not estimable**
Total events:	0		0			

Heterogeneity: Not applicable
Test for overall effect: Not applicable
Test for subgroup differences: Not applicable



APPENDICES

Appendix 1. Cochrane Gynaecology and Fertility Group Specialised Register search strategy

PROCITE platform

Searched 15 October 2019

Paper 4: VALUE study: a protocol for a qualitative semi-structured interview study of IVF add-ons use by patients, clinicians and embryologists in the UK and Australia

Following the completion and publication of papers 1, 2 and 3, my mind turned to exploring the reasons why patients and professionals in the IVF sphere opt to use add-ons. By this time (the beginning of 2020), there was a growing body of evidence to suggest that add-ons added no advantage in terms of livebirth, over IVF without add-ons, so it seemed curious to me that they were still as popular as ever [10, 106]. Academic debate at the time suggested that there was more to the story than naïve, poorly informed patients accepting add-ons from clinics who saw the pecuniary advantages of offering add-ons, and I wanted to design a study that revealed the driving factors behind their use [36, 107]. Section 6.7 reveals the existing qualitative evidence behind add-on use, which was limited to a survey of 1017 patients commissioned by the HFEA [27]. There had been no semi-structured interview studies exploring the motivating factors for using add-ons from a patient and professional perspective.

My quantitative work, and the presentation of findings at meetings around the world had introduced me to people from different disciplines who were interested in the same question. One important collaborator was Dr Sarah Lensen, whom I had co-authored a Cochrane review with on the topic of endometrial scratching for IVF (not included in this thesis); another example of an add-on [108]. Dr Lensen was based at the University of Melbourne and was keen to explore the views of those undergoing IVF in Australia alongside those in the UK. We were intrigued to understand whether the views of professionals and patients would differ between the two settings, or whether they would be harmonious given that the process of IVF is broadly the same in both countries. Therefore, we resolved to design a study.

I gathered a group of co-authors with qualitative expertise, and clinical and laboratory experience in IVF, to develop the protocol for a semi-structured interview study of IVF add-ons use by patients, clinicians and embryologists in the UK and Australia. We named the study 'The VALUE Study', and the protocol was published in BMJ Open following peer review [109].

BMJ Open

VALUE study: a protocol for a qualitative semi-structured interview study of IVF add-ons use by patients, clinicians and embryologists in the UK and Australia

Sarah C Armstrong [1], Sarah Lensen [2], Emily Vaughan [3], Elaine Wainwright [4], Michelle Peate [2], Adam H Balen [5], Cynthia M Farquhar [6], Allan Pacey [1]

To cite: Armstrong SC, Lensen S, Vaughan E, *et al*. VALUE study: a protocol for a qualitative semi-structured interview study of IVF add-ons use by patients, clinicians and embryologists in the UK and Australia. *BMJ Open* 2021;**11**:e047307. doi:10.1136/bmjopen-2020-047307

► Prepublication history and supplemental material for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2020-047307).

Received 27 November 2020
Revised 28 April 2021
Accepted 29 April 2021

For numbered affiliations see end of article.

Correspondence to
Dr Sarah C Armstrong; scarmstrong1@sheffield.ac.uk

ABSTRACT

Introduction For couples undergoing assisted reproduction, a plethora of adjuncts are available; these are known as 'add-ons'. Most add-ons are not supported by good quality randomised trial evidence of efficacy, with some proven to be ineffective. However, estimates suggest that over 70% of fertility clinics provide at least one add-on, often at extra cost to the patient. This study has three aims. First, to undertake a survey of in vitro fertilisation (IVF) clinics in the UK to ascertain which add-ons are being offered and at what cost. Second, to undertake qualitative semi-structured interviews of patients, clinicians and embryologists, to explore their opinions and beliefs surrounding add-ons. Third, to review the interpretation of the Human Fertilisation and Embryology Authority traffic light system, to better understand the information required by IVF patients, clinicians and embryologists when making decisions about add-ons.

Methods and analysis All UK IVF clinics will be contacted by email and invited to complete an online survey. The survey will ask them which add-ons they offer, at what cost per cycle and how information is shared with patients. Semi-structured interviews will be conducted in the UK and Australia with three groups of participants: (i) fertility patients; (ii) clinicians and (iii) embryologists. Participants for the interviews will be recruited via social media channels, website adverts, email and snowball sampling. Up to 20 participants will be recruited for each group in each country. Following an online consent process, interviews will be conducted via video-conferencing software, transcribed verbatim and data subjected to inductive thematic analysis.

Ethics and dissemination Ethical approval has been granted by the Universities of Sheffield, Bath Spa and Melbourne. Findings will be published in a peer-reviewed journal and disseminated to regulatory bodies in the UK and Australia. A lay summary of findings will be shared via Fertility Network, UK.

Strengths and limitations of this study

- ► VALUE is the first study to explore, through in-depth, semi-structured qualitative interviews, the driving factors behind the use of in vitro fertilisation add-ons by patients, clinicians and embryologists.
- ► Early and in-depth patient and public involvement was used to ensure the study's acceptability, use and relevance to the target population.
- ► Purposive sampling in two different healthcare systems, encompassing both private and state funded fertility services will be conducted to capture a wide range of patient, clinician and embryologist experiences.
- ► We will mitigate the risk that interviewees adjust their responses in light of interviewers being medical professionals working in fertility, by training interviewers and highlighting their neutrality at the start of the interview.
- ► Recruitment via social media may limit the recruitment to a particularly motivated, engaged and media literate group of participants.

INTRODUCTION

Undergoing fertility treatment can involve physical, mental and financial stress, with patients often desperate to explore any options which might confer greater chance of treatment success. Over recent years there has been an increase in medical and non-medical in vitro fertilisation (IVF) treatment adjuncts available; these are commonly known as 'add-ons'.[1]

The UK regulatory body for assisted reproduction, Human Fertilisation and Embryology Authority (HFEA) describes add-ons as 'optional extras you may be offered on top of your normal fertility treatment, often at an additional cost. They're sometimes emerging techniques that may have shown some promising results in initial studies, or they may have been around for a number of years, but haven't necessarily been proven to improve pregnancy or birth rates'.[2] In some

cases, add-ons have become 'routine practice', with costs embedded into the fertility package fee as opposed to being charged in addition. For example, embryo incubation using time-lapse technology is routine in some centres, and is optional in others.[3]

Assisted reproduction is a fast-paced area of medicine, with growing demand for treatment, accompanied by rapid innovation.[4] There is growing recognition from the assisted reproduction community of the paucity of evidence surrounding the use of add-ons, most add-ons are not supported by good quality randomised trial evidence.[2 4 5] There has been much speculation and interest in the driving forces behind add-ons' popularity, both factors of supply (IVF clinics offering or advertising add-ons) and demand (IVF patients requesting add-ons).[6–9] However, thus far, there has been no research specifically focused on why patients, clinicians and embryologists opt to offer or use them. There is a lack of research into the views of these groups, particularly surrounding their interpretation of evidence of efficacy of add-ons, and information sources for decision-making about their use.

The HFEA have provided a website designed for patients regarding add-ons, with a traffic light rating of red, amber and green to denote the quality of evidence on efficacy and safety of use. However, there is little information about the utility of this system, and how patients interpret the different traffic light colours. For patients in Australia, there is no such similar patient directed website.

The practice of medicine rests open three main principles of ethics. First beneficence (the moral obligation to act for the benefit of others), second non-maleficence (requires that medical professionals prevent harm to the health and well-being of patients) and third autonomy (patients have a right to self-determination, or choices in their care). IVF add-ons raise an interesting ethical dilemma, given that add-ons have not been conclusively proven to make IVF more effective, or reduce the risk of harms, such as miscarriage. However, denying a patient's autonomy in opting to use add-ons may also be seen as unethical.[10] In order for autonomy to be executed, the patient must have informed consent, that is, an understanding of the potential benefits and risks of any given add-on.[11]

The VALUE study is important because it will help inform how patients, clinicians and embryologist weigh up the factors that relate to these three pillars of medical ethics when thinking about their experience of using add-ons. It will also explore what information is important to these three stakeholder groups when participating in informed consent. It is hoped that the information from VALUE will support caregivers to provide the best possible ethical care to their patients, and improve the quality of the informed consent process for patients to better support them in making informed decisions.

Aims

This study aims to first undertake a survey of IVF clinics in the UK to ascertain which add-ons are being offered and at what cost to the patient. Second, through qualitative semi-structured individual interviews of assisted reproduction patients, clinicians and embryologists, it will then explore the opinions and beliefs surrounding add-ons and any evidence for efficacy. Finally, the interviews will also be used to review the interpretation of information provided by regulatory bodies in order to optimise provision of information for these groups when making future decisions about IVF add-ons.

Objectives

1. Provide information on availability of add-ons in UK and the costs that are charged for them.
2. To understand how people make decisions about using or recommending IVF add-ons.
3. To understand where information about add-ons is sought, and to understand the role and importance of information such as safety and effectiveness when considering their use.
4. To explore participants' understanding and interpretation of the HFEA traffic light system for add-ons.

METHODS AND ANALYSIS

Study design

Part 1: UK clinic survey

A list of all licensed IVF clinics in the UK will be compiled using public data from the HFEA website.[12] Then, the medical director of each clinic will be contacted and invited to complete an online survey. The online survey will ask the following questions: (i) number of IVF and intracytoplasmic sperm injection (ICSI) cycles performed in year January 2019–January 2020; (ii) whether the clinic treats National Health Service (NHS) and/or private privates; (iii) which add-ons they offer at their clinic; (iv) the cost per-cycle to patients for the use of each add on; (v) whether written information regarding add-ons is offered, and the form of this information (ie, published by the clinic, or published by Industry) and (vi) whether any of the listed add-ons are included as part of an NHS funded cycle, or a private cycle (ie, are used routinely). The clinic survey is only taking place in the UK because a similar survey has already taken place in Australia.[13 14]

In order to improve the response rates, we will use an evidence-based strategy of survey recruitment. A prenotification email will be sent to the medical director 1 week prior to the survey opening outlining the survey and informing them that following completion of the survey they can choose to be entered into a prize draw for three £50 Love2shop vouchers. One week later the link to the survey will be emailed with a follow-up email at week 2. In week 3 or 4 we will send a further follow-up email and phone call to the clinic. In week 6 the survey will close and the prize draw winners announced. Those who complete the survey will be sent a follow-up email thanking them,

and asking if they would be happy to share their patient information leaflets on add-ons with us.

Part 2: qualitative in-depth semi- Research Ethics approval was not required for the PPI interviews

Interview schedule design

The interview schedule was iteratively developed with our patient and public involvement (PPI) group and is underpinned by preidentified domains of interest within the academic and grey literature: (i) sources of information; (ii) the decision-making process and (iii) importance of evidence. The HFEA 2018 pilot national fertility patient survey[15] revealed important areas where information on add-ons is lacking including where patients seek information from, whether information on the efficacy, cost-effectiveness and side effects of add-ons is provided. Through our semi-structured interview design, we will therefore explore participants' personal experiences in deciding whether to use or offer add-ons. We will explore factors that are important to them in making this decision and their sources of information as well as explore how participants in the UK and Australia interpret the HFEA's traffic light system and its role to guiding decision-making around add-ons.[2]

Patient and public involvement

PPI has taken place to tailor the study design to ensure it is addressing important research questions and that the study materials are presented in a clear and understandable format. A 'study-focussed framework' will be supported, whereby two patients will follow the research cycle from initial PPI stages through to disseminating findings and achieving impact.[16]

The PPI process included patients, clinicians and embryologists with two PPI groups in both the UK and Australia. PPI sessions were conducted separately in the UK and Australia due to subtle differences in demographic questions. Due to concerns about a power dynamic between professionals and patients possibly preventing participants from feeling able to free express themselves[17] we held one focus group for patients in each location and a separate group for clinicians and embryologists. Participants were recruited through social media and engaged in an hour-long teleconference with other panel members and the research team. Each participant was provided with a draft set of interview questions ahead of the meeting and were asked to comment on them during the teleconference. In addition, they were asked to review the consent form, the information sheet and the study website. A series of questions about the coordination and practical running of the study were also posed.

The teleconferences were recorded following agreement from participants and followed strict General Data Protection Regulation (GDPR) guidance. Participants were offered either a £20 love2shop e-voucher or a $50 Australian supermarket voucher as a thank you for their time. All participants consented to being acknowledged in resulting publications.

Patients PPI

Two patient participants were sought in each country and patients were required to have undergone assisted reproduction (IVF or ICSI) in the past 2 years.

As a result of patient PPI, the wording of some of the qualitative questions was altered, and prompts were added where necessary. The panel felt that the patient interviews should be divided in two to enable time for the participant to browse the HFEA website prior to questions on this topic. The feedback was that asking the participant to familiarise themselves with the website during the hour-long interview was too stressful and would put the participant under undue pressure. The panel was in agreement that two shorter interviews were no more onerous or inconvenient than 1-hour long interview. In response to feedback, a table of 'commonly used terms' was added to the preinterview demographic questions (table 1) and the website was altered slightly to improve readability.

Embryologists and clinicians PPI

A minimum of two embryologists and two clinicians in each country were sought for PPI. Professional databases were checked to ensure that those taking part were registered doctors or embryologists delivering fertility treatment in the UK or Australia. In the UK, two embryologists and one reproductive medicine specialist doctor joined the teleconference, and a separate teleconference was undertaken with one other reproductive medicine specialist doctor due to clinical commitments. In Australia, one PPI panel was convened, consisting of two embryologists and two reproductive medicine specialist doctors.

As a result of PPI, the preinterview demographic questions were altered to accurately reflect clinicians' job titles and questions regarding ethnicity and religion were removed. Following panel input, the questions were reordered to improve the flow of the interview and the wording of some questions changed to remove any negative connotations towards add-ons. In addition, lay descriptions of add-ons were added to the website following feedback that this would enable patients to more easily identify which add-ons they had used or considered. This panel explained that part 1 and part 2 of the interview should not be split into two separate interviews for because it was too time consuming and may deter clinicians and embryologists from participating. The feedback was that being given the chance to look at the HFEA website prior to the interview would be preferable to being asked to look at it mid-interview. The study protocol has been altered to reflect these changes.

VALUE study eligibility criteria

Inclusion criteria

Patients

Adult women, men or couples (18+ years of age); who have undergone IVF or ICSI in the past 2 years (any number of cycles); publicly funded (NHS funded in the UK, or Medicare in Australia) or privately funded; using

Table 1 Table of commonly used terms

Term we use	What it stands for	Description of term
IVF	In vitro fertilisation	The process of stimulating the woman's ovaries, collection of eggs, mixing of egg/s with sperm to make embryos, incubation of embryos and replacement of embryos into the woman.
ICSI	Intracytoplastic sperm injection	The process described above, except instead of mixing the woman's eggs with sperm, a single sperm is selected to be injected into the egg.
A cycle of IVF or ICSI		One cycle of IVF or ICSI includes all the steps involved in IVF or ICSI described above, plus the replacement of any resulting embryos from that cycle (fresh or frozen transfer). A cancelled cycle, or a cycle where no embryos can be transferred both count as a cycle.
Embryo transfer		Embryo transfer refers to the process of replacing an embryo that results from an IVF or ICSI cycle. Embryo transfers can be single, where one embryo is transferred, or double, where two embryos are transferred. No matter how many embryos are replaced, these all count as *one* embryo transfer procedure.
Ovulation induction		The process of stimulating the ovaries to release an egg each month. This can be done using tablets such as clomiphene citrate, or injections. The couple conceive the baby through sexual intercourse.

ICSI, intracytoplasmic sperm injection; IVF, in vitro fertilisation.

either autologous oocytes and sperm, or donor oocytes and sperm; and who have considered using, or had used, one or more add-ons as part of their treatment.

Clinicians

Registered doctors involved in the care of patients or couples undergoing assisted reproduction. Doctors can be consultant fertility specialists, staff-grade fertility specialists or General Practitioners (GPs) who specialise in reproductive medicine and work in fertility clinics.

Embryologists

Registered embryologists involved in decisions regarding the assessment of embryos, who have direct interaction with patients or couples undergoing IVF or ICSI.

Exclusion criteria

Those who are non-fluent English speakers owing to the financial cost and logistics of arranging appropriate translation assistance during interviews. Those who are donating oocytes or sperm therefore undergoing assisted reproduction themselves.

Recruitment

In both countries, patient participants will be recruited via broad ranging social media advertising, including the websites and social media of patient support groups such as Fertility Network in the UK. Recruiting participants in this way aims to include those from a diverse range of socioeconomic and ethnic backgrounds and geographic locations. Additionally, this approach should include patients or couples who are at varying stages of their IVF experience, including those undergoing their first cycle, to those embarking on repeated cycles and those who have and have not experienced success from IVF.

Clinicians and embryologists will also be recruited via websites, newsletters and social media, but in this case with the assistance of professional bodies such The British Fertility Society and the Association of Reproductive and Clinical Scientists in the UK. The Fertility Society of Australia (FSA) will advertise the study in Australia. Recruiting in this manner enables sampling from a broad geographical range of clinicians and embryologists, working in different clinics, with difference practices.

Both patient participants and professionals may also be recruited using a snowballing technique, where at the end of the interview existing participants are asked to nominate others to be approached for participation. Snowball sampling is a valid technique for participant recruitment in qualitative research and allows researchers to reach populations who otherwise would have been hard to reach.

Interested participants in both the UK and Australia will be directed to the VALUE study website (www.value-study.org) where they can express interest in the study using the 'contact us' form embedded in the 'patient' webpage (www.valuestudy.org/for-patients) and the 'professionals' webpage (www.valuestudy.org/for-professionals). Researchers will then confirm eligibility and obtain informed consent via a secure online form, and schedule a time to undertake the interview (figure 1). A list of examples of add-ons is provided on the website, and has been published as online supplemental table 1.

Sampling strategy and size

Approximately 60 interviews will be conducted in both the UK and Australia (20 per participant group) and the collection and analysis of data will be done iteratively to consider when sufficiently robust codes and themes have been created.[18] A sample of n=20 per group has been based on similar studies,[18–22] however, it is recognised that deep analysis is more important than number of interviews and sample size will be determined by data

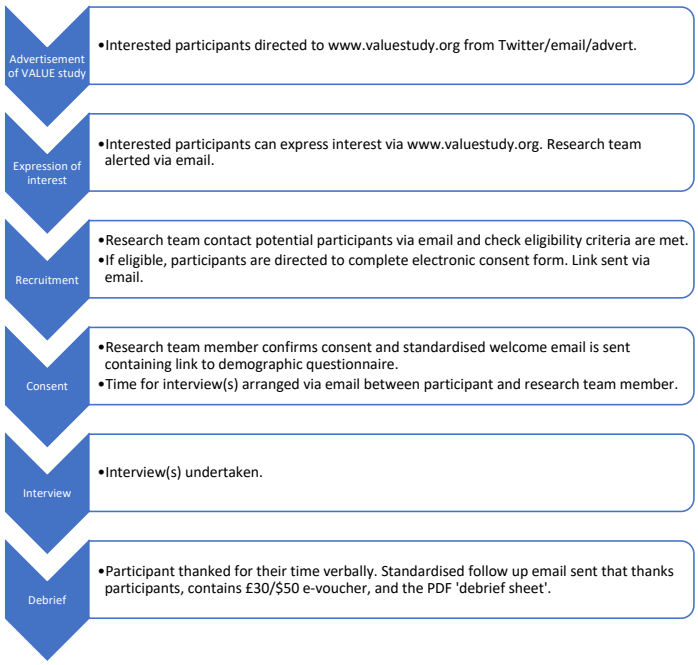


Figure 1 Flow of participants through the VALUE study.

saturation when no further themes are generated.[23] Couples who wish to be interviewed together will be considered as n=1 towards the sample size, however, if they wish to be interviewed separately, then they will be considered as two participants.

