Scientific and Clinical Advances Advisory Committee (SCAAC) - Agenda *(Hybrid)*

Tuesday 25 July 2023, 10:30am – 12:30pm
Wandle room 40 & 41, 2nd Floor, 2 Redman Place, London, E20 1JQ & MS Teams

Agenda item	Time
Refreshments available on arrival – from 10:00am	*10:00am (30')*
1. Mitochondrial donation – background for the committee Tim Child, Chair	*10:30am (15')*
2. Mitochondrial donation discussion **Joining online at 10:45:** Mary Herbert, Louise Hyslop, Jane Stewart, and Rekha Neelakanta Pillai - Newcastle Fertility Centre International Centre for Life	*10:45am(1'15")*
3. Meeting summary and close (SCAAC members only)	*12:00pm (30')*

Scientific and Clinical Advances Advisory Committee (SCAAC) - Agenda *(Hybrid)*

Tuesday 25 July 2023, 1:00pm – 5:00pm
Wandle room 40 & 41, 2nd Floor, 2 Redman Place, London, E20 1JQ & MS Teams

Agenda item	Time
Lunch available on arrival – from 12:30pm	*12:30pm (30')*
1. Welcome, apologies, declarations of interest	*1:00pm (5')*
2. Review of ratings for treatment add-ons Beth Lockwood (HFEA), Andy Vail (The University of Manchester), and Paul Cannon (University of Glasgow)	*1:05pm (1'55")*
Break	*3:00pm (15')*
3. Review of ratings for treatment add-ons Beth Lockwood (HFEA), Andy Vail (The University of Manchester), and Paul Cannon (University of Glasgow)	*3:15pm (1'35")*
4. Any other business	*4:50pm (5')*
5. Meeting summary and close	*4:55pm (5')*

Treatment add-ons rating review – July 2023

Details about this paper

Area(s) of strategy:	The right information
Meeting:	Scientific and Clinical Advances Advisory Committee (SCAAC)
Agenda item:	1
Paper number:	HFEA (25/07/2023) 001
Meeting date:	25 July 2023
Author:	Bethany Lockwood, Policy Manager
Annexes	Annex A. Evidence decision tree for rating add-ons Annex B. References of reviewed studies The independent reviewers report and results of analysis are available as separate meeting papers

Output from this paper

For information/ recommendation?	For recommendation
Recommendation:	Members are asked to: • consider the quality of evidence for each treatment add-on based on the findings from an independent assessor; • agree and recommend ratings for each treatment add-on based on the outcome of live birth rate for the general population; and • agree and recommend ratings for each additional outcome(s) and population(s) relevant to specific treatment add-ons.
Resource implications:	In budget
Implementation date:	Recommendations will be implemented as soon as feasible
Communication(s):	Updates to the HFEA's website information on treatment add-ons and communication of updates to the sector, patients and public.
Organisational risk:	Low

1. Introduction

1.1. The Authority met in **July 2022** and agreed:

- The definition of treatment add-ons that the HFEA will provide information for.
- To move to a five-category rating scale.
- To rate additional outcomes, such as miscarriage, and outcomes for specific patient groups, such as male-factor infertility, in addition to live births for specific add-ons.
- To expand the evidence base in line with SCAAC's recommendation that in the absence of high-quality randomised controlled trials (RCTs) or systematic reviews the evidence base should be expanded to non-randomised studies of intervention (NRSIs).

1.2. At the **February 2023** SCAAC meeting, the new rating system was applied for the first time and the committee were asked to consider the methodology and recommend ratings for each add-on. The methodology for the new rating system was discussed in detail and recommendations were made to ensure completeness of the review process. As a result, ratings were not assigned during this meeting, and it was agreed to assign ratings after the process had been reviewed.

1.3. The Committee agreed to the following next steps:

- The Executive to instruct a specialist librarian to assemble a list of search terms and recommend a methodology to perform the literature searches.
- To share the updated search terms with the review panel to suggest any changes. The panel consisted of:
  - The Chair of SCAAC;
  - At least one member of SCAAC who is a clinician;
  - At least one member of SCAAC who is involved in clinical research/embryology; and
  - One person from the HFEA who is either a member of the scientific policy team or is a member of the Register Research Panel.
- The Executive to conduct the literature searches using the updated search terms.
- The list of papers resulting from the search to be sent to the whole of SCAAC for review. SCAAC to highlight any missing papers to the Executive. SCAAC to take ownership of the list of papers.
- After SCAAC's review, the papers to be sent to the external independent reviewer who will analyse the quality of the evidence base and produce a report with a recommendation for ratings of each treatment add-on.
- The independent reviewer's report to be circulated before the SCAAC meeting.
- SCAAC to hold an additional meeting specifically to assign the new add-ons ratings.