There will be purposive sampling within our inclusion criteria, to provide a variety of participants so that readers can assess transferability to a wider population of clinicians, embryologists and patients undergoing IVF.[23] The VALUE study aims to interview patients who have had government subsidised as well as privately funded cycles. It also aims to interview clinicians and embryologists working in the public and private sector and to include both senior and junior staff, the importance of which was highlighted by the PPI panel. Timely thematic analysis of the first 20 interviews will be undertaken and if an appropriate spread of patients and professionals has not been included, we will use a sampling framework for maximum variation for the next 20 interviews prior to consent stage (table 2).

There will be complete transparency with potential participants that we may not need to interview them depending on their answers, but that we appreciate their interest and time in getting in touch. We will explain that

Table 2 Framework sampling questions

Question	Patients	Professionals
Targeted questions to be asked prior to consent for potential interviewees if analysis of the first 10 interviews is suggestive of lack of diverse respondents.		
1	Please can we ask how many cycles of IVF or ICSI you have undergone?	Please can you share your clinical title?
2	Please can we ask whether you have received NHS funded or privately funded IVF or ICSI? Perhaps you have had both?	How many years have you worked in the discipline of reproductive medicine? (clinicians only)
3		Please can we ask whether you see and treat NHS funded or privately funded IVF or ICSI? Perhaps you treat had both? Please can you explain.

ICSI, Intracytoplasmic sperm injection; IVF, in vitro fertilisation; NHS, National Health Service.

their answers to these questions will not be recorded as part of the study.

Interviews

Interviews will be held remotely using video-conferencing software and will be recorded to aid transcription. Patients will be interviewed twice. First with nine questions, lasting approximately 45 min, following this they will be asked to review the HFEA website prior to the second interview of eight questions, lasting approximately 30 min. Clinicians and embryologists will participate in a single interview of approximately 60 min interview containing 15 questions. Interview schedule will not be made available until after all interviews have been conducted so as not bias responses from participants having seen the questions in advance of the interviews from this publication. However, they will be available on request after the interviews have been completed.

The interview will be conducted by members of the research team who have undergone training in conducting semi-structured interviews about potentially upsetting topics. At the beginning of the interview participants will be asked to try to avoid mentioning their names or those of IVF clinics or staff; although, the onus will be on the research team to fully anonymise subsequent transcriptions. Participants will be reminded that involvement in the research is entirely voluntary and that they can withdraw at any point during the interview. For clinicians and embryologists, they will be reminded prior to the interview that it is not a test of their clinical knowledge and that all information shared will be kept confidential.

Patients, clinicians and embryologists will be offered a £30 e-Gift Card for love2shop or a $50 Australian supermarket voucher as a thank you for their time. National Institute for Health Research recommend rewarding public participation in research and vouchers of this value are an appropriate thank you for their time.[24]

Transcription

Audio recordings will be kept on secure servers and will undergo transcription by a third-party confidential and secure password protected transcription service. Transcription of audio recording will be checked by the in-country research team to ensure that all identifiably data are removed and the transcript deidentified.

Analysis

The clinic survey data and demographic data from interview participants will be exported to a password protected Excel spreadsheet and will undergo descriptive analysis.

The interview data will undergo inductive thematic analysis to identify descriptive labels (codes) through repeated analysis. Codes will be used to group data into subthemes and further overarching themes to produce a complex account of data that is both rich and detailed and appropriate to purpose.[25] Thematic analysis covers a range of epistemological and ontological decisions; we will use it as a 'contextualist' method within a critical realist paradigm.[24 26] Thematic analysis is an appropriate framework to use for data collection and analysis as it enables a detailed account of data that is both descriptive and interpretive.[27] It can acknowledge how people make sense of their experiences as well as how broad social structures interact with these.[28] It should enable an overarching understanding of the experience of the three groups being interviewed in this study.

Analysis will begin with listening to interview recordings and reading each transcript many times to establish familiarity with the whole interview and become immersed in the data, noting initial interpretations. Initial codes (salient features) will be created, to arrange the data into meaningful segments. In the main analytic phase, different codes will be reviewed and combined to form broader themes. The first set of coding and themes will be reflexively considered until consensus is reached to define, name and exemplify all themes.

Reducing bias

We acknowledge that some of the authors of this study have been involved in the publication of evidence that does not support the routine use of IVF add-ons. Every effort has been made to be aware of this and mitigate it in the planning, execution and analysis of VALUE. The interview questions have undergone a robust PPI process, and were also subject to close scrutiny by the ethical review bodies at the Universities of Sheffield and Melbourne. Changes were made to the wording of questions as a result of feedback from these processes where there was felt to be any implied judgement. In addition, interviewers have undergone the planned training on undertaking qualitative interviews. Furthermore, double coding on a proportion of the interview data is being undertaken by Dr Wainwright, who was brought into the project as someone experienced in PPI and qualitative methods but who has not been involved in the publication of evidence that does not support the routine use of IVF add-ons.

Data protection

All data from the VALUE study will be stored securely on password protected encrypted servers. No hard copies of data will be kept. Demographic data, interview recordings and transcripts will be stored in the country of origin (UK participants' data will be stored at the University of Sheffield, and Australia participants' data will be stored at the University of Melbourne). Only deidentified interview transcripts will be shared between the UK and Australia sites and uploaded to form part of qualitative analysis on using secure password protected analytic application. All recordings will be deleted after the transcripts have been checked by the respective country's research team and are fully anonymised.

The VALUE study will not release anonymised transcripts for future research. This decision has been made in light of the sensitive nature of the topic and in response to PPI feedback which suggested that participants may

feel inhibited to speak openly due to the nature of their stories being potentially identifiable.

Ethics and dissemination

Ethical considerations

Research ethics approval was not required for the PPI phase of this study. In the UK, ethical approval has been obtained from the University of Sheffield (reference: 036268) and Bath Spa University (BSU-20-205) and in Australia ethical approval has been obtained from University of Melbourne (2057434.1). Participants will receive comprehensive information leaflets prior to the study and participants will undergo an online written consent process prior to interview with all participant information treated confidentially. Participants are free to withdraw from the study at any time.

Output and dissemination

Results will be published in a peer-review journal and disseminated to regulatory bodies such as the HFEA, The National Institute for Care and Clinical Excellence, the Victorian Assisted Reproduction Treatment Authority (Australia) and the FSA in order to help shape future information about IVF add-ons. A lay summary of findings will be shared with participants from our PPI panel, patients interviewed and via fertility UK to highlight results from the work to the wider public.

The VALUE study aims that rich qualitative data from this research will help improve communication of clinical impact of IVF add-ons to patients in future. It also hopes to analyse understanding and interpretability of a traffic light system in conveying information to patients and professionals, generating information which can be used to inform the use of the traffic light system in regulatory bodies in other countries.

Limitations

Recruitment via social media aims to facilitate purposive sampling of participants from different geographical locations, and different socioeconomic backgrounds. However, a significant limitation of this approach is that it may attract a particularly information technology literate, motivated group of individuals. One concern is that patient participants who are looking at fertility websites and social media outlets, may be more likely to be further into their fertility journey, and less likely to be undergoing their first cycle of assisted reproduction. We aim to ameliorate this by using a variety of social media outlets, plus websites and emails.

Researchers involved in VALUE have been involved in novel research that has thrown into question the rationale of the routine use of some add-ons. This involvement in research may be known to some participants, and one limitation is the risk that participants may alter their responses in light of this. The qualitative interviews have been carefully designed to demonstrate equipoise and to not introduce any form of value judgement.

Author affiliations
[1]Oncology and Metabolism, The University of Sheffield, Sheffield, UK
[2]Obstetrics and Gynaecology, The University of Melbourne, Melbourne, Victoria, Australia
[3]Academic Women's Health Unit, North Bristol NHS Trust, Bristol, UK
[4]Psychology Department, Bath Spa University, Bath, UK
[5]Obstetrics and Gynaecology, Leeds Teaching Hospitals NHS Trust, Leeds, UK
[6]Obstetrics and Gynaecology, The University of Auckland, Auckland, New Zealand

Twitter Sarah C Armstrong @ArmstrongSarahC, Cynthia M Farquhar @CindyFarquhar and Allan Pacey @allanpacey

Acknowledgements Our heartfelt thanks to our PPI panel who gave up their time to contribute to the development of VALUE. UK: Isabella Dash, Jennifer Nisbett, Hannah Reid, Ally Richardson, Victoria Thomas, Bassel Wattar. Australia: Katherine Gobbi, Hilary Smith, Natasha Devetak, Alex Polyakov, Anna Ninnis, Lisa Lee, Vadim Mirmilstein.

Contributors SCA will take the role of principle investigator. In the UK and Australia respectively, SCA and SL were responsible for applying for ethics approval. They planned and coordinated the local PPI and wrote the initial draft of questions for the semi-structured interviews. They will conduct the interviews, collate the results and be involved in the thematic analysis of results, and write up the results for publication. EV will be involved at all stages of the study. She was involved in the PPI and formulation of questions for the semi-structured interviews. EV will conduct the interviews and help in the thematic analysis of results. She will be involved in write up and editing of the drafts of the VALUE study for publication. EW will be the qualitative research expert for the study. She advised and participated in the PPI, helped with the methodological rigor of planning questions, and commented on ethics. EW will oversee the thematic analysis of results and edit drafts of the VALUE study for publication. CMF, AP, MP and AHB are experienced researchers who contributed to the design of the study. They will be involved in interpretation of results, as well as editing drafts of the study for publication.

Funding This research was funded by an Obstetrics and Gynaecology Innovation Grant from the University of Melbourne Australia. MP is supported by a University of Melbourne Department of Obstetrics and Gynaecology MCR Fellowship.

Competing interests SCA, SL, EV, MP, CMF and EW have no competing interests. AP and AHB are trustees of the British Fertility Society (Charity number 1075661) and AP is also a trustee of the Progress Educational Trust (Charity number 1139856).

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Sarah C Armstrong http://orcid.org/0000-0001-7308-520X
Sarah Lensen http://orcid.org/0000-0002-1694-1142
Emily Vaughan http://orcid.org/0000-0001-8843-2497
Elaine Wainwright http://orcid.org/0000-0001-5678-4412
Michelle Peate http://orcid.org/0000-0003-2903-4688
Adam H Balen http://orcid.org/0000-0001-5938-8319
Cynthia M Farquhar http://orcid.org/0000-0002-3685-3553
Allan Pacey http://orcid.org/0000-0002-4387-8871

REFERENCES

1 RCOG. Additional treatments www.rcog.org.uk: RCOG, 2020. Available: https://www.rcog.org.uk/en/patients/fertility/treatment/additional/
2 HFEA. Treatment add-ons, 2020. Available: www.hfea.gov.uk/treatments/explore-all-treatments/treatment-add-ons/
3 Armstrong S, Bhide P, Jordan V, *et al*. Time-Lapse systems for embryo incubation and assessment in assisted reproduction. *Cochrane Database Syst Rev* 2019;5:CD011320.
4 Armstrong S, Atkinson M, MacKenzie J, *et al*. Add-ons in the laboratory: hopeful, but not always helpful. *Fertil Steril* 2019;112:994–9.
5 Heneghan C, Spencer EA, Bobrovitz N, *et al*. Lack of evidence for interventions offered in UK fertility centres. *BMJ* 2016;355:i6295–4.
6 Tsigdinos PM. *The big IVF add-on racket*. New York: New York Times, 2019.
7 Harper J, Jackson E, Sermon K, *et al*. Adjuncts in the IVF laboratory: where is the evidence for 'add-on' interventions? *Hum Reprod* 2017;32:485–91.
8 Spencer EA, Mahtani KR, Goldacre B, *et al*. Claims for fertility interventions: a systematic assessment of statements on UK fertility centre websites. *BMJ Open* 2016;6:e013940.
9 Macklon NS, Ahuja KK, Fauser B. Building an evidence base for IVF 'add-ons'. *Reprod Biomed Online* 2019;38:853–6.
10 Zemyarska MS. Is it ethical to provide IVF add-ons when there is no evidence of a benefit if the patient requests it? *J Med Ethics* 2019;45:346–50.
11 Beauchamp BCJ. *Principles of biomedical ethics*. 7th Edition. Oxford, UK: Oxford University Press, 2013.
12 HFEA. Welcome to the HFEA, 2020. Available: https://www.hfea.gov.uk/
13 Lensen S, Chen S, Goodman L, *et al*. Ivf add-ons in Australia and New Zealand: a systematic assessment of IVF clinic websites. *Aust N Z J Obstet Gynaecol* 2021. doi:10.1111/ajo.13321. [Epub ahead of print: 16 Feb 2021].
14 Lensen S, Hammarberg K, Polyakov A, *et al*. How common is add-on use and how do patients decide whether to use them? a national survey of IVF patients. *Hum Reprod* 2021. doi:10.1093/humrep/deab098. [Epub ahead of print: 04 May 2021].
15 HFEA. Pilot national fertility patient survey hfea.co.uk, 2018. Available: https://www.hfea.gov.uk/media/2702/pilot-national-fertility-patient-survey-2018.pdf
16 Greenhalgh T, Hinton L, Finlay T, *et al*. Frameworks for supporting patient and public involvement in research: systematic review and co-design pilot. *Health Expect* 2019;22:785–801.
17 Nimmon L, Stenfors-Hayes T. The "Handling" of power in the physician-patient encounter: perceptions from experienced physicians. *BMC Med Educ* 2016;16:114.
18 Baker SE, Edwards R, Doidge M. How many qualitative interviews is enough? Expert voices and early career reflections on sampling and cases in qualitative research. *National Centre for Research Methods Review Paper* 2012.
19 Wainwright E, Looseley A, Mouton R, *et al*. Stress, burnout, depression and work satisfaction among UK anaesthetic trainees: a qualitative analysis of in-depth participant interviews in the satisfaction and wellbeing in anaesthetic training study. *Anaesthesia* 2019;74:1240–51.
20 O'Reilly M, Parker N. 'Unsatisfactory Saturation': a critical exploration of the notion of saturated sample sizes in qualitative research. *Qualitative Research* 2013;13:190–7.
21 Saunders B, Sim J, Kingstone T, *et al*. Saturation in qualitative research: exploring its conceptualization and operationalization. *Qual Quant* 2018;52:1893–907.
22 Nelson J. Using conceptual depth criteria: addressing the challenge of reaching saturation in qualitative research. *Qualitative Research* 2017;17:554–70.
23 Patton MQ. *Qualitative research & evaluation methods*. Fourth edition. Thousand Oaks: SAGE Publishing, 2014.
24 Braun V, Clarke V. *Successful qualitative research: a practical guide for beginners*. Thousand Oaks: SAGE Publishing, 2013.
25 Craig P, Dieppe P, Macintyre S, *et al*. Developing and evaluating complex interventions: the new medical Research Council guidance. *Int J Nurs Stud* 2013;50:587–92.
26 Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;3:77–101.
27 Boyatzis RE. *Transforming qualitative information: thematic analysis and code development*. Thousand Oaks: SAGE Publishing, 1998.
28 Archer MBR, Collier A, Lawson T. *Critical realism: essential readings*. London: Routledge, 2013.

Supplementary Table 1: Table of examples of add-ons

Add-ons for eggs, sperm, embryos	
Egg activation	Trying to stimulate egg activation with a substance called calcium ionophore which the embryo is treated with.
Intrauterine culture	Fertilising the egg in the lab then transferring the embryo in an intrauterine culture device into the womb where it stays for a few hours during embryo development. It is then removed and the embryo put back in an incubator.
Embryo Glue	EmbryoGlue contains a substance called hyaluronan, which aims to improve the chance of the embryo implanting in the womb.
Elective freeze-all cycles	Creating embryos then freezing them all so none are transferred in the 'fresh' cycle.
Assisted hatching	Using acid, lasers or other tools to thin or make a hole in the zona pellucida (the layer surrounding an embryo) in order to try to help the embryo 'hatch'.
Preimplantation genetic testing (PGT)	Checking the genes or chromosomes of the embryos for abnormalities before deciding which embryo to transfer.
Sperm DNA test	Analysing the DNA in sperm for damage.
Embryo culture media containing growth factors (BlastGen, EmbryoGen)	Adding growth factors to the solution used to bathe the embryos as they grow in the lab.
Intra-cytoplasmic morphologically selected sperm injection (IMSI)	Using a high-power microscope to look at the sperm to try to help with selection of the best sperm prior to ICSI.
SpermSlow	A solution containing hyaluronic acid to try and help select the best sperm prior to ICSI.
Incubator	
Time-lapse imaging (Embryoscope, Primovision, CAREmaps)	A process that enables many images of the developing embryos to be taken without removing them from the incubator. It also has the ability to help the embryologist decide which is the best embryo to replace
Medications, including tablets and drips	
Intravenous immunoglobulin (IVIG)	A blood product containing antibodies given through a drip to try to help the immune system not to reject an embryo.
Tumour necrosis factor alpha blocking agents	Medicine given either as an injection under the skin or into a vein to try help the immune system not to reject an embryo.
Intralipid infusion	Medicine given through a drip to reduce the activity of NK cells in the immune system to try to improve IVF outcomes.
Quad therapy: aspirin, heparin, progesterone and prednisolone	A combination of medicines to try to help implantation and the early growth of an embryo.
Platelet rich plasma	A blood product infused either into the uterus or injected into the ovaries to try to improve egg quality or the chance of an embryo implanting into the lining of the womb.

Testosterone or androgens (DHEA, androderm patch)	A hormone given to try to improve the number and quality of eggs and embryos.
Procedures	
Endometrial scratching	A procedure carried out before IVF where the lining of the womb is deliberately scratched to try and make the womb lining more receptive to the embryo implanting.
Endometrial receptivity array (ERA)	A genetic test undertaken from a sample of the lining of the womb to try and help with timing of embryo transfer.
Alternative therapies	
Chinese medicine	The use of herbal medicines to try and improve fertility treatment outcomes.
Acupuncture	Inserting small needles into the skin at specific places on the body to try to improve fertility outcomes.

Paper 5: Patient and professional perspectives about using in vitro fertilisation add-ons in the UK and Australia: a qualitative study

As Principal Investigator, I drove the VALUE study from start to finish, which included conducting the study throughout an unanticipated pandemic. In 2019 when I conceived the idea for VALUE, I had tentatively planned to undertake interviews in the UK face to face and had drafted a proposed budget which reflected the travel costs. I had sought funding for this through the Wellbeing of Women Award, which was unsuccessful. I then applied for funding locally in Bristol through 'Above and Beyond', a hospital charity, which was rejected owing to the high costs related to travel to interview participants from across all four nations of the UK. Then the Covid 19 pandemic hit, and it became clear that interviews could be conducted easily remotely via encrypted and secure video-conferencing software. This meant the budget for VALUE was suddenly much smaller and more manageable. Funding for the study was achieved through the University of Melbourne which allowed for the development of a study website (www.valuestudy.org), thank you vouchers for participants, transcription of audio files, and the use of the qualitative online tool 'Dedoose^{TM}' to facilitate coding of transcripts.

Recruitment and interviewing of participants in the UK and Australia occurred between January and May 2021, with subsequent coding of transcripts by myself and three other researchers. Two coding trees were developed: one for professionals and one for patients. The thematic analysis of these two coding trees took place between a core group of four co-authors which included me as the main coordinator and chair. The wider research team commented upon and debated the themes and subthemes, which are presented here.

The findings of VALUE have been presented at several international conferences including ESHRE in 2022, The Reproductive Medicine Winter Symposium in 2022 where it was awarded the Lawrence Shaw Medal, and the Oxford Scientific Forum in Obstetrics and Gynaecology in 2022 where it attracted first prize. This paper is currently undergoing peer review with BMJ Open.

Patient and professional perspectives about using in vitro fertilisation add-ons in the UK and Australia: a qualitative study

Dr Sarah Armstrong http://orcid.org/0000-0001-7308-520X

Dr Emily Vaughan http://orcid.org/0000-0001-8843-2497

Full address: NIHR Academic Clinical Fellow, University of Bristol, UK. Academic Women's Health Unit, North Bristol NHS Trust, Bristol, UK.

Dr Sarah Lensen http://orcid.org/0000-0002-1694-1142

Full address: Level 7, Royal Women's Hospital, University of Melbourne, Parkville, Victoria, Australia, 3052

Ms Lucy Caughey http://orchid.org/0000-0002-4461-7374

Full address: Department of Obstetrics and Gynaecology, Level 7, Royal Women's Hospital, University of Melbourne, Parkville, Victoria, Australia, 3052

Prof Cynthia Farquhar http://orcid.org/0000-0002-3685-3553

Full address: M&HS Building 507, 28 Park Ave, Grafton, Auckland, 1023, New Zealand

Prof Allan Pacey http://orcid.org/0000-0002-4387-8871

Full address: Department of Oncology and Metabolism, The University of Sheffield, The Medical School, Beech Hill Road, Sheffield, S10 2SF

Prof Adam Balen http://orcid.org/0000-0001-5938-8319

Full address: Leeds Teaching Hospitals NHS Trust, Great George St, Leeds, UK

Dr Michelle Peate http://orcid.org/0000-0003-2903-4688

Full address: Department of Obstetrics and Gynaecology, Level 7, Royal Women's Hospital, University of Melbourne, Parkville, Victoria, Australia, 3052

Dr Elaine Wainwright http://orcid.org/0000-0001-5678-4412

Full address: Epidemiology Group, School of Medicine, Medical Sciences and Nutrition, University of Aberdeen, UK (0.8FTE); Department for Health, University of Bath, UK (0.2 FTE); Honorary Research Fellow, Bath Spa University.

Abstract

Study Question:
What are the drivers for IVF add-on use amongst patients, clinicians and embryologists in the UK and Australia and how is their safety and effectiveness weighed up during the decision-making process?

Summary answer:
Our findings show that the desperation patients experience during IVF gives rise to reaching for add-ons as a source of hope which is prioritised over considerations of safety, efficacy, or cost. For professionals, VALUE shows the tension that add-ons create in the context of traditional evidence-based medicine in the unique sphere of IVF.

What is known already:
There has been very limited qualitative research exploring the attitudes and beliefs of patients and professionals about the use of IVF add-ons. One semi-structured interview study explored the experience of patients and professionals focusing on the use of time-lapse imaging but was limited to patients who had experienced success and professionals working in the public sector.

Study design, size, duration:
'VALUE' is a qualitative semi-structured interview study of both patients (n=25) and health professionals (embryologists (n=25) and clinicians (n=24)) in the UK and Australia. Interviews were conducted between January and May 2021.

Participants/materials, setting, methods:
Participants were recruited in a variety of ways including broad-ranging social media advertising, invitation via professional associations, via an established research panel, and snowball recruitment. The sampling framework included men and women having state subsidised and privately funded cycles, professionals working in the public and private sector, geographical location, and professionals of all grades. Two separate inductive thematic

analyses of anonymised transcriptions were performed; one for professionals and one for patients.

Main results:

Patients often made decisions about add-ons based on hope, minimising considerations of safety, efficacy, or cost, whereas professionals sought the best outcomes for their patients and wanted to avoid them wasting their money. The driving forces behind add-on use differed: for patients, a professional opinion was the most influential reason, whereas for professionals it was seen as patient driven. For both groups, applying the available evidence to individual circumstances was very challenging, especially in the sphere of IVF medicine, where the stakes were viewed as very high.

Limitations, reasons for caution:

Some study authors have previously written quantitative and opinion pieces about add-ons. These potential biases have been acknowledged and managed by including authors who have no affiliation with the add-ons debate. Other limitations include the generalisability of VALUE's findings to other countries. Despite differences in reproductive care in the UK and Australia, we have shown that the participant experience is similar.

Wider implications of the findings:

The VALUE study provides new insights and understanding of how patients and professionals make decisions about IVF add-ons. It is hoped that the findings will expand the existing and controversial debate surrounding add-ons which has often portrayed professionals as mercenary and patients as naïve. Given that the complex landscape of add-ons is likely to grow, there is scope to build on the quality of the discourse between patients and professionals to ensure informed consent it met.

Study funding/competing interests:

Department of Obstetrics and Gynaecology Innovation Grant (University of Melbourne) and an NHMRC Investigator Grant (APP1195189).

Study registration number: osf.io/vnyb9

Introduction

The advent of in vitro fertilisation (IVF) in 1978 was a breakthrough for infertile people, but current live birth rates per cycle initiated are more or less static. In 2019 fewer than 26% of women had a baby with each cycle [1, 2]. The decline in fecundability with advancing female age makes IVF a time-sensitive treatment and this in combination with limited state funding means that many patients often pay for IVF themselves. The pressure to improve IVF outcomes has led to a search for additional or adjunct procedures known as "add-ons". Add-ons have been widely introduced without evaluation and are usually an additional cost to patients [3]. IVF clinics who offer them have been described as 'mercenary' or 'exploitative' [4-10]. The lack of evidence and concerns about informed consent has further highlighted the debate about their merit [11, 12].

For health professionals, a frequent rationale for offering add-ons is a simple response to market forces because patients demand them [27, 110, 111]. Almost three quarters of those undergoing IVF choose to use at least one add-on [27, 28]. However, there is limited research exploring the attitudes and beliefs of patients and professionals about their use. Qualitative research is limited and has not comprehensively included women who have had failed treatment, the views of fertility clinic staff, or considered a range of add-ons [27, 112-114].

We developed a qualitative semi-structured interview study (The VALUE study) for both patients and health professionals (embryologists and clinicians) in the UK and Australia. The aims were to: understand the decision-making process regarding using or recommending add-ons; report sources of information for add-ons; and explore concerns for safety and effectiveness when considering their use.

Materials and methods

The protocol for the VALUE study has been published previously [109], but an amendment was made to exclude those patients in active treatment as it was recognised that the interview may represent an additional psychological burden. The interview schedules can be viewed in supplementary material 1. In both countries, participants were recruited via broad-ranging social media advertising. In addition, the British Fertility Society, and the Association of Reproductive and Clinical Scientists sent emails to its members highlighting the study, and the

charity Fertility Network UK advertised VALUE on its website. In Australia, some participants were recruited via an established Research Panel at the University of Melbourne. Snowball recruitment was utilised across both countries. A purposive sampling strategy was undertaken, whereby participants were selected for in-depth interview when they were deemed likely to be information rich due to their experiences [115]. The sampling framework included people having state subsidised and privately funded cycles, professionals working in the public and private sector, geographical location, and professionals of all grades.

Interviews were conducted by SA and EV (UK) and SL and LC (Australia) between January and May 2021. Semi-structured interviews took place remotely using recorded video-conferencing software, were transcribed verbatim, and then anonymised (Tables I, II, and supplementary Tables I and II).

Concurrent iterative analyses of all transcripts were undertaken using Dedoose™ to organise coding (SocioCultural Research Consultants, Manhattan Beach, California). Embryologist and clinician transcripts were coded together, with the patients occupying a different coding tree because it was reasoned that the responses from professionals were likely to be different. Recruitment of participants ceased once data saturation was achieved through thematic and code saturation which was continually discussed and debated iteratively.

Two separate thematic analyses took place [116, 117]. An inductive approach was adopted, whereby themes were generated from the data, opposed to being mapped to a pre-conceived coding scheme. Before commencement of coding, the coders (SA, EV, DW, and SL), immersed themselves in the data through repeated readings of transcripts and initial thoughts were noted. Coders embarked on the analytic phase together to combine codes into broader themes, and sub-themes, which were discussed, debated, and named. The wider research team then commented upon and debated the themes and sub-themes, which were settled upon by consensus, with only minor changes in the naming of one theme.

Strategies were employed to ensure transparency, credibility, and quality of the research process, especially considering our own professional and research backgrounds [118]. The imperative to be reflexive and open about our own and others' perceptions throughout each

stage of VALUE was acknowledged given some members of the research team (SA, CF, SL, AP, and AB) had published papers about add-ons, while others (EW, EV, MP, LC) had had no previous research with them. We also acknowledged the research team's experience with fertility services, both on a personal and professional level. The most senior qualitative researcher (EW) double-coded 12 transcripts and high agreement between coders was reached, supporting the validity of the results. Regular meetings took place to appraise the sample size, data collection, analyses, and research reflexivity.

The University of Melbourne provided funding. There were no restrictions or requirements affiliated with the funding. All participants were offered a £30 love2shop voucher or a $50 supermarket voucher as a thank you for their time, and a post-interview leaflet was emailed to patient participants which signposted to sources of support. Ethical approval was obtained from the Universities of Sheffield (036268), Bath Spa (BSU-20-205) and Melbourne (2057434.1).

Results

A total of 25 patients were interviewed (11 UK and 14 Australia), 25 embryologists (13 UK and 12 Australia), and 24 clinicians (11 UK and 13 Australia) (Table I, Table II). Interviews lasted an average of 69 minutes (patients) and 45 minutes (professionals). There was a demographic spread of patient (Supplementary Table I) and professional (Supplementary Table II) participants from across the UK and Australia. However, no participants were recruited from Northern Ireland, and within Australia participants were recruited from five of the eight states and territories (Supplementary Tables I and II). Analyses identified five key themes for patients and professionals (Table III) which are compared and contrasted below.

Table I Patients

	Interviews UK (n=11)	Interviews Australia (n=14)	Totals (n/%)
Gender			
Male	0	2	8%
Female	11	12	92%
Age (years)			
25-30 years	1	1	8%
31-35 years	4	4	32%
36-40 years	5	6	44%
41-45 years	1	3	16%
Relationship status			
Single	1	0	4%
In partnership	10	14	96%
Gender of partner			
Female	1	3*	16%
Male	10	11	84%
Treatment undertaken			
IVF/ICSI One – two cycles	4	8	48%
IVF/ICSI Three – four cycles	6	3	36%
IVF/ICSI ≥ five cycles	1	3	16%
IUI (any number of cycles)	2	2	16%
Number of embryo transfer procedures			
None	1	2	12%
One – two	3	5	32%
Three - four	4	4	32%
≥ five	3	3	24%
Cumulative period undergoing fertility treatment (IVF/ICSI, IUI, OI)			
1-2 years	3	6	36%
3-4 years	5	5	40%
5 years or longer	3	3	24%

Table II Professionals

	Interviews UK (n=24)	Interviews Australia (n=25)	Totals (n/%)
Profession			
Clinician	11	13	49%
Embryologist	13	12	51%
Years working in assisted reproduction			
1-5 years	3	1	8%
6-10 years	5	5	21%
11-15 years	6	6	24%
≥ 15 years	10	13	47%
Seniority embryologist (n=25)			
Scientific director	6	3	36%
Laboratory manager	1	0	4%
Senior qualified embryologist	4	5	36%
Qualified embryologist	2	3	20%
Missing	0	1	4%
Seniority clinician (n=24)			
Consultant (O&G) fertility specialist	7	12	79%
Staff grade fertility specialist	1	0	4%
Clinical Fellow/trainee reproductive medicine	3	1	17%

Desperation and the compulsion to treat it

Patients were vulnerable and had a strong sense of desperation. Desperation was illustrated with several examples of 'bargaining', with patients willing to suffer theoretical hardships offered by professionals if it meant a successful outcome: *"If they'd said, I don't know, stand on your head for an hour and that would work, I would have done that. You know, it, it just leaves patients very vulnerable, I think"* (UK patient 4). Decision making in the context of desperation gave rise to examples of being willing to try any add-on, no matter how small the additional chance of success. One participant compared her situation to a patient she cared for in her role at work: *"We had a little boy, he went off to China for some weird therapy because of a 1% chance it might work, and I could never understand that. But I kind of do now, because you get to a point, you're so desperate why wouldn't you? If you have that 1% chance*

of it working, you'd throw that 1% at it" (UK patient 11). The goal of parenthood for IVF patients was profoundly important, and fertility treatment often left them feeling out of control, powerless, and at the mercy of chance. Opting to use add-ons was a way of bringing about purpose and control (Table IV, UK patient 7).

In addition, add-ons provided renewed hope by offering a bespoke addition to the cycle that may bring about a successful IVF outcome. Safety and efficacy were minimized in favour of hope: *"I couldn't care less, you could've told me the risk was really high, and honestly, I just couldn't have cared. Because if you could guarantee me a baby, if it was a 100% guarantee, I just had to chop off my left arm, I would have been, no worries, just chop it off. If it was 100% guarantee, so, the risk, even if there was risks, I couldn't have cared less, what any of, probably, the side-effects or risks were"* (Australia patient 3). Nearly all patients indicated that it was unacceptable for a clinician to offer an add-on based on false hope, with many citing they held doctors to a higher standard of honesty (Table IV, Australia patient 11).

For professionals, some expressed that hope from utilizing an add-on was beneficial to patients, whereas other felt that hope was false, burdensome, and left patients vulnerable to exploitation: *"Disadvantages are that it [the add-on] might give them false hope, and I think it's really important not to do that. We've got to be honest. Let's not take hope away, but we've got to be honest with people"* (UK clinician 5). Professionals acknowledged the desperation their patients experience, particularly after unsuccessful cycles. Add-ons offered the patient a change to the subsequent cycle, and in the absence of any other evidence-based interventions, were a reasonable option: *"I think for the rest of the add-ons, it's really when the consultation turns into a sort of, consultation of desperation. Like the patients had several failed cycles, and her NHS funding is just about to finish or perhaps she has just one cycle left with the NHS. I think that's the point when, if we can improve something, if we can change something in a treatment protocol without essentially incurring extra cost then it's, sort of a, why not? Sort of consultation"* (UK clinician 8). Add-ons also offer their patients the opportunity to feel that they and the clinic have tried their best, even if the cycle ends in disappointment. It was also believed to absolve patients of regret at not having 'tried everything' (Table V, Australia embryologist 12).

	Themes	Sub-themes
Patient thematic analysis	Patient 1: Vulnerability	1.1 Desperate for success 1.2 Add-ons lend hope and a sense of control 1.3 Safety and efficacy ranked lower than hope
	Patient 2: Power of the trusted professional opinion	2.1 Must be in my best interest 2.2 Unaware add-on was optional 2.3 Supports patient autonomy 2.4 Informed consent important
	Patient 3: The evidence doesn't apply to me	3.1 Tension between EBM and bespoke care 3.2 Power of personal testimonies
	Patient 4: Acceptability of add-on	4.1 Risks perceived as low 4.2 Costs worth it: stakes are high
	Patient 5: Role of previous experience	5.1 Previously used and had success 5.2 Previous unsuccessful cycle
Professional thematic analysis	Professional 1: Treating desperation	1.1 In absence of anything else, it is reasonable to offer add-ons 1.2 Allows patients to exhaust every avenue 1.3 Hope versus false hope
	Professional 2: The patient shopper	2.1 Patients drive use following personal research 2.2 Allows patients autonomy to choose provided informed consent undertaken 2.3 Not being cutting-edge risks losing patients
	Professional 3: Tensions within evidence-based practice	3.1 'One size fits all' versus 'bespoke tailoring' 3.2 Continuum of approaches to the evidence 3.3 Being evidence-based in IVF is challenging
	Professional 4: Potential for harm	4.1 Add-on potentially harmful 4.2 Discomfort with performing some lab-based add-ons
	Professional 5: Success, not profits	5.1 Genuine desire to help, and avoid wasting patients' money 5.2 Other clinics exploit patients 5.3 Discomfort in charging for add-ons 5.4 Add-ons increase workload for clinic

Table III: Themes and subthemes

Professionals versus patients: who is driving add-on use?

The driving forces behind the use of add-ons differed between patients and professionals. For patients, a professional opinion was felt to be the most influential reason for opting to use them. Such recommendations held a lot of sway and were hard to disregard. Patients described how the add-on was in their best interest and a bespoke element of care: *"That all came from the clinic. I hadn't heard of either of them before, natural killer cells or the ERA [endometrial receptivity array] testing. And it was more a, we're going to do this. We're testing this and then when the results came, they were like, we're going to do this. We're changing this, do these medications. Obviously, we had the choice, but for us it was a no-brainer. If that's what your specialist is telling you to do, then we're doing it"* (Australia patient 1).

The power of the professional opinion wasn't limited to clinicians. Some participants described how important the opinions of their nurse or embryologist were: "*And I remember the nurse had said to us, you know, if it failed, would you consider that you've done everything that you've possibly could? And then we were like, all right, yeah, no, we should go with the options that you've given us…"* (UK patient 7). The importance of the professional opinion also holds when the recommendation is to reject an add-on, providing patients with the freedom to stop considering it (Table IV, Australia patient 3).

The power of the professional opinion sometimes extended to patients not realising the optional nature of add-ons, on the basis that if it had been offered, then it must be an essential element of care. Learning about the additional cost was sometimes only revealed when they came to pay (Table IV, UK patient 2). Whilst the professional opinion was important, patients also expressed the desire for autonomy with add-on choices: *"For my fifth transfer I wanted to try something. I'd had four failed and I think I was quite happy to try it, so they agreed to do the scratch"* (Australia patient 11). The need for adequate counselling about the risks to make an informed decision was deemed important by over two thirds of participants. (Table IV, UK patient 5).

Contrasting 'power of the trusted professional opinion' is the professional theme 'the patient shopper'. Professionals described the well-informed patient, who had undertaken extensive

reading online, and had clear preferences regarding add-ons: *"You know, it, it used to be the case that they [patients] would leave their brain at the door and just walk in and do as they're told. And now I think, I absolutely don't think that's the case with a large proportion of patients. I think they come in through the door knowing what they want and often having researched it"* (UK embryologist 2).

Professionals described the importance of listening to patient requests about add-ons but caveated this with the need to maintain the core ethical principle of informed consent: *"When they come and talk to me about growth factor, I show them that paper, and say, look, it really has not shown any benefit, it costs as much as another IVF cycle. You know, if you wanted to use it, that's fine, but there's been a proven study, that hasn't shown a benefit from it, it's enormously expensive, and, and you'll, you'll get much more chance if you do another IVF cycle"* (Australia clinician 9). Being able to offer add-ons provides professionals with the benefit of appearing modern and innovative. The patient shopper drives add-on use which clinics respond to in the hope of attracting new clients and keeping those who might go elsewhere, possibly for less ethical treatment (Table V, UK embryologist 9).

Add-ons and evidence

For both patients and professionals, applying the available evidence to individual circumstances was very challenging. Patients overwhelmingly appreciated the importance of evidence to inform healthcare decisions, but when it came to decisions about themselves, tailored care took precedence. Their clinician's opinion and personal experience trumped evidence from randomised controlled trials (RCTs), which were felt to not represent their unique clinical circumstances (Table IV, Australia patient 5). There also was a tension between evidence-based medicine and personal testimony read online. They described how the blog of an unknown IVF patient, especially one with a similar set of circumstances to them, was very compelling: *"…that's why the Janet from Birmingham comes in useful. Because she will say, I've done five cycles with rubbish eggs, rubbish embryos, everything was terrible, but then I did that [add-on], and look what happened. And so, for me, that individual story of, similar to me, for instance, has done this, tried that, and it's worked, is a lot more helpful, even though it's one person, than knowing what happened to a hundred. When you read that Janet from*

Birmingham did this and got pregnant, you're, like, oh, my god, it's going to work for me. So, I'd say, I can honestly say that is the most powerful, powerful influencer of all" (UK patient 5).

For professionals, there was also a profound tension within evidence-based practice. Professional attitude to evidence about add-ons sat on a spectrum between scepticism and trust and seemed linked to whether the professional subscribed to a 'one size fits all' or a 'bespoke tailoring' approach. Those who subscribed to the former appeared to believe that patients have a largely common set of problems that explain their infertility, and when cycles failed, this was due to chance, or more specifically, aneuploidy within the embryo. Thus, their preference was to repeat cycles without add-ons in the hope that eventually a euploid embryo would be replaced and result in a pregnancy. They were also concerned that add-ons might exhaust funding better spent on an additional cycle of IVF. (Table V, Australia clinician 6). One-size-fits-all practitioners were critical of add-ons which they felt lacked evidence of efficacy, and described changing their clinical practice in response to RCTs: *"I've prescribed growth hormone, I don't know, not more than a half dozen times in my life but, [fertility clinic] did a study which I think was called the [name of study] and that sort of refuted any perceived benefits so I stopped doing that."* (Australia clinician 13).

Professionals who were 'bespoke tailors' believed that patients (especially those with multiple failed cycles) have specific problems, that may be identified through extensive diagnostic testing, and remedied by add-ons. They held the available evidence with scepticism, which was criticised for being underpowered thus ruling out the identification of efficacy in sub-groups, and for 'cherry picking' good prognosis patients (Table V, Australia clinician 8). They described feeling uneasy about the ethics of conducting RCTs on technologies already available and described suspicion of research groups' objectives and publishing journals' political stance on add-ons (Table V, Australia embryologist 8). In the absence of compelling evidence from RCTs, their practice was based on scientific plausibility and on evidence gleaned from their own clinic's data: *"It's even difficult to prove something that's ineffective because the trials that are required are often very expensive, large and might not be applicable to a particular patient. So, if you group all patients together in a trial you may not find evidence of effectiveness, but if you looked at some subgroups perhaps you would. So, a trial is not real life. You know, evidence-based medicine, you have to take the best*

evidence and then apply it to a patient in front of you who may not be the same as patients in the trial" (Australia clinician 8). Many professionals occupied points between 'one size fits all' and 'bespoke tailoring', however even the staunchest 'one size fits all' professionals described being willing to provide add-ons under specific clinical circumstances.

Table IV Patient Quotes

Patient quotes		
Theme	Participant	Illustrative quotations
Patient theme 1: 'Vulnerability'	UK patient 7	"Because if someone's offering me something and they, they say, well, it could work but it might not work, you cling on to the it could work. So, yeah, I'd probably still go ahead with it. But, you know, looking back now, uhm, it [evidence of effectiveness] does matter. But, you know, it doesn't make a difference when you're, kind of, in the flows of it and there were lots of emotions flying around"
	Australia patient 11	"If a doctor wants to offer some hope because they genuinely think that might work for you, wonderful, but I don't think you can give people false hope because it will make the doctor feel better. You know, if a doctor, there's no point doing it just to make the doctor feel better about giving a patient hope. They've got to truly believe it would actually work"
Patient theme 2: 'Power of the trusted professional opinion'	Australia patient 3	"And he said that's got nothing to do with it, you're just throwing money down the drain. You might as well just stop [DHEA], and it was really blunt, and I did actually just stop. And then he said, [clinician quotes study regarding melatonin]. And he said, you're wasting your money on that too. And I said, okay terrific, so I stopped both of them [DHEA and melatonin], which was fine, I suppose it saved me some money"
	UK patient 2	"I thought that was just standard, to be honest. I didn't realise that [time-lapse imaging] was an option. I mean, it came up, sort of, itemised on our bill so maybe I should've guessed from that that I could've taken it off."
	UK patient 5	"...I think it's a bewildering, overwhelming world of stuff that lay people wouldn't necessarily understand. And yes, it doesn't seem fair that they include them [add-ons] as standard when people can't make the active choice, based on research, whether to go ahead or not"
Patient theme 3: 'The evidence	Australia patient 5	Australia patient 5: "So I think a lot of these, even the ERA test, the endometrial scratching, a lot of them are actually not proven to guarantee success. It's just, I think because

doesn't apply to me'		everybody's different, everybody responds differently to treatment, I don't think there's ever going to be a definite answer, scientifically proven answer, for every single person."
Patient theme 4: 'Acceptability of add-on'	Australia patient 6	"I also looked at the dangers of PGT testing. So, let's say wrong results come back. And it was only quite low, so I was okay with that. I also looked at the risk of it being, like, what if it will harm the embryo unnecessarily and actually make the embryo unusable? But I think the risk of that is also quite low, so I was comfortable with that."
	UK patient 6	"Cost was a massive thing for us. We ended up re-mortgaging our house to pay for our treatment….."
Patient theme 5: 'Role of previous experience'	Australia patient 13	"But that said, one thing that did work really well for us, which was not noted down, was, a song called [name]. We played that before we went to the clinic, and that worked for our first daughter. And then the second time around… Of course, the same cocktail of different, different combinations. We also played that song again on the way on the way to the insemination clinic, and it worked two times [laughter]. For our friends, we said, we know you don't like this music at all, but put this song on on the way to the IVF clinic, and see if it works, and it did. Three for three, scientifically proven [laughing]. You should play this. I'm just throwing it out there, so there's… That's three for three"

Consideration of risks

Add-ons were acceptable to patients, particularly in the context of a professional recommendation, because they were perceived as low risk and worth the cost (Table IV, Australia patient 6). Although the additional cost was a burden, the goal of parenthood was more important and therefore worth the financial strain. For some, using add-ons left them in significant debt, with participants remortgaging their home or borrowing money to fund them (Table IV, UK patient 6). The substantial cost of IVF was used as an 'anchor' to reference and justify the relatively low cost of add-ons: *"….it just got to the point because they were all, it's not thousands of pounds each one, is it? It's like, the scratch is, I don't know, a couple of hundred pounds. None of it was so expensive that it was, that it made you think. It seemed like a drop in the ocean I guess to the thousands of pounds that we'd already paid"* (UK patient 9).

In contrast to the patients, professionals held concerns about the potentially harmful nature of certain add-ons, particularly assisted hatching, pre-implantation genetic testing (PGT-A), and immunological therapies (Table V, UK clinician 1). For one embryologist, the requirement to undertake PGT-A was the catalyst to change employer: *"And I felt very uncomfortable. The way we presented it was, you know, those are things that might help, but we're not sure that they will. But what we do know, or what we used to say is that we did know that it wouldn't do any harm. And I now feel uncomfortable about that as well, particularly about PGT-A, because you're really putting embryos in a very sort of stressful situation, with no evidence that what you're doing will make a difference to the outcome to the patient. And you're mutilating the embryo. And also, you're taking cells that might not be representative of what the fetus's cells will be like. So, this is one of the ones that I felt most uncomfortable with, and part of the reason why I left where I was working"* (UK embryologist 1).

Role of previous experience

Use of an add-on in a previous successful cycle was an important driver for patients. Deviating from a 'tried and tested' formula was difficult as it was impossible to tease out whether it was the add-on that had led to success: *"I'm glad we used the scratch, very glad, because whilst I can't say it was what caused us to conceive, I can't say it didn't. If I was in the position where I needed to do IVF again, I would definitely pay for it every time"* (UK patient 6). For some this becomes a superstition, even when it can be rationalised that the add-on is unlikely to be helpful (Table IV, Australia patient 13).

Table V Professional Quotes

Professional quotes		
Theme	Participant	Illustrative quotations
Professional theme 1: Treating desperation	Australia embryologist 12	"I think having add-ons gives them that slight feeling that they have opened all the doors. They have explored all the avenues. And then maybe they will be able to, you know, complete their IVF journey at least with the satisfaction that they know they have tried everything. They have given it all"
Professional theme 2: 'The patient shopper'	UK embryologist 9	"But I think patients as well are becoming a lot more informed and a lot more are aware of what is available. And certainly I, I think a proportion of patients, you know, if they are not able to have certain add-ons at a particular

		clinic they can probably take their business elsewhere as well."
Professional theme 3: 'Tensions within evidence-based practice'	Australia clinician 6	"Everyone talks about how expensive IVF is and, you know, that it's $10,000 a cycle. And many people say it has a low success rate, which is not true. So, if a patient has a certain amount of money, there is no question that the greatest likelihood of getting pregnant is the more IVF they have, the more cycles they have, the more eggs that are collected, the more embryos that are made. So, when patients are wasting their money on totally charlatan, unethical treatments, they are using their pool of money towards something that is not making them pregnant. For example, something like embryo biopsy that might double the cost of the cycle. Where, in fact, they would have been better off having two cycles"
	Australia clinician 8	"To have a proper randomised trial that can prove efficacy, let's say improving the chances from 2% to 3%, that's a 50% increase, but to actually have that show significance in a randomised control trial would be impossible to do. It's not a trial that you can do because it requires thousands and thousands of patients. Nobody can run such a trial. So, uh, our ability to prove efficacy of any add-on is very limited."
	Australia embryologist 8	"And all the scientific journal papers are skewed, as well, depending on the clinic, depending on who's studying it, and depending on who's, who's publishing it. I think the results vary too much at this stage with a lot of different add-ons."
Professional theme 4: 'Potential for harm'	UK clinician 1	UK clinician 1: "…there will be the case where you will damage some embryos in the process of biopsying them for example. And you will, you'll have some abnormal, you'll, you may be damaging the occasional normal embryo. And I always say to patients about mosaicism, just because the results are abnormal doesn't mean to say that that baby's abnormal. So, you may be causing harm, but that's a discussion we have with them."
Professional theme 5: 'Success not profits'	Australia clinician 10	"I can't begin to tell you the anger that I see in my rooms when for whatever reason I say to them, well, natural killer cells are elevated, or you have tissue compatibly, or you have a balance translocation of your chromosomes which is why you're not getting pregnant. But nobody has done these tests and they've had IVF treatments without success. And they're very angry that they've wasted all this money on previous cycles"
	Australia embryologist 9	"I actually feel quite angry sometimes when you hear of patients that have gone to other clinics and been sold all this stuff that's really, you do wonder if it's doing more

		harm than good. And you just think if, your problem is, is quite simple, cut away all of that and just focus on the basic science that we know is working, save your money. I think, I, I do think that there is a bit of exploitation going on."
	UK embryologist 10	"But there's other add-ons that, they don't cost any money to the clinic, like assisted hatching. That, the cost for the lab is zero. Obviously, you can always factor in the knowledge of the embryologist, the equipment calibrations, blah, blah, blah. But the cost is essentially to buy anything. So, for that one, for example, there shouldn't be a charge at all"

Success, not profits

Analyses showed professionals wanted the best outcomes for their patients alongside a genuine desire to avoid wasting their money. For those who subscribed to the 'one size fits all' approach, add-ons were avoided when they were not clinically relevant. Conversely, for 'bespoke tailors', the lack of add-ons in previous unsuccessful cycles was deemed as failing to optimise all variables (Table V, Australia clinician 10).

Holding their patients' best interests at heart was expressed universally, however they observed that some other clinics used add-ons unethically for financial gain, including clinics they had previously worked for (Table V, Australia embryologist 9).

Professionals expressed discomfort at charging patients for add-ons in various contexts including when they were used routinely (e.g., time-lapse incubation of embryos), or when the cost of the add-on technology had already been met by the clinic (Table V, UK embryologist 10). Professionals described the paradox of charging for add-ons that are believed to be effective whilst also charging for add-ons that were not. Many argued that add-ons were only ever offered in a 'success not profit' context by expressing how they increase the burden on clinics. Keeping track of individual add-ons increased complexity in the laboratory and heightened workload around managing patients' expectations: *"It is a bit time consuming, to be honest, to go through the list of adds-on with patients, in particular, the ones that are very well, you know, brainwashed by Google, and they know everything, and*

they just start from the beginning. And so it is, it does add time to the consultation" (UK clinician 6).

Discussion

The VALUE study provides new insights and understanding of how patients and professionals make decisions about IVF add-ons. Patients describe the importance of hope, which is ranked higher than considerations of efficacy, safety, and cost to frame their choices, particularly after previously unsuccessful cycles. Choosing an add-on offers a sense of control, with the possibility of overcoming problems encountered previously. The driver for add-ons from a professional perspective is ascribed to patients, however, patients describe the power of the professional opinion, but also acknowledge that seeking add-ons is a quest for hope sought after learning of success stories online.

VALUE's findings are at odds with the debate surrounding add-ons, which often portrays professionals within fertility services as having commercial incentives [4-6]. We found no 'smoking gun' to suggest that professionals saw add-ons as a means of generating revenue, however there was an acceptance that unethical practices do exist, with examples of embryologists being uncomfortable performing some laboratory-based add-ons, believing them to be potentially harmful to embryos. We show here the significant tension that exists between traditional evidence-based medicine and IVF in the era of add-ons. Clinicians and embryologists sit on a spectrum regarding their approach: 'one size fits all' versus 'bespoke tailors'. However, even for the staunchest evidence-based medicine advocates, there were caveats where add-ons would be offered. For patients, the stakes are so high, that evidence concerning efficacy and safety although important, are not the most important factor when deciding upon an add-on, alongside a belief that the evidence doesn't reflect their unique clinical circumstances.

VALUE explores the patient-professional decision-making dyad regarding add-ons for the first time. When compared to other high-stakes, time-critical clinical situations, such as cancer treatment decisions, we find that there is a distinct difference. For routine cancer care, shared decision making is rarely implemented [119]. In contrast, we found that people undergoing IVF

are often actively engaged with clinical decisions, consuming information online, and are supported by professionals to exercise autonomy regarding add-ons.

There is limited qualitative evidence exploring the patient and professional perspective surrounding add-ons. One semi-structured interview study analysed how professionals legitimise the use of one add on: time-lapse imaging [112]. The authors suggest that professionals create legitimation arguments for its use, downplaying the values of traditional evidence-based medicine to evaluate its worth [112]. VALUE goes further to show that professionals occupy a spectrum of approaches to the evidence, which furthermore guides their clinical practice. Previous studies have found hope to be of critical appeal for patients, and important for persevering against adversity [113, 120]. VALUE offers a broader explanation as to the appeal of add-ons, including the role of desperation, and hierarchy of hope over other priorities. Regarding patients, this is the first study that offers the perspective of those who have had and not had, success from IVF, including those who considered, but didn't use add-ons.

VALUE's strengths lie in its robust design and development which included the opinions of patients and professionals. We interviewed a broad range of participants spread across two countries. Some study authors have previously written quantitative and opinion pieces about add-ons. These potential biases have been acknowledged and managed by including authors who have no affiliation with the add-ons debate and by involving qualitative research experts (EW and MP) at every stage. Other limitations include the generalisability of our findings to other countries. Despite differences in reproductive care in the UK and Australia, VALUE has shown that the participant experience is similar. This may reflect the comparable availability of add-ons in both settings.

VALUE shows that the IVF add-on debate is more nuanced than merely predatory clinics and naïve patients. Add-ons strike to the heart of traditional evidence-based medicine. With their presence ever expanding in the largely privatised sphere of IVF, policy makers and regulators will be tasked with establishing how they will reconcile add-ons in an evidence-based medicine world.

Acknowledgements:

Our thanks to all the patients, embryologists, and doctors, who participated in VALUE and our PPI panel who gave up their time to contribute to the development of VALUE. PPI panel UK: Isabella Dash, Jennifer Nisbett, Hannah Reed, Ally Richardson, Victoria Thomas, Bassel Wattar. Australia: Katherine Gobbi, Hilary Smith, Natasha Devetak, Alex Polyakov, Anna Ninnis, Lisa Lee, Vadim Mirmilstein. Thank you to Dr David Wainwright, University of Bath, for undertaking the coding of some of the professional UK transcripts.

Contributions of authors:

All authors except LC were involved in the design of VALUE and publication of the protocol. SA, EV, SL and LC undertook participant interviews. Analysis was undertaken by SA, EV, SL and EW. SA wrote the initial draft which was edited by all authors. All authors involved in analysis (SA, SL, EV, EW) had full access to the anonymised data set and accept responsibility for publication.

Funding

The University of Melbourne provided funding. There were no restrictions or requirements affiliated with the funding.

Conflicts of interest:

SA is a subspecialty trainee in reproductive medicine and surgery and works within a fertility clinic setting. She is an editor for Cochrane Gynaecology and Fertility, and Human Fertility. In the past 12 months she been an invited speaker for ESHRE, receiving expenses for travel. She has no financial relationships or affiliations with any commercial companies.

EV is an O&G registrar in Severn Deanery, she is an NIHR Academic Clinical Fellow with the University of Bristol. She has no financial relationships or affiliations with any commercial companies.

SL is a clinical trialist and methodologist, employed by the University of Melbourne. She has a strong interested in randomised controlled trials and systematic reviews and is an editor for the Cochrane Gynaecology and Fertility group. She has no financial relationships or affiliations with any commercial companies and has received research funding from only Government and not-for-profit organisations.

LC has no conflicts of interest to declare.

CF is a Professor of Obstetrics and Gynaecology, University of Auckland, NZ. She is a fertility specialist, Clinical Director of Gynaecology and Fertility Plus, Auckland District Health Board, NZ and Co-ordinating Editor of the Cochrane Gynaecology and Fertility Group. She has no financial relationships with any commercial companies or fertility clinics. She is an invited speaker for ESHRE 2022, receiving expenses for travel.

AAP is the Editor in Chief of Human Fertility, Trustee of the Progress Educational Trust, and Chairman of the advisory committee of the UK National External Quality Assurance Schemes in Andrology (all unpaid). In the last 24 months, he has undertaken paid consultancy, Speaker fees or Contributor fees from Cryos International, Cytoswim Ltd, Exseed Health, Merck Serono Limited, but all monies associated with this are paid to The University of Sheffield.

MP is behavioural scientist, employed by the University of Melbourne. She is President-Elect for the Australian Society for Psychosocial Obstetrics and Gynaecology and is on the editorial board for Journal of Psychosocial Oncology Research and Practice. She is an invited speaker for ESHRE 2022, receiving expenses for travel. She has no financial relationships or affiliations with any commercial companies and has received research funding from only Government and not-for-profit organisations.

AHB is an NHS Consultant in Reproductive Medicine, Honorary Chair and Clinical lead for Leeds Centre for Reproductive Medicine, Leeds Teaching Hospitals NHS Trust. He is also Consultant at CARE Fertility, Leeds, and Chair of The CARE Fertility UK Innovation and Research Board. He is Fellows' Representative RCOG Council for Northern, Yorkshire and Humber, a Trustee of The British Fertility Society (BFS), a Member of the World Health Organisation Guideline Development Group on Infertility. He is a director of Balance Reproductive Health Ltd and a member of the advisory board on PCOS and weight reduction, NovoNordisk Pharmaceuticals, for which he has received speaker's fees for attending meetings.

EW is a Chartered psychologist (academic) employed by the Universities of Aberdeen and Bath and with an Honorary Research Fellowship at Bath Spa University. She has a strong interest and expertise in qualitative methods as applied to health services research. She has no financial relationships or affiliations with any commercial companies and has received research funding from only Government and not-for-profit organisations.

Supplementary Table I Patients			
	Interviews UK (n=11)	**Interviews Australia (n=14)**	**Totals (%)**
Ethnicity (UK)			
White British	10	-	91%
Mixed or Multiple ethnic groups	1	-	9%
Ethnicity (Australia)			
Caucasian	-	5	36%
Australian	-	5	36%
Other (free text responses)	-	3	21%
Missing	-	1	7%
Highest educational qualification			
Secondary school	0	1	4%
Diploma or certificate	0	2	8%
Undergraduate university degree	8	5	52%
Postgraduate degree/qualification	3	6	36%
Religion			
Christian	5	6	44%
Other	0	1	4%
No religion	6	6	48%
Missing	0	1	4%
Average household income per year before tax (GBP and AUD)			
£20,000-£40,000/$40,000-$120,000	2	0	8%
£40,001-£60,000/$80,001 -$120,000	1	6	28%
£60,001-£80,000/$120,001-$160,000	1	1	8%
£80,001-£100,000/$160,001-$200,00	4	3	28%
£100,001-£150,000/$200,001-$300,000	1	2	12%
>£150,000	1	0	4%
Prefer not to say/missing	1	2	12%
Country participant lives in (UK)			
Scotland	1	-	9%
England	9	-	82%

Yorkshire and The Humber	2	-	-
South East	4	-	-
North West	2	-	-
East Midlands	1	-	-
Northern Ireland	0	-	0%
Wales	1	-	9%
States participant lives in (Australia)			
New South Wales	-	3	22%
Queensland	-	2	14%
Northern Territory	-	0	0%
Western Australia	-	0	0%
South Australia	-	1	7%
Victoria	-	6	43%
Australian Capital Territory	-	2	14%
Tasmania	-	0	0%
Donor eggs used in treatment	0	0	0%
Donor sperm used in treatment	2	3	20%
Children			
Child/children naturally conceived	2	0	8%
Child/children conceived through IVF, ICSI or IUI	8	8	64%
Child/children conceived through ovulation induction	0	1	4%
Foster or adopted child/children	0	0	0%
Stepchild/children	1	0	4%

Supplementary Table II Professionals			
	Interviews UK (n=24)	**Interviews Australia (n=25)**	**Totals (%)**
Clinical practice			
Only private	5	NA	-
Only public	2	NA	-
Mixture of private and public	17	NA	-
Country participant practices in (UK)			
England	22	-	92%
Wales	1	-	4%
Scotland	1	-	4%
State participant practices in (Australia)			
Australian Capital Territory	-	1	4%
New South Wales	-	4	16%
South Australia	-	2	8%
Western Australia	-	1	4%
Victoria	-	15	60%
Queensland	-	2	8%

Supplementary Material I

Qualitative question schedule patient participants, VALUE study

1. You indicated that you have considered using [add-on] as part of your fertility treatment. Can you talk me through your decision making about whether to use this/these add-on(s)?

Prompts:

a. Who was involved in the decision?

b. What were your thoughts on the potential benefits?

c. What were your thoughts on the potential risks?

d. What about cost? Did that feature?

e. Any other factors you can think of?

f. Did you know that the [add-ons] you used/thought about were optional, or are often optional?

g. Can you tell me how your first came to consider using [add-on]

h. What happened next?

i. Can you tell me a bit more about that?

2. Can you tell me if you received or sought any information about [add-ons], and if so where you got the information from?

Prompts:

a. Your fertility clinic website, information brochures?

b. Your fertility doctor or another staff member (verbally)?

c. Internet searches/blogs?

d. What about family advice, or information from family?

e. How about social media or internet advertising?

f. Was the information helpful/useful/did it help with making the decision?

3. Reflecting on your experience, how do you feel about the decision to use (or not) use [insert name add-on] in your IVF treatment?

Prompts:

a. Did you get pregnant that cycle?

b. Would you recommend it to someone else?

c. What advice would you give someone else considering this add-on?

4. Sometimes add-ons are available at fertility clinics before they have been thoroughly studied to check that they increase the chances of IVF being successful. What are your thoughts on this?

Prompts:

a. Does whether they are proven to work or not make any difference to you?

b. Do you think scientific evidence of benefit matters to patients?

c. Is whether it costs patients money, and the amount, important?

d. What if the add-on is used as routine or standard at specific clinics, for instance included in a treatment package?

5. Sometimes add-ons are available at fertility clinics before they have been thoroughly studied to check they are safe, for example whether they increase the risk of miscarriage, stillbirth, or poor outcomes for the baby. What are your thoughts on fertility clinics offering add-ons in these cases?

Prompts:

a. Does whether they are proven to be safe make any difference to you?

b. Do you think scientific evidence of safety matters to patients?

6. Now I'd like you to pretend for a moment that you are considering using a brand-new add-on in your IVF cycle that has not yet been scientifically proven as effective and safe.

I will show you a slide and would like it if you could let me know which of these you agree with? Perhaps you could talk me through your thoughts?

Slide:

a. OK if it is free to use

b. OK if is low risk

c. OK if the doctor has a hunch that it might be effective

d. OK if the doctor wants to offer some hope

e. OK if the doctor has used it before and a patient got pregnant

f. OK if the patient wants to use it

g. OK if if it's being studied as part of a research project

7. If you were given an information leaflet about an IVF add-on, what sort of information would you want it to include?

Prompts:

a. How about risks and benefits?

b. Would how it works and how long it's been around for matter to you?

c. Would you be interested in success rates? Cost?

8. What do you think about using the term 'add-on'?

Prompts:

a. Do you have any alternate suggestions for terms?

b. Is it a good description?

c. What do you think about the terms 'adjuvant' or 'adjunct' instead?

9. Is there anything else I haven't asked you about add-ons, that you'd like to share or talk about?

Qualitative question schedule professional participants, VALUE study

1. Can you tell me how you would define an add-on? For example, what components would you include in a definition?

Prompts:

a. evidence of effectiveness, evidence of safety, costs, being optional?

b. What things do you think definitely aren't add-ons?

c. What things do you think definitely are add-ons?

2. What do you think about using the term 'add-on'?

Prompts:

a. Do you have any alternate suggestions for terms?

b. Is it a good description?

c. What do you think about the terms 'adjuvant' or 'adjunct' instead?

3. Please can you tell me about your experience of offering or using add-ons recently, say in the last couple of years?

Prompts:

a. Can you give some examples of add-ons you talk about with your patients (or colleagues if embryologist)?

b. Can you recall a particular case in which you did this?

c. Can you tell me a bit more about that?

d. Have you any other thoughts on that topic?

e. If patient drivers: How do you weigh up patient autonomy versus your duty to beneficence and non-maleficence when you know there is no clinical benefit?

f. Who makes the decisions regarding laboratory-based add-ons? How does this make you feel?

4. What advantages or benefits do you think add-ons provide to patients?

Prompts:

a. Can you give some examples of add-ons you think offer advantages to patients?

b. Can you think of some instances where you might you recommend or suggest an add-on to your patient?

c. What advantages or benefits do you think add-ons provide to clinicians or the clinic?

5. What disadvantages or risks do you think add-ons pose to patients?

Prompts:

a. Can you give some examples of add-ons you think are particularly risky?

b. Can you recall any particular instances when adverse events arose as a result of using an add-on?

c. Can you think of some instances where you might advise a patient against using an add-on?

d. Can you think of any disadvantages that add-ons pose to clinicians or the clinic?

6. Where do you seek information on add-ons from?

Prompts:

a. What about where you seek information on a technique or how to use/perform the add-on (such as the dose, timing of procedure etc.)?

b. What about information on evidence such as safety and effectiveness?

7. What kind of information do you or your clinic aim to give patients regarding add-ons?

Prompts:

a. Do you ever direct them to any resources?

b. Do you ever provide written information?

c. Can you tell me a bit more about that?

7. Conclusions and areas for future research

The aims of this MD that I set out in section 6.8 were threefold: (i) to present the highest quality quantitative and qualitative evidence on IVF add-ons; (ii) to advocate for evidence-based medicine regarding add-ons and; (iii) to be a source of trusted information. I believe papers 1 to 5 presented here represent the highest quality, transparent evidence that achieve these aims. However, there is still more that can be done. To fully achieve aims two and three, I need to continue engaging in scientific debates regarding add-ons, and to reach out to patient facing resources to ensure that the findings of all five papers are accessible. Our decision to ensure papers are open access has been by design to ensure that the findings are freely available to all. For aim one to be a lasting legacy of this MD, I am resolved to continue to periodically update the Cochrane reviews on GM-CSF containing culture media, and time-lapse systems. Cochrane systematic reviews are only relevant so long as they are current, and therefore need revisiting every two to three years to establish what further evidence can be included and analysed. This is an ongoing commitment to evidence-based medicine in the sphere of add-ons.

The governance of add-ons in the future

Reproductive medicine, in particular IVF, occupies a unique area of medicine, where the usual paradigm of evidence-based medicine becomes more blurred. This is due to the interplay between the high-stakes and emotive nature of infertility, private medicine, and commercial industry who generate technologies which often side-step the usual Medicines and Healthcare products Regulatory Agency (MHRA) approval [121]. Add-ons flourish in this sphere, and VALUE has shown that there is an appetite for them from both patients and professionals, in a quest to improve outcomes, no matter how small the additional improvement may be [110]. There is no doubt that add-ons will continue to be developed by industry, who see the commercial benefit in improving IVF success rates. As of 2019, the global IVF market was worth US$25bn and is projected to reach a value of US$41bn by 2026 [122]. No wonder investors are willing to spend time and resources on add-ons in such a lucrative market.

In the UK, the HFEA's involvement in add-ons to date has stopped at the development of a webpage which utilises a traffic-light system to denote the quality of the evidence to reflect

the effectiveness of add-ons. The HFEA is unlikely to ever regulate or stimy the availability of add-ons, as the Human Fertilisation and Embryology Act 2008 does not mandate the licensing of IVF technologies, instead this lies with the MHRA[123].

Governance of add-ons will therefore have to fall within the remit of a body that is responsible for upholding the law regarding consumers. This is where the UK's Competition and Markets Authority (CMA) responsibility lies. The CMA is a non-governmental organisation, whose role is to protect consumers and ensure businesses comply with consumer law. In February 2020 the CMA raised concerns about add-ons surrounding misleading advertising regarding success rates. It was the first time that the fertility sector was made explicitly aware of its role in adhering to consumer law when it came to add-ons. The following year, the CMA published guidance for fertility clinics to make clear clinics' legal obligation to treat patients fairly [124]. It also published a video and guide to help IVF patients understand their consumer rights [125].

Recently, the CMA published a report on clinics' compliance with consumer law [126]. It revealed that there were compliance issues with most clinics reviewed: clinics failed to provide information about the evidence for, or risks associated with, certain add-ons; and clinics made claims that link success rates to the use of certain add-ons without any, or adequate, explanation of the basis on which the claims were made [126]. The report also highlighted the positive changes that some clinics made in response to concerns regarding compliance. For example, clinics had updated their webpages to provide additional information so that the benefits and risks of add-ons as well as the view of the HFEA were more clearly explained. In addition, some clinics made it clearer where information about the basis for claims that link the use of add-ons to successful treatment outcomes came from. The CMA's recent work has helped to shine a light on clinics' responsibility to be open and honest about the paucity of good quality evidence to support the use of add-ons in improving IVF outcomes. Their task will now be to continue to uphold the standards expected of clinics regarding the advertisement and selling of add-ons through regular audit of clinic websites and materials.

Evidence generated from papers 2 and 3 in this thesis will be useful in guiding clinics who will look to the evidence of safety and efficacy behind time-lapse and GM-CSF containing culture media when describing them on their websites. It is hoped that paper 5 will provide support

the CMA's role in protecting customers (those who undergo IVF in this case), by providing evidence that patients prioritise hope of a child over considerations of safety, efficacy or costs when considering add-ons. It is hoped that the HFEA will continue to use evidence from papers 2 and 3 to inform its traffic light system for add-ons, and that paper 5 will provide robust qualitative evidence to inform debate within the Scientific and Clinical Advances Advisory Committee (SCAAC). Such debate might include considering the findings of VALUE when deciding how unproven treatment add-ons might be introduced into clinical practice, and what level of evidence of safety and efficacy is required prior to their adoption.

Future research

Settling on a scientific definition of 'add-on' will be vital to ensuring that novel add-ons are identified in a timely manner, and subject to scrutiny regarding effectiveness and safety. A definition has been attempted by the HFEA, however is fraught with flaws. The HFEA definition rests upon the presence of three criteria: (i) being 'optional'; (ii) claiming to be effective at improving chances of success; and (iii) costing the patient extra. However, where does this leave add-ons that are absorbed into the headline cost of IVF treatment if the clinic uses them 'routinely' for all private patients? According to the HFEA definition, they would not be deemed an add-on. Therefore, explicit consent would not be sought for their use. The costs would inevitably be passed onto patients through the overall cost of IVF, but without their explicit understanding that the add-on is non evidence-based and optional.

To reach a definition in an evidence-based manner, a Delphi method could be adopted, whereby a structured group of multi-disciplinary experts involved in IVF and add-ons research around the globe would convene to undertake questionnaires in two or more rounds. After each round, a facilitator would provide an anonymised summary of the experts' answers from the previous round as well as the reasons they provided for their judgements. Experts would then encourage individuals to revise their earlier answers in light of the replies from other members of the panel. The process would stop after a predefined stop criterion such as number of rounds or achievement of consensus, and the mean or median scores of the final rounds determine the result [127].

The Delphi method has been successfully adopted to reach consensus in other controversial areas of medicine. For example, it was used to develop a core outcome set for future infertility trials in 2021[128]. This Delphi study reached conclusions on the definition of core outcomes deemed of critical importance to patients, to help harmonise and strengthen the impact of forthcoming studies and reduce research waste in fertility. Over 80 specialty journals have committed to implementing this core outcome set [128].

Further quantitative research regarding the effectiveness and safety of all add-ons is needed in the form of properly powered, clinically meaningful RCTs. Add-ons will cease to be add-ons once they are of proven effectiveness. Until this time, they remain of unknown or uncertain effectiveness, and will remain as such until reliable studies are undertaken. The current major flaw with the available evidence is that the studies are often small and inadequately powered to answer the questions that patients, embryologists, and clinicians need the answers to: does the add-on improve livebirth rates and is it safe, i.e. does it worsen miscarriage rates beyond baseline. In addition to being underpowered, there are often methodological flaws within the available studies, such as the problem with selection bias. Many studies opt to randomise oocytes opposed to women, or calculate livebirth rates per embryo transferred opposed to per woman randomised. The outcomes measured are often disparate and difficult to combined in systematic reviews. For example, there are myriad different definitions of clinical pregnancy, miscarriage and even of livebirth. Encouraging study collaborators to use the aforementioned core outcome set would help standardise outcome selection, data collection and reporting [128].

At the top of the evidence pyramid lie systematic reviews combining RCTs. High-quality, methodologically robust systematic reviews, such as those undertaken by Cochrane Gynaecology and Fertility need to continue to be produced on novel add-ons emerging onto the scene, and established reviews need to be periodically updated when new trials are published. A clearly defined protocol should be published before any review is undertaken to ensure transparency and avoid 'fishing' for statistically significant results. Given the plethora of low quality, small RCTs available, future systematic reviews should consider confining meta-analysis to trials of low risks of bias. The difficulty with combining all available trials is that the heterogeneity between studies becomes so high, that drawing meaningful

conclusions becomes impossible. One example of this is the Cochrane review on endometrial scratching before IVF [108]. There were 38 eligible studies, many of which were very small and had various biases detected. When all studies were combined, the I^2 result, which is a measure of heterogeneity, was so high that drawing conclusions from the summary estimate of the meta-analysis was impossible. This meant that a decision needed to be made regarding which studies would be included within the meta-analysis. It was decided that only studies with low selection bias would be included in the meta-analysis which reduced heterogeneity, making the drawing of conclusions more meaningful [108]. Despite the need for high-quality systematic reviews, there is also the need to avoid research waste, especially on poorly designed reviews, where the methodology means that their publication only confuses the picture, making it harder for readers to navigate the evidence. Therefore, my decision to continue to periodically update the two Cochrane titles published in this thesis will be important to provide reproducible, robust evidence for the world.

The future of evidence-based medicine within IVF

Reproductive medicine and IVF stands at a crossroads philosophically and ethically. Throughout its relatively short history, it has relied upon innovation, experimentation, and development of techniques, most of which never underwent formal assessment of effectiveness or safety in the context of RCTs [129]. This is often the argument put forth by those who rebuff the notion that add-ons should not be introduced until RCT evidence is available. In 1978, when IVF was being pioneered, the ethical argument for undertaking non evidence-based treatments was that the only alternative was adoption or childlessness. Reproductive medicine now finds itself at a stage where IVF represents the most effective fertility treatment available for most causes of infertility, offering an average 26% livebirth rate per cycle started [1, 2]. With that in mind, is it right that novel IVF technologies, treatments, and procedures should not be subject to the same robust assessment of safety and efficacy as set by other spheres of medicine?

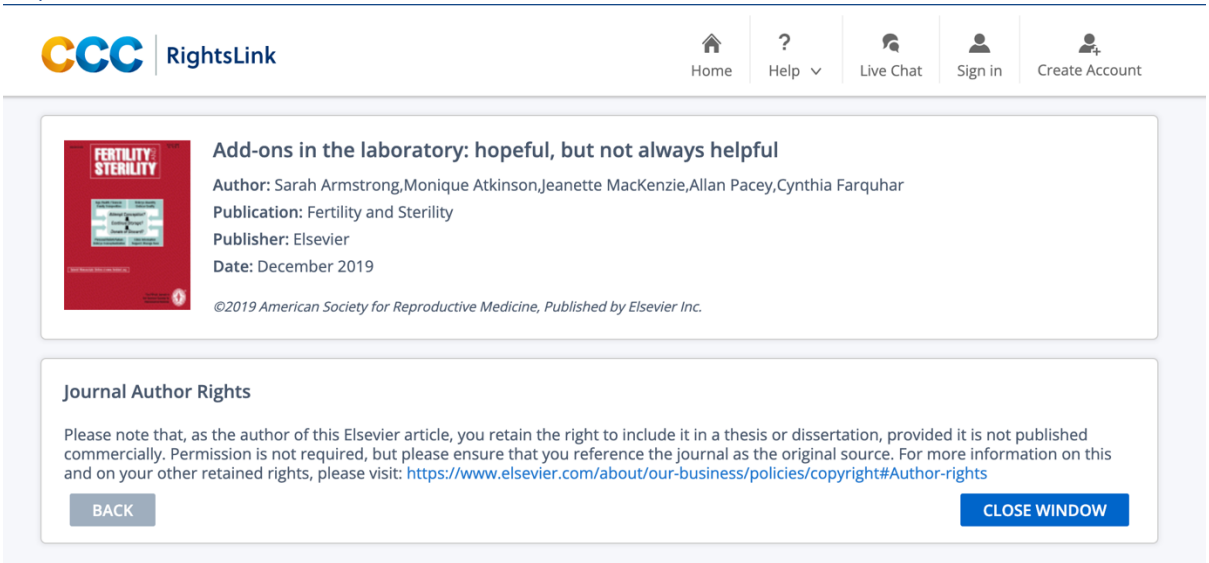
Trials assessing these outcomes in IVF are costly and time-consuming owing to the lengthy follow-up time to livebirth, plus the challenge of collecting data on cumulative outcomes from embryo transfers which can be difficult to keep track of. However, despite these difficulties,

we must remember that at the core of all that we do are the patients we serve. Patients will continue to opt for non evidence-based extras if they offer hope above and beyond what conventional IVF can offer.

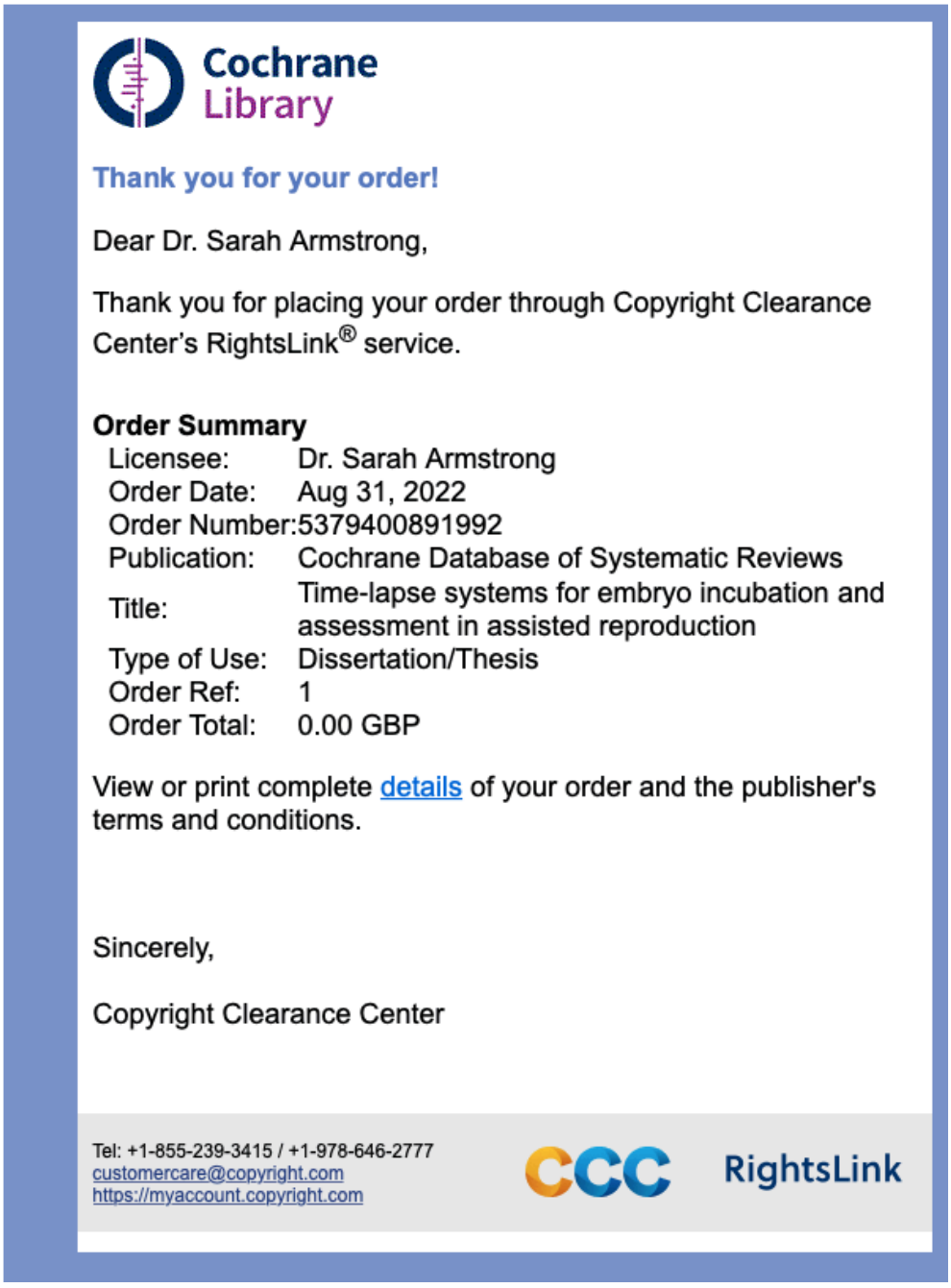
The drive to assess add-ons will not come from industry itself given that the requirements for licensing will have often already been met (i.e., repurposing of an established medical treatment for a novel fertility purpose) or the add-on will not be required to undergo phase 3 trials prior to licensing owing to its definition as a device. Therefore, it is beholden on the scientific community, clinicians, embryologists and fertility nurses to strive for high quality evidence on add-ons to allow patients to be fully informed in their decision making, and to support those practising within IVF clinics to fulfil their duties of informed consent.

Appendix 1: RightsLink Licences

Paper 1:



Paper 2:



Cochrane Library

Thank you for your order!

Dear Dr. Sarah Armstrong,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

Licensee:	Dr. Sarah Armstrong
Order Date:	Aug 31, 2022
Order Number:	5379400891992
Publication:	Cochrane Database of Systematic Reviews
Title:	Time-lapse systems for embryo incubation and assessment in assisted reproduction
Type of Use:	Dissertation/Thesis
Order Ref:	1
Order Total:	0.00 GBP

View or print complete details of your order and the publisher's terms and conditions.

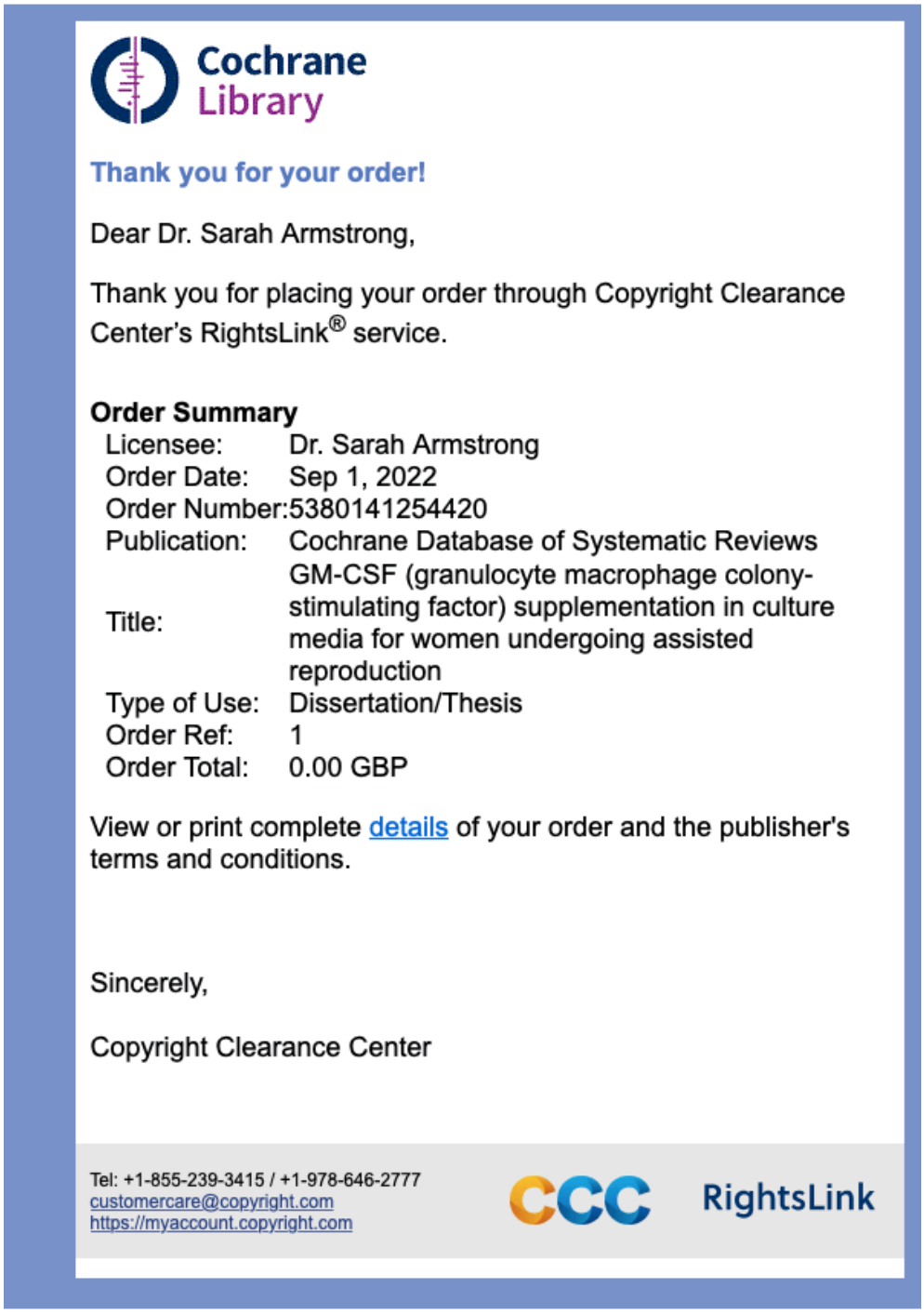
Sincerely,

Copyright Clearance Center

Tel: +1-855-239-3415 / +1-978-646-2777
customercare@copyright.com
https://myaccount.copyright.com

CCC RightsLink

Paper 3:



Cochrane Library

Thank you for your order!

Dear Dr. Sarah Armstrong,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

Licensee:	Dr. Sarah Armstrong
Order Date:	Sep 1, 2022
Order Number:	5380141254420
Publication:	Cochrane Database of Systematic Reviews
Title:	GM-CSF (granulocyte macrophage colony-stimulating factor) supplementation in culture media for women undergoing assisted reproduction
Type of Use:	Dissertation/Thesis
Order Ref:	1
Order Total:	0.00 GBP

View or print complete details of your order and the publisher's terms and conditions.

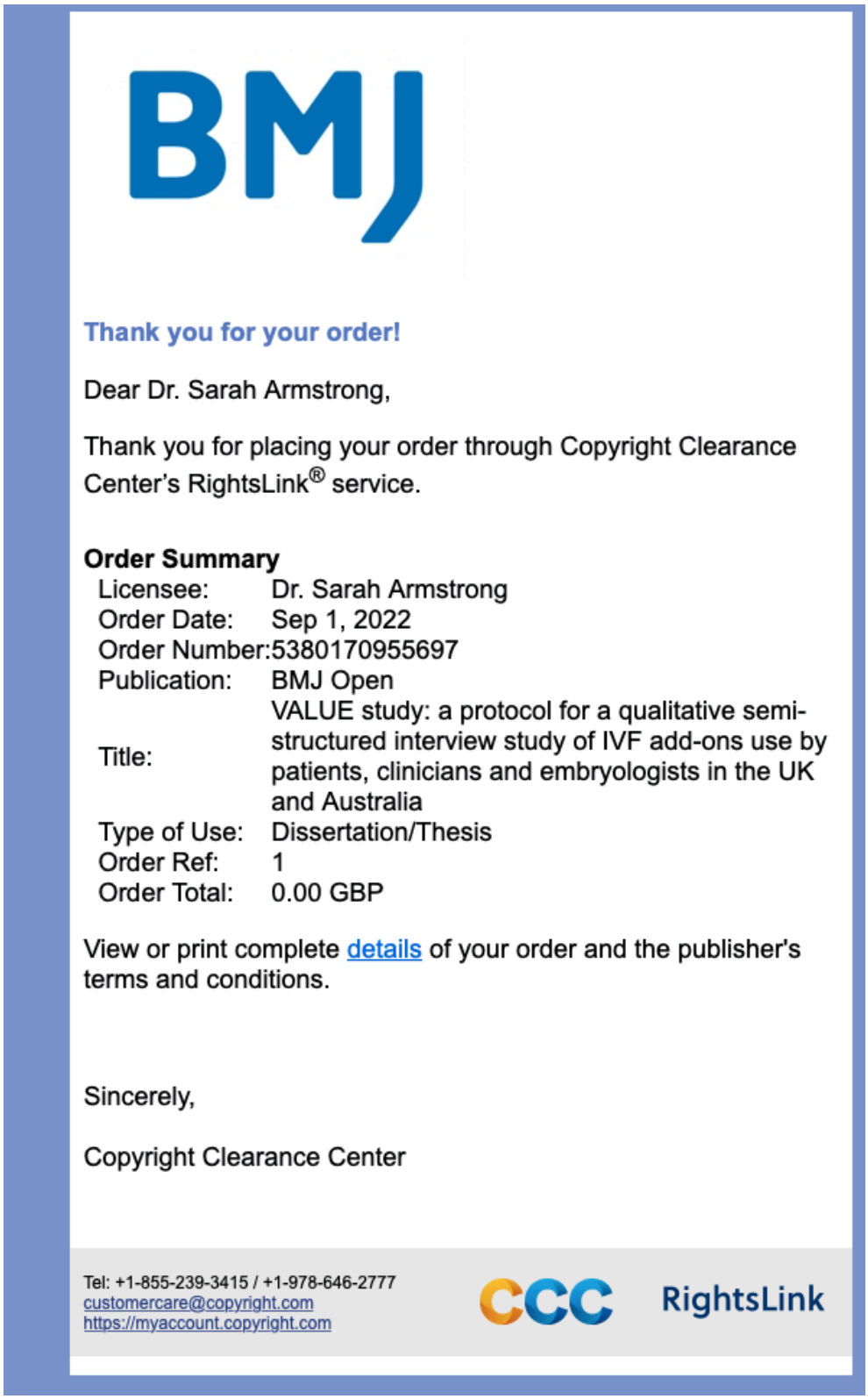
Sincerely,

Copyright Clearance Center

Tel: +1-855-239-3415 / +1-978-646-2777
customercare@copyright.com
https://myaccount.copyright.com

CCC RightsLink

Paper 4:



BMJ

Thank you for your order!

Dear Dr. Sarah Armstrong,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

Licensee:	Dr. Sarah Armstrong
Order Date:	Sep 1, 2022
Order Number:	5380170955697
Publication:	BMJ Open
Title:	VALUE study: a protocol for a qualitative semi-structured interview study of IVF add-ons use by patients, clinicians and embryologists in the UK and Australia
Type of Use:	Dissertation/Thesis
Order Ref:	1
Order Total:	0.00 GBP

View or print complete details of your order and the publisher's terms and conditions.

Sincerely,

Copyright Clearance Center

Tel: +1-855-239-3415 / +1-978-646-2777
customercare@copyright.com
https://myaccount.copyright.com

CCC RightsLink

Permission from HFEA to use traffic light image

From: Sarah C Armstrong <scarmstrong1@sheffield.ac.uk>
Sent: 03 February 2023 12:15
To: Communications <Communications@HFEA.GOV.UK>
Subject: Use of screengrab

Dear HFEA communications team,

Re: screen grab of traffic lights and add-ons

I would like to use a JPEG image of the traffic light webpage in my MD thesis with your permission. My MD is titled 'IVF add-ons: The Quantitative and Qualitative Evidence Behind Their Use'.

I attach the image I would like to use.

Best wishes
Sarah

From: Lauren Snaith <Lauren.Snaith@HFEA.GOV.UK>
Subject: RE: Use of screengrab
Date: 3 February 2023 at 15:44:46 GMT
To: Sarah C Armstrong <scarmstrong1@sheffield.ac.uk>, Communications <Communications@HFEA.GOV.UK>

Dear Sarah,

Thank you for your email.

Yes this is absolutely fine. Thank you for checking.

If it's possible, could you provide a reference to the treatment add-ons ratings system on the HFEA website so people can find this if needed?

Best wishes
Lauren

Lauren Snaith (she/her)
External Communications Manager

Human Fertilisation and Embryology Authority
020 7291 8944
www.hfea.gov.uk
@HFEA

Appendix 2: Co-author permissions

Dear Allan,

As you know, I am in the process of apply for an MD with the University of Sheffield ('IVF add-ons: the quantitative and qualitative evidence behind their use'). My thesis is in publication format, and as a co-author of the included publications below, I'd be grateful if you could indicate your consent to me including these papers by signing below and returning by email at your earliest convenience.

Thank you.
Best wishes,

Sarah Armstrong

List of co-authored papers:

Paper 1: **Armstrong S**, Atkinson M, MacKenzie J, Pacey A, Farquhar C. Add-ons in the laboratory: hopeful, but not always helpful. *Fertility and Sterility* 2019, 112(6):994-999. DOI: https://doi.org/10.1016/j.fertnstert.2019.10.031

Paper 2: **Armstrong S**, Bhide P, Jordan V, Pacey A, Marjoribanks J, Farquhar C. Time-lapse systems for embryo incubation and assessment in assisted reproduction. *Cochrane Database of Systematic Reviews* 2019, Issue 5. Art. No.: CD011320. DOI: 10.1002/14651858.CD011320.pub4.

Paper 3: **Armstrong S**, MacKenzie J, Woodward B, Pacey A, Farquhar C. GM-CSF (granulocyte macrophage colony stimulating factor) supplementation in culture media for women undergoing assisted reproductive technology (ART). *Cochrane Database of Systematic Reviews* 2019, Issue 12. Art. No.:CD013497. DOI: 10.1002/14651858.CD013497.

Paper 4: **Armstrong SC**, Lensen S, Vaughan E, Wainwright E, Peate M, Balen A, Farquhar C, Pacey A. VALUE study: a protocol for a qualitative semi-structured interview study of IVF add-ons use by patients, clinicians and embryologists in the UK and Australia *BMJ Open* 2021;**11:**e047307. doi: 10.1136/bmjopen-2020-047307

Paper 5: **Armstrong SC,** Vaughan E, Lensen S, Caughey L, Farquhar C, Pacey A, Balen A, Peate M, Wainwright E. Patient and professional perspectives about using in vitro fertilisation add-ons in the UK and Australia: a qualitative study. Undergoing editorial review with Human Reproduction

I agree to the inclusion of the above listed published papers in Dr Sarah Armstrong MD thesis, to be submitted to the University of Sheffield.

Name: Allan Pacey

Signed:

Date: 22nd September 2022

Dear Priya

I am in the process of apply for an MD with the University of Sheffield ('IVF add-ons: the quantitative and qualitative evidence behind their use'). My thesis is in publication format, and as a co-author of the included publications below, I'd be grateful if you could indicate your consent to me including these papers by signing below and returning by email at your earliest convenience.

Thank you.
Best wishes,

Sarah Armstrong

List of co-authored papers:

Paper 2: **Armstrong S**, Bhide P, Jordan V, Pacey A, Marjoribanks J, Farquhar C. Time-lapse systems for embryo incubation and assessment in assisted reproduction. *Cochrane Database of Systematic Reviews* 2019, Issue 5. Art. No.: CD011320. DOI: 10.1002/14651858.CD011320.pub4.

I agree to the inclusion of the above listed published paper in Dr Sarah Armstrong's MD thesis, to be submitted to the University of Sheffield.

Name: Priya Bhide

Signed: 

Date:21/09/2022

Dear Monique

I am in the process of apply for an MD with the University of Sheffield ('IVF add-ons: the quantitative and qualitative evidence behind their use'). My thesis is in publication format, and as a co-author of the included publications below, I'd be grateful if you could indicate your consent to me including these papers by signing below and returning by email at your earliest convenience.

Thank you.
Best wishes,

Sarah Armstrong

List of co-authored papers:

Paper 1: **Armstrong S**, Atkinson M, MacKenzie J, Pacey A, Farquhar C. Add-ons in the laboratory: hopeful, but not always helpful. *Fertility and Sterility* 2019, 112(6):994-999. DOI: https://doi.org/10.1016/j.fertnstert.2019.10.031

I agree to the inclusion of the above listed published paper in Dr Sarah Armstrong's MD thesis, to be submitted to the University of Sheffield.

Name: Monique Atkinson

Signed:

Date: 27/09/2022

Dear Lucy

I am in the process of apply for an MD with the University of Sheffield ('IVF add-ons: the quantitative and qualitative evidence behind their use'). My thesis is in publication format, and as a co-author of the included publications below, I'd be grateful if you could indicate your consent to me including these papers by signing below and returning by email at your earliest convenience.

Thank you.
Best wishes,

Sarah Armstrong

List of co-authored papers:

Paper 5: **Armstrong SC,** Vaughan E, Lensen S, Caughey L, Farquhar C, Pacey A, Balen A, Peate M, Wainwright E. Patient and professional perspectives about using in vitro fertilisation add-ons in the UK and Australia: a qualitative study. Undergoing editorial review with Human Reproduction

I agree to the inclusion of the above listed published papers in Dr Sarah Armstrong MD thesis, to be submitted to the University of Sheffield.

Name: Lucy Caughey

Signed: L Caughey

Date: 21 September 2022

Dear Bryan

I am in the process of apply for an MD with the University of Sheffield ('IVF add-ons: the quantitative and qualitative evidence behind their use'). My thesis is in publication format, and as a co-author of the included publications below, I'd be grateful if you could indicate your consent to me including these papers by signing below and returning by email at your earliest convenience.

Thank you.
Best wishes,

Sarah Armstrong

List of co-authored papers:

Paper 3: **Armstrong S**, MacKenzie J, Woodward B, Pacey A, Farquhar C. GM-CSF (granulocyte macrophage colony stimulating factor) supplementation in culture media for women undergoing assisted reproductive technology (ART). *Cochrane Database of Systematic Reviews* 2019, Issue 12. Art. No.:CD013497. DOI: 10.1002/14651858.CD013497.

I agree to the inclusion of the above listed published papers in Dr Sarah Armstrong's MD thesis, to be submitted to the University of Sheffield.

Name: Bryan Woodward

Signed: 

Date: 21.09.22

Dear Cindy

I am in the process of apply for an MD with the University of Sheffield ('IVF add-ons: the quantitative and qualitative evidence behind their use'). My thesis is in publication format, and as a co-author of the included publications below, I'd be grateful if you could indicate your consent to me including these papers by signing below and returning by email at your earliest convenience.

Thank you.
Best wishes,

Sarah Armstrong

List of co-authored papers:

Paper 1: **Armstrong S**, Atkinson M, MacKenzie J, Pacey A, Farquhar C. Add-ons in the laboratory: hopeful, but not always helpful. *Fertility and Sterility* 2019, 112(6):994-999. DOI: https://doi.org/10.1016/j.fertnstert.2019.10.031

Paper 2: **Armstrong S**, Bhide P, Jordan V, Pacey A, Marjoribanks J, Farquhar C. Time-lapse systems for embryo incubation and assessment in assisted reproduction. *Cochrane Database of Systematic Reviews* 2019, Issue 5. Art. No.: CD011320. DOI: 10.1002/14651858.CD011320.pub4.

Paper 3: **Armstrong S**, MacKenzie J, Woodward B, Pacey A, Farquhar C. GM-CSF (granulocyte macrophage colony stimulating factor) supplementation in culture media for women undergoing assisted reproductive technology (ART). *Cochrane Database of Systematic Reviews* 2019, Issue 12. Art. No.:CD013497. DOI: 10.1002/14651858.CD013497.

Paper 4: **Armstrong SC**, Lensen S, Vaughan E, Wainwright E, Peate M, Balen A, Farquhar C, Pacey A. VALUE study: a protocol for a qualitative semi-structured interview study of IVF add-ons use by patients, clinicians and embryologists in the UK and Australia *BMJ Open* 2021;**11:**e047307. doi: 10.1136/bmjopen-2020-047307

Paper 5: **Armstrong SC,** Vaughan E, Lensen S, Caughey L, Farquhar C, Pacey A, Balen A, Peate M, Wainwright E. Patient and professional perspectives about using in vitro fertilisation add-ons in the UK and Australia: a qualitative study. Undergoing editorial review with Human Reproduction

I agree to the inclusion of the above listed published papers in Dr Sarah Armstrong MD thesis, to be submitted to the University of Sheffield.

Name: Cynthia Farquhar

Signed: 22/9/22

Dear Jeanette

I am in the process of apply for an MD with the University of Sheffield ('IVF add-ons: the quantitative and qualitative evidence behind their use'). My thesis is in publication format, and as a co-author of the included publications below, I'd be grateful if you could indicate your consent to me including these papers by signing below and returning by email at your earliest convenience.

Thank you.
Best wishes,

Sarah Armstrong

List of co-authored papers:

Paper 1: **Armstrong S**, Atkinson M, MacKenzie J, Pacey A, Farquhar C. Add-ons in the laboratory: hopeful, but not always helpful. *Fertility and Sterility* 2019, 112(6):994-999. DOI: https://doi.org/10.1016/j.fertnstert.2019.10.031

Paper 3: **Armstrong S**, MacKenzie J, Woodward B, Pacey A, Farquhar C. GM-CSF (granulocyte macrophage colony stimulating factor) supplementation in culture media for women undergoing assisted reproductive technology (ART). *Cochrane Database of Systematic Reviews* 2019, Issue 12. Art. No.:CD013497. DOI: 10.1002/14651858.CD013497.

I agree to the inclusion of the above listed published papers in Dr Sarah Armstrong's MD thesis, to be submitted to the University of Sheffield.

Name: Jeanette MacKenzie

Signed:

Date: 22nd Sep 2022

Dear Vanessa

I am in the process of apply for an MD with the University of Sheffield ('IVF add-ons: the quantitative and qualitative evidence behind their use'). My thesis is in publication format, and as a co-author of the included publication below, I'd be grateful if you could indicate your consent to me including this paper by signing below and returning by email at your earliest convenience.

Thank you.
Best wishes,

Sarah Armstrong

List of co-authored papers:

Paper 2: **Armstrong S**, Bhide P, Jordan V, Pacey A, Marjoribanks J, Farquhar C. Time-lapse systems for embryo incubation and assessment in assisted reproduction. *Cochrane Database of Systematic Reviews* 2019, Issue 5. Art. No.: CD011320. DOI: 10.1002/14651858.CD011320.pub4.

I agree to the inclusion of the above listed published paper in Dr Sarah Armstrong's MD thesis, to be submitted to the University of Sheffield.

Name: Vanessa Jordan

Signed: 

Date: 22/09/22

Dear Emily,

I am in the process of apply for an MD with the University of Sheffield ('IVF add-ons: the quantitative and qualitative evidence behind their use'). My thesis is in publication format, and as a co-author of the included publications below, I'd be grateful if you could indicate your consent to me including these papers by signing below and returning by email at your earliest convenience.

Thank you.
Best wishes,

Sarah Armstrong

List of co-authored papers:

Paper 4: **Armstrong SC**, Lensen S, Vaughan E, Wainwright E, Peate M, Balen A, Farquhar C, Pacey A. VALUE study: a protocol for a qualitative semi-structured interview study of IVF add-ons use by patients, clinicians and embryologists in the UK and Australia *BMJ Open* 2021;**11:**e047307. doi: 10.1136/bmjopen-2020-047307

Paper 5: **Armstrong SC,** Vaughan E, Lensen S, Caughey L, Farquhar C, Pacey A, Balen A, Peate M, Wainwright E. Patient and professional perspectives about using in vitro fertilisation add-ons in the UK and Australia: a qualitative study. Undergoing editorial review with Human Reproduction

I agree to the inclusion of the above listed published papers in Dr Sarah Armstrong's MD thesis, to be submitted to the University of Sheffield.

Name: Emily Vaughan

Signed: Emily Vaughan

Date:
22/9/22

Dear Elaine,

I am in the process of apply for an MD with the University of Sheffield ('IVF add-ons: the quantitative and qualitative evidence behind their use'). My thesis is in publication format, and as a co-author of the included publications below, I'd be grateful if you could indicate your consent to me including these papers by signing below and returning by email at your earliest convenience.

Thank you.
Best wishes,

Sarah Armstrong

List of co-authored papers:

Paper 4: **Armstrong SC**, Lensen S, Vaughan E, Wainwright E, Peate M, Balen A, Farquhar C, Pacey A. VALUE study: a protocol for a qualitative semi-structured interview study of IVF add-ons use by patients, clinicians and embryologists in the UK and Australia *BMJ Open* 2021;**11:**e047307. doi: 10.1136/bmjopen-2020-047307

Paper 5: **Armstrong SC,** Vaughan E, Lensen S, Caughey L, Farquhar C, Pacey A, Balen A, Peate M, Wainwright E. Patient and professional perspectives about using in vitro fertilisation add-ons in the UK and Australia: a qualitative study. Undergoing editorial review with Human Reproduction

I agree to the inclusion of the above listed published papers in Dr Sarah Armstrong's MD thesis, to be submitted to the University of Sheffield.

Name: Elaine Wainwright

Signed: 

Date: 210922

Dear Adam,

I am in the process of apply for an MD with the University of Sheffield ('IVF add-ons: the quantitative and qualitative evidence behind their use'). My thesis is in publication format, and as a co-author of the included publications below, I'd be grateful if you could indicate your consent to me including these papers by signing below and returning by email at your earliest convenience.

Thank you.
Best wishes,

Sarah Armstrong

List of co-authored papers:

Paper 4: **Armstrong SC**, Lensen S, Vaughan E, Wainwright E, Peate M, Balen A, Farquhar C, Pacey A. VALUE study: a protocol for a qualitative semi-structured interview study of IVF add-ons use by patients, clinicians and embryologists in the UK and Australia *BMJ Open* 2021;**11:**e047307. doi: 10.1136/bmjopen-2020-047307

Paper 5: **Armstrong SC,** Vaughan E, Lensen S, Caughey L, Farquhar C, Pacey A, Balen A, Peate M, Wainwright E. Patient and professional perspectives about using in vitro fertilisation add-ons in the UK and Australia: a qualitative study. Undergoing editorial review with Human Reproduction

I agree to the inclusion of the above listed published papers in Dr Sarah Armstrong's MD thesis, to be submitted to the University of Sheffield.

Name: Adam Balen

Signed:  21.09.2022

Date:

Dear Sarah,

I am in the process of apply for an MD with the University of Sheffield ('IVF add-ons: the quantitative and qualitative evidence behind their use'). My thesis is in publication format, and as a co-author of the included publications below, I'd be grateful if you could indicate your consent to me including these papers by signing below and returning by email at your earliest convenience.

Thank you.
Best wishes,

Sarah Armstrong

List of co-authored papers:

Paper 4: **Armstrong SC**, Lensen S, Vaughan E, Wainwright E, Peate M, Balen A, Farquhar C, Pacey A. VALUE study: a protocol for a qualitative semi-structured interview study of IVF add-ons use by patients, clinicians and embryologists in the UK and Australia *BMJ Open* 2021;11:e047307. doi: 10.1136/bmjopen-2020-047307

Paper 5: **Armstrong SC,** Vaughan E, Lensen S, Caughey L, Farquhar C, Pacey A, Balen A, Peate M, Wainwright E. Patient and professional perspectives about using in vitro fertilisation add-ons in the UK and Australia: a qualitative study. Undergoing editorial review with Human Reproduction

I agree to the inclusion of the above listed published papers in Dr Sarah Armstrong's MD thesis, to be submitted to the University of Sheffield.

Name: Sarah Lensen

Signed: 

Date: 26 September 2022

Dear Michelle,

I am in the process of apply for an MD with the University of Sheffield ('IVF add-ons: the quantitative and qualitative evidence behind their use'). My thesis is in publication format, and as a co-author of the included publications below, I'd be grateful if you could indicate your consent to me including these papers by signing below and returning by email at your earliest convenience.

Thank you.
Best wishes,

Sarah Armstrong

List of co-authored papers:

Paper 4: **Armstrong SC**, Lensen S, Vaughan E, Wainwright E, Peate M, Balen A, Farquhar C, Pacey A. VALUE study: a protocol for a qualitative semi-structured interview study of IVF add-ons use by patients, clinicians and embryologists in the UK and Australia *BMJ Open* 2021;**11**:e047307. doi: 10.1136/bmjopen-2020-047307

Paper 5: **Armstrong SC,** Vaughan E, Lensen S, Caughey L, Farquhar C, Pacey A, Balen A, Peate M, Wainwright E. Patient and professional perspectives about using in vitro fertilisation add-ons in the UK and Australia: a qualitative study. Undergoing editorial review with Human Reproduction

I agree to the inclusion of the above listed published papers in Dr Sarah Armstrong's MD thesis, to be submitted to the University of Sheffield.

Name: Michelle Peate

Signed:

Date: 26/9/22

References

1. Newman JE PR, Chambers GM. Assisted reproductive technology in Australia and New Zealand 2019. Sydney: National Perinatal Epidemiology and Statistics Unit, the University of New South Wales; 2021.

2. HFEA. Fertility treatment 2019: trends and figures HFEA: HFEA; 2021 [Available from: https://www.hfea.gov.uk/about-us/publications/research-and-data/fertility-treatment-2019-trends-and-figures/.

3. Armstrong S, Atkinson M, MacKenzie J, Pacey A, Farquhar C. Add-ons in the laboratory: hopeful, but not always helpful. Fertil Steril. 2019;112(6):994-9.

4. McAuley J. Panorama: Inside Britain's Fertility Business. 2016.

5. Belli E L-WF. Pregnancy: Calls to regulate 'cruel' IVF add-ons. BBC news2022.

6. Tsigdinos P. The Big IVF Add-On Racket. The New York Times. 2019 12/12/19.

7. Heneghan C, Spencer EA, Bobrovitz N, Collins DRJ, Nunan D, Plüddemann A, et al. Lack of evidence for interventions offered in UK fertility centres. BMJ. 2016;355:392-4.

8. Harper J, Jackson E, Sermon K, Aitken RJ, Harbottle S, Mocanu E, et al. Adjuncts in the IVF laboratory: where is the evidence for 'add-on' interventions? Hum Reprod. 2017;32(3):485-91.

9. Repping S. Evidence-based medicine and infertility treatment. Lancet. 2019;393(10170):380-2.

10. Cochrane. Special Collection: In vitro fertilisation - effectiveness of add-ons Cochranelibrary.com2019 [Available from: https://www.cochrane.org/news/special-collection-vitro-fertilisation-effectiveness-add-ons.

11. Authority CaM. Fertility treatment: A guide for Clinics. https://www.gov.uk/government/publications/fertility-treatment-a-guide-for-clinics: Competition and Markets Authority; 2021 [

12. HFEA. Fertility 2022 - Julia Chain, Chair of the HFEA https://www.hfea.gov.uk/about-us/news-and-press-releases/2022-news-and-press-releases/fertility-2022-julia-chain-chair-of-the-hfea/: Human Fertilisation and Embryology Authority; 2022 [

13. HFEA. Treatment add-ons with limited evidence 2021 [Available from: https://www.hfea.gov.uk/treatments/treatment-add-ons/.

14. HFEA. The responsible use of treatment add-ons in fertility services: a consensus statement 2019 [Available from: www.hfea.gov.uk/media/2792/treatment-add-ons-consensus-statement-final.pdf.

15. HFEA. Treatment add-ons 2020 [Available from: www.hfea.gov.uk/treatments/explore-all-treatments/treatment-add-ons/.

16. HFEA. Cost and funding 2018 [Available from: https://www.hfea.gov.uk/treatments/explore-all-treatments/costs-and-funding/.

17. NHS. IVF 2021 [03/2/23]. Available from: https://www.nhs.uk/conditions/ivf/.

18. Fertilitynetworkuk. NHS Funding 2018 [Available from: http://fertilitynetworkuk.org/for-those-trying-to-become-parents/nhs-funding/?gclid=Cj0KCQjw6J7YBRC4ARIsAJMXXsfQG3sEKQMdIl_HXIk221D9mljhMf8SwcVpe0BZRYWE3cOzhs5m-TAaAr_9EALw_wcB.

19. NICE. Fertility Problems: assessment and treatment (CG156) 2017 [Available from: https://www.nice.org.uk/guidance/CG156.

20. fertilitynetworkuk. NHS Funding in England https://fertilitynetworkuk.org/access-support/nhs-funding/england/2022 [Available from: https://fertilitynetworkuk.org/access-support/nhs-funding/england/.

21. fertilitynetworkuk. NHS Funding in Scotland https://fertilitynetworkuk.org/access-support/nhs-funding/scotland/: fertilitynetworkuk; 2022 [Available from: https://fertilitynetworkuk.org/access-support/nhs-funding/scotland/.

22. fertilitynetworkuk. NHS Funding in Northern Ireland https://fertilitynetworkuk.org/access-support/nhs-funding/northern-ireland/#:~:text=NHS%20fertility%20treatment%20is%20not,for%20commissioning%20of%20fertility%20services.: fertilitynetworkuk; 2022 [Available from: https://fertilitynetworkuk.org/access-support/nhs-funding/northern-ireland/#:~:text=NHS%20fertility%20treatment%20is%20not,for%20commissioning%20of%20fertility%20services.

23. fertilitynetworkuk. NHS funding in Wales https://fertilitynetworkuk.org/access-support/nhs-funding/wales/2022 [Available from: https://fertilitynetworkuk.org/access-support/nhs-funding/wales/.

24. HFEA. Fertility treatment 2019: trends and figures Hfea.gov.uk2021 [Available from: https://www.hfea.gov.uk/about-us/publications/research-and-data/fertility-treatment-2019-trends-and-figures/.

25. YouGov. Pilot national fertility patient survey. HFEA; 2018.

26. HFEA. Fertility treatment 'add-ons'. 2017.

27. HFEA. Pilot national fertility patient survey hfea.co.uk2018 [Available from: https://www.hfea.gov.uk/media/2702/pilot-national-fertility-patient-survey-2018.pdf.

28. Lensen S, Hammarberg K, Polyakov A, Wilkinson J, Whyte S, Peate M, et al. How common is add-on use and how do patients decide whether to use them? A national survey of IVF patients. Hum Reprod. 2021;36(7):1854-61.

29. Heneghan C, Spencer EA, Bobrovitz N, Collins DRJ, Nunan D, Plüddemann A, et al. Lack of evidence for interventions offered in UK fertility centres. BMJ. 2016;355.

30. Nardo LG, El-Toukhy T, Stewart J, Balen AH, Potdar N. British Fertility Society Policy and Practice Committee: Adjuvants in IVF: Evidence for good clinical practice. Human Fertility. 2015;18(1):2-15.

31. Nardo LG, Granne I, Stewart J, Policy, Practice Committee of the British Fertility S. Medical adjuncts in IVF: evidence for clinical practice. Human fertility (Cambridge, England). 2009;12(1):1-13.

32. HFEA. How we regulate HFEA2023 [Available from: https://www.hfea.gov.uk/about-us/how-we-regulate/.

33. ESHRE. Regulation: ESHRE reaffirms its position on IVF add-ons but little regulatory consensus on donor anonymity https://www.focusonreproduction.eu/article/ESHRE-News-Regulation-PET: Focus on Reproduction; 2021 [Available from: https://www.focusonreproduction.eu/article/ESHRE-News-Regulation-PET.

34. BFS. Public information 2015 [Available from: https://britishfertilitysociety.org.uk/public-resources/.

35. Heneghan C, Spencer EA, Bobrovitz N, Collins DR, Nunan D, Pluddemann A, et al. Lack of evidence for interventions offered in UK fertility centres. BMJ. 2016;355:i6295.

36. Balen A. Re: Lack of evidence for interventions. BMJ. 2016;355(i6295).

37. Scott L. The biological basis of non-invasive strategies for selection of human oocytes and embryos. Hum Reprod Update. 2003;9(3):237-49.

38. Hong KH, Werner MD, Franasiak JM, Forman EJ, Prodoehl A, Upham K, et al. Embryologist interpretation of time-lapse imaging parameters at the blastocyst stage do not alter selection among transferred euploid blastocysts. Fertil Steril. 2014;1):e305.

39. Scott L. Pronuclear scoring as a predictor of embryo development. Reprod Biomed Online. 2003;6(2):201-14.

40. Shoukir Y, Campana A, Farley T, Sakkas D. Early cleavage of in-vitro fertilized human embryos to the 2-cell stage: a novel indicator of embryo quality and viability. Human reproduction (Oxford, England). 1997;12(7):1531-6.

41. Alpha-Scientists. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. Human reproduction (Oxford, England). 2011;26(6):1270-83.

42. Cutting R, Morroll D, Roberts SA, Pickering S, Rutherford A, Bfs, et al. Elective single embryo transfer: guidelines for practice British Fertility Society and Association of Clinical Embryologists. Human fertility (Cambridge, England). 2008;11(3):131-46.

43. Meseguer M, Rubio I, Cruz M, Basile N, Marcos J, Requena A. Embryo incubation and selection in a time-lapse monitoring system improves pregnancy outcome compared with a standard incubator: a retrospective cohort study. Fertil Steril. 2012;98(6):1481-+.

44. Payne D, Flaherty SP, Barry MF, Matthews CD. Preliminary observations on polar body extrusion and pronuclear formation in human oocytes using time-lapse video cinematography. Hum Reprod. 1997;12(3):532-41.

45. Bhide P, Maheshwari A, Cutting R, Seenan S, Patel A, Khan K, et al. Time lapse imaging: is it time to incorporate this technology into routine clinical practice? Hum Fertil. 2017;20(2):74-9.

46. Vitrolife. General 2015 [Available from: http://www.vitrolife.com/en/Products/EmbryoScope-Time-Lapse-System/Support-Material/FAQ/General/.

47. Armstrong S, Vail A, Mastenbroek S, Jordan V, Farquhar C. Time-lapse in the IVF-lab: how should we assess potential benefit? Human reproduction (Oxford, England). 2015;30(1):3-8.

48. Conaghan J, Chen AA, Willman SP, Ivani K, Chenette PE, Boostanfar R, et al. Improving embryo selection using a computer-automated time-lapse image analysis test plus day 3 morphology: results from a prospective multicenter trial. Fertil Steril. 2013;100(2):412-9.e5.

49. Meseguer M, Herrero J, Tejera A, Hilligsoe KM, Ramsing NB, Remohi J. The use of morphokinetics as a predictor of embryo implantation. Human reproduction (Oxford, England). 2011;26(10):2658-71.

50. Kirkegaard K, Hindkjaer JJ, Grondahl ML, Kesmodel US, Ingerslev HJ. A randomized clinical trial comparing embryo culture in a conventional incubator with a time-lapse incubator. J Assist Reprod Genet. 2012;29(6):565-72.

51. Mio Y, Maeda K. Time-lapse cinematography of dynamic changes occurring during in vitro development of human embryos. Am J Obstet Gynecol. 2008;199(6):5.

52. Kirkegaard K, Agerholm IE, Ingerslev HJ. Time-lapse monitoring as a tool for clinical embryo assessment. Hum Reprod. 2012;27(5):1277-85.

53. Wong C, Chen AA, Behr B, Shen S. Time-lapse microscopy and image analysis in basic and clinical embryo development research. Reprod Biomed Online. 2013;26(2):120-9.

54. Lemmen JG, Agerholm I, Ziebe S. Kinetic markers of human embryo quality using time-lapse recordings of IVF/ICSI-fertilized oocytes. Reprod Biomed Online. 2008;17(3):385-91.

55. Campbell A, Fishel S, Bowman N, Duffy S, Sedler M, Thornton S. Retrospective analysis of outcomes after IVF using an aneuploidy risk model derived from time-lapse imaging without PGS. Reprod Biomed Online. 2013;27(2):140-6.

56. Cruz M, Gadea B, Garrido N, Pedersen KS, Martinez M, Perez-Cano I, et al. Embryo quality, blastocyst and ongoing pregnancy rates in oocyte donation patients whose embryos were monitored by time-lapse imaging. J Assist Reprod Genet. 2011;28(7):569-73.

57. Rubio I, Galan A, Larreategui Z, Ayerdi F, Bellver J, Herrero J, et al. Clinical validation of embryo culture and selection by morphokinetic analysis: a randomized, controlled trial of the EmbryoScope. Fertil Steril. 2014;102(5):1287-94.e5.

58. Kovacs P, Matyas S, Forgacs V, Sajgo A, Rarosi F, Pribenszky C. Time-lapse embryo selection for single blastocyst transfer-results of a multicenter, prospective, randomized clinical trial. Fertil Steril. 2013;1):S90.

59. Yang L, Kong X, Zhang S, Dai J, Gong F, Lu G, et al. Single embryo transfer on cleavage-stage(D3) using Time-lapse selection VS on blastocyst(D5) using traditional morphological selection in patients with good prognosis: a prospective randomized controlled trial. Hum Reprod. 2017;32:102-3.

60. Park H, Bergh C, Selleskog U, Thurin-Kjellberg A, Lundin K. No benefit of culturing embryos in a closed system compared with a conventional incubator in terms of number of

good quality embryos: results from an RCT. Human reproduction (Oxford, England). 2015;30(2):268-75.

61. Kahraman S, Çetinkaya M, Pirkevi C, Yelke H, Kumtepe Y. Comparison of Blastocyst Development and Cycle Outcome in Patients with eSET Using Either Conventional or Time Lapse Incubators. A Prospective Study of Good Prognosis Patients. Journal of Reproductive and Stem Cell Biotechnology. 2013;3(2):55-61.

62. Wu YG, Lazzaroni-Tealdi E, Wang Q, Zhang L, Barad DH, Kushnir VA, et al. Different effectiveness of closed embryo culture system with time-lapse imaging (EmbryoScope(TM)) in comparison to standard manual embryology in good and poor prognosis patients: a prospectively randomized pilot study. Reproductive biology and endocrinology : RB&E. 2016;14(1):49.

63. Kaser DJ, Bormann CL, Missmer SA, Farland LV, Ginsburg ES, Racowsky C. A pilot randomized controlled trial of Day 3 single embryo transfer with adjunctive time-lapse selection versus Day 5 single embryo transfer with or without adjunctive time-lapse selection. Hum Reprod. 2017;32(8):1598-603.

64. Goodman LR, Goldberg J, Falcone T, Austin C, Desai N. Does the addition of time-lapse morphokinetics in the selection of embryos for transfer improve pregnancy rates? A randomized controlled trial. Fertil Steril. 2016;105(2):275-+.

65. Armstrong S, Arroll N, Cree LM, Jordan V, Farquhar C. Time-lapse systems for embryo incubation and assessment in assisted reproduction. Cochrane Database of Systematic Reviews. 2015(2).

66. Pribenszky C, Nilselid AM, Montag M. Time-lapse culture with morphokinetic embryo selection improves pregnancy and live birth chances and reduces early pregnancy loss: a meta-analysis. Reprod Biomed Online. 2017;35(5):511-20.

67. Chen M, Wei S, Hu J, Yuan J, Liu F. Does time-lapse imaging have favorable results for embryo incubation and selection compared with conventional methods in clinical in vitro fertilization? A meta-analysis and systematic review of randomized controlled trials. PloS one. 2017;12(6):e0178720.

68. Polanski LT, Coelho Neto MA, Nastri CO, Navarro PA, Ferriani RA, Raine-Fenning N, et al. Time-lapse embryo imaging for improving reproductive outcomes: systematic review and meta-analysis. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2014;44(4):394-401.

69. Kaser DJ, Racowsky C. Clinical outcomes following selection of human preimplantation embryos with time-lapse monitoring: a systematic review. Hum Reprod Update. 2014;20(5):617-31.

70. Steptoe PC, Edwards RG. Birth after the reimplantation of a human embryo. Lancet. 1978;2(8085):366.

71. Chronopoulou E, Harper JC. IVF culture media: past, present and future. Hum Reprod Update. 2015;21(1):39-55.

72. Youssef MM, Mantikou E, van Wely M, Van der Veen F, Al-Inany HG, Repping S, et al. Culture media for human pre-implantation embryos in assisted reproductive technology cycles. Cochrane Database Syst Rev. 2015(11):CD007876.

73. Nelissen EC, Van Montfoort AP, Coonen E, Derhaag JG, Geraedts JP, Smits LJ, et al. Further evidence that culture media affect perinatal outcome: findings after transfer of fresh and cryopreserved embryos. Human reproduction (Oxford, England). 2012;27(7):1966-76.

74. El Hajj N, Haaf T. Epigenetic disturbances in in vitro cultured gametes and embryos: implications for human assisted reproduction. Fertil Steril. 2013;99(3):632-41.

75. Mantikou E, Youssef MA, van Wely M, van der Veen F, Al-Inany HG, Repping S, et al. Embryo culture media and IVF/ICSI success rates: a systematic review. Hum Reprod Update. 2013;19(3):210-20.

76. Dumoulin JC, Land JA, Van Montfoort AP, Nelissen EC, Coonen E, Derhaag JG, et al. Effect of in vitro culture of human embryos on birthweight of newborns. Human reproduction (Oxford, England). 2010;25(3):605-12.

77. Sjoblom C, Wikland M, Robertson SA. Granulocyte-macrophage colony-stimulating factor promotes human blastocyst development in vitro. Human reproduction (Oxford, England). 1999;14(12):3069-76.

78. Robertson SA, Seamark RF. Granulocyte-macrophage colony stimulating factor (GM-CSF): one of a family of epithelial cell-derived cytokines in the preimplantation uterus. Reprod Fertil Dev. 1992;4(4):435-48.

79. Jasper MJ, Brannstrom M, Olofsson JI, Petrucco OM, Mason H, Robertson SA, et al. Granulocyte-macrophage colony-stimulating factor: presence in human follicular fluid, protein secretion and mRNA expression by ovarian cells. Mol Hum Reprod. 1996;2(8):555-62.

80. Sjoblom C, Wikland M, Robertson SA. Granulocyte-macrophage colony-stimulating factor (GM-CSF) acts independently of the beta common subunit of the GM-CSF receptor to prevent inner cell mass apoptosis in human embryos. Biol Reprod. 2002;67(6):1817-23.

81. Robertson SA, Sjoblom C, Jasper MJ, Norman RJ, Seamark RF. Granulocyte-macrophage colony-stimulating factor promotes glucose transport and blastomere viability in murine preimplantation embryos. Biol Reprod. 2001;64(4):1206-15.

82. Siristatidis C, Vogiatzi P, Salamalekis G, Creatsa M, Vrachnis N, Glujovsky D, et al. Granulocyte macrophage colony stimulating factor supplementation in culture media for subfertile women undergoing assisted reproduction technologies: a systematic review. Int J Endocrinol. 2013;2013:704967.

83. Perricone R, De Carolis C, Giacomelli R, Guarino MD, De Sanctis G, Fontana L. GM-CSF and pregnancy: evidence of significantly reduced blood concentrations in unexplained recurrent abortion efficiently reverted by intravenous immunoglobulin treatment. Am J Reprod Immunol. 2003;50(3):232-7.

84. Robertson SA, Roberts CT, Farr KL, Dunn AR, Seamark RF. Fertility impairment in granulocyte-macrophage colony-stimulating factor-deficient mice. Biol Reprod. 1999;60(2):251-61.

85. Richter KS. The importance of growth factors for preimplantation embryo development and in-vitro culture. Curr Opin Obstet Gynecol. 2008;20(3):292-304.

86. Agerholm I, Loft A, Hald F, Lemmen JG, Munding B, Sorensen PD, et al. Culture of human oocytes with granulocyte-macrophage colony-stimulating factor has no effect on embryonic chromosomal constitution. Reprod Biomed Online. 2010;20(4):477-84.

87. Market Velker BA, Denomme MM, Mann MR. Loss of genomic imprinting in mouse embryos with fast rates of preimplantation development in culture. Biol Reprod. 2012;86(5):143, 1-16.

88. Young LE, Fernandes K, McEvoy TG, Butterwith SC, Gutierrez CG, Carolan C, et al. Epigenetic change in IGF2R is associated with fetal overgrowth after sheep embryo culture. Nat Genet. 2001;27(2):153-4.

89. Origio. EmbryoGen and BlastGen 2018 [Available from: https://www.origio.com/products/embryogen-blastgen/.

90. Carefertility.com. Cost of IVF and fertility treatment 2018 [Available from: https://www.carefertility.com/costs/costs-of-ivf-and-fertility-treatment/.

91. NHS. PR 2016-29: Culture media containing granulocytemacrophage colony stimulating factor (GM-CSF) for in vitro fertilisation (IVF). 2016.

92. Kim D, Kim M, Kang H, Lee H, Park W, Kwon H. The supplementation of granulocyte-macrophage colony-stimulating factor (GM-CSF) in culture medium improves the pregnancy rate in human ART programs. Fertil Steril. 2001;76(3):S6.

93. Shapiro BS, Richter KS, Daneshmand ST, Quinn P, Behr B. Granulocyte-macrophage colony-stimulating factor enhances human embryo development to the blastocyst stage: a randomized study. Fertil Steril. 2003;79:15-6.

94. Ziebe S, Loft A, Povlsen BB, Erb K, Agerholm I, Aasted M, et al. A randomized clinical trial to evaluate the effect of granulocyte-macrophage colony-stimulating factor (GM-CSF) in embryo culture medium for in vitro fertilization. Fertil Steril. 2013;99(6):1600-9.

95. Sfontouris I, Lainas G, Anagnostara K, Kolibiankis E, Lainas T. Effect of granulocyte-macrophage colony-stimulating factor (GM-CSF) on pregnancy rates in patients with multiple unsucessful IVF attempts. Hum Reprod. 2013;28(suppl_1):i60-i2.

96. Zemyarska MS. Is it ethical to provide IVF add-ons when there is no evidence of a benefit if the patient requests it? J Med Ethics. 2019;45(5):346-50.

97. Tsigdinos PM. The Big IVF Add-On Racket. The New York Times. 2019 12/12/2019.

98. Blackwell RE, Carr BR, Chang RJ, DeCherney AH, Haney AF, Keye WR, Jr., et al. Are we exploiting the infertile couple? Fertil Steril. 1987;48(5):735-9.

99. Beauchamp B CJ. Principles of Biomedical Ethics. 7th Edition ed. Oxford, UK: Oxford University Press; 2013.

100. VARTA. Fertility treatment explained 2022 [Available from: https://www.varta.org.au/fertility-treatment/fertility-treatment-explained.

101. ESHRE. Draft Guideline: Good practice recommendations for add-ons in reproductive medicine. 2022.

102. F&S. Fertility and Sterility 2023 [Available from: https://www.fertstert.org/.

103. Armstrong S, Arroll N, Cree LM, Jordan V, Farquhar C. Time-lapse systems for embryo incubation and assessment in assisted reproduction. Cochrane Database Syst Rev. 2015(2):CD011320.

104. Armstrong S, Bhide P, Jordan V, Pacey A, Marjoribanks J, Farquhar C. Time-lapse systems for embryo incubation and assessment in assisted reproduction. Cochrane Database Syst Rev. 2019;5:CD011320.

105. Armstrong S, MacKenzie J, Woodward B, Pacey A, Farquhar C. GM-CSF (granulocyte macrophage colony-stimulating factor) supplementation in culture media for women undergoing assisted reproduction. Cochrane Database Syst Rev. 2020;7:CD013497.

106. Lensen SU, N. IVF add-ons: the latest Cochrane evidence Evidently Cochrane: Cochrane UK; 2020 [Available from: https://www.evidentlycochrane.net/ivf-add-ons-the-latest-cochrane-evidence/#:~:text=These%20Cochrane%20reviews%20found%20that,1%2F4%20%3D%2025%25).

107. HFEA. The responsible use of treatment add-ons in fertility services: a consensus statement hfea.gov.uk2019 [Available from: https://www.hfea.gov.uk/media/2792/treatment-add-ons-consensus-statement-final.pdf.

108. Lensen SF, Armstrong S, Gibreel A, Nastri CO, Raine-Fenning N, Martins WP. Endometrial injury in women undergoing in vitro fertilisation (IVF). Cochrane Database Syst Rev. 2021;6:CD009517.

109. Armstrong SC, Lensen S, Vaughan E, Wainwright E, Peate M, Balen AH, et al. VALUE study: a protocol for a qualitative semi-structured interview study of IVF add-ons use by patients, clinicians and embryologists in the UK and Australia. BMJ Open. 2021;11(5):e047307.

110. HFEA. The responsible use of treatment add-ons in fertility services: a consensus statement. HFEA; 2019.

111. Lensen S, Chen S, Goodman L, Rombauts L, Farquhar C, Hammarberg K. IVF add-ons in Australia and New Zealand: A systematic assessment of IVF clinic websites. Aust N Z J Obstet Gynaecol. 2021;61(3):430-8.

112. 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.

113. Perrotta M, Hamper J. The crafting of hope: Contextualising add-ons in the treatment trajectories of IVF patients. Soc Sci Med. 2021;287:114317.

114. Iacoponi O, van de Wiel L, Wilkinson J, Harper JC. Passion, pressure and pragmatism: how fertility clinic medical directors view IVF add-ons. Reprod Biomed Online. 2022;45(1):169-79.

115. Patton MQ. Qualitative research & evaluation methods : integrating theory and practice. Fourth edition. ed. Thousand Oaks, California: SAGE Publications, Inc.; 2015. xxi, 806 pages p.

116. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology. 2006;3(2):77-101.

117. Kiger ME, Varpio L. Thematic analysis of qualitative data: AMEE Guide No. 131. Med Teach. 2020;42(8):846-54.

118. Shenton A. Strategies for ensuring trustworthiness in qualitative research projects. Education for Information. 2004;22:63-75.

119. Hahlweg P, Harter M, Nestoriuc Y, Scholl I. How are decisions made in cancer care? A qualitative study using participant observation of current practice. BMJ Open. 2017;7(9):e016360.

120. Franklin S. Embodied progress : a cultural account of assisted conception. London: Routledge; 1997. xii, 252 p. p.

121. Patrizio P, Albertini DF, Gleicher N, Caplan A. The changing world of IVF: the pros and cons of new business models offering assisted reproductive technologies. J Assist Reprod Genet. 2022.

122. Business. The fertility business is booming. The Economist. 2019.

123. Human Fertilisation and Embryology Act, (2008).

124. gov.uk. Fertility treatment: A guide for Clinics: Competition and Markets Authority; 2021 [Available from: https://www.gov.uk/government/publications/fertility-treatment-a-guide-for-clinics.

125. gov.uk. Fertility treatment: A guide to your consumer rights: Competition and Markets Authority; 2021 [Available from: https://www.gov.uk/government/publications/fertility-treatment-a-guide-to-your-consumer-rights.

126. gov.uk. Self-funded IVF: consumer law guidance: Competition and Markets Authority; 2022 [Available from: https://www.gov.uk/cma-cases/self-funded-ivf-consumer-law-guidance#full-publication-update-history.

127. Barrett D, Heale R. What are Delphi studies? Evid Based Nurs. 2020;23(3):68-9.

128. Duffy JMN, AlAhwany H, Bhattacharya S, Collura B, Curtis C, Evers JLH, et al. Developing a core outcome set for future infertility research: an international consensus development study. Fertil Steril. 2021;115(1):191-200.

129. Yovich JL, Craft IL. Founding Pioneers of IVF: Independent innovative researchers generating livebirths within 4 years of the first birth. Reprod Biol. 2018;18(4):317-23.