1.4. Ratings were discussed in more depth for two add-ons: physiological intracytoplasmic sperm injection (PICSI) and time-lapse at the **February 2023** SCAAC meeting. It was

agreed that these ratings would be reassessed at the next SCAAC meeting after the updates to the methodology had been implemented.

1.5. ESHRE are developing a **good practice recommendation paper** which outlines a set of treatment add-on tests and treatments, describes their rationale and any evidence of their efficacy and safety, and provides a recommendation for clinical practice. This is currently a draft paper and the recommendations therein have not been considered by our expert reviewer but may be of interest to the committee.

2. Literature search – updated process

2.1. The interface MEDLINE (Ovid) will be used to carry out the searches, along with two clinical trial registries in line with Cochrane (International Clinical Trials Registry Platform (ICTRP) and ClinicalTrials.gov), as recommended by a medical librarian. Systematic reviews are screened by a specialist biostatistician (independent reviewer) to ensure that all randomised studies are considered in the current assessment.

2.2. The literature is first searched for randomised controlled trials (RCTs) and systematic reviews. If fewer than three RCT studies are identified, then the search will be expanded to non-randomised studies of intervention (NRSIs). NRSIs are limited to case/cohort/control studies (as followed by NICE and agreed by the SCAAC in **February 2023**). Pre-prints will not be considered as part of the evidence base, as they may never get published, and neither are abstracts as they do not provide sufficient information. This is in line with the decision tree found at Annex A that was agreed by the SCAAC in **February 2023**.

2.3. At the **February 2017 SCAAC meeting**, it was agreed that evidence published in the last 10 years would be sent for review. Due to the updated process, all searches for this literature review went back to 2007, in line with the current process of reviewing studies in the last 10 years since 2017. For subsequent reviews the literature will be searched for publications since the last review.

2.4. The Executive will continue to apply this when an add-on is introduced to the HFEA's list, i.e. the last 10 years of evidence will be considered.

2.5. The decision tree for determining how evidence will be used by SCAAC when assigning add-ons rating reflects the agreed process and can be found at Annex A.

3. Independent assessment of the quality of evidence

3.1. In order to categorise the treatment add-ons under consideration, it is necessary not only to identify the published evidence on each treatment add-on, but also to assess the quality of that evidence. For this reason, we seek advice from an expert in systematic reviews and evidence assessment to carry out an independent assessment of the quality of evidence (using the GRADE methodology[1]) for each treatment add-on.

3.2. The independent reviewer reassessed the traffic light ratings in light of the new five-category rating system and additional studies identified using the updated search terms.

[1] GRADE is an approach for grading the quality of evidence and the strength of recommendations. It was developed by the Grading of Recommendations, Assessment, Development and Evaluation Working Group.

3.3. The critical review of studies included assessment of risk of bias from allocation method, blinding, selective reporting, unexplained attrition, unplanned interim analysis and other miscellaneous errors in the design, conduct or reporting of results.

3.4. The findings of the assessment for each add-on and the independent reviewer's recommended ratings can be found as a separate paper, with the current traffic light rating agreed by the committee, last in **October 2021**.

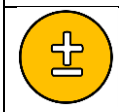
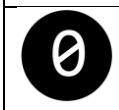
3.5. The assessments made by the independent reviewer are from a methodological perspective without expertise in the clinical or scientific context. The independent reviewer's report can be found as a separate paper.

4. Assigning ratings to add-ons

4.1. The Authority approved a five category rating system with the following symbols/colours in **July 2022** and the SCAAC were updated in **October 2022**.

4.2. At the **February 2023** SCAAC meeting it was agreed that the decision tree terminology at Annex A and the definitions of the ratings would be updated in line with GRADE methodology. The table below has also been updated to reflect these changes (highlighted changes shown in yellow).

4.3.

	On balance, **findings from high quality evidence** shows **this add-on is effective** at improving the treatment outcome**.** An add-on can be rated green if at least one moderate/high quality RCT focuses on LBR.
	On balance, **it is not clear whether this add-on is effective** at improving the treatment outcomes. This is because there **is conflicting moderate/high quality evidence.** In some studies, the add-on has been found to be effective, but in other studies it has not.
	**We cannot rate the effectiveness of this add-on** at improving the treatment outcome as there **is insufficient moderate/high quality evidence.** If an insufficient number of publications can be identified as per the evidence decision tree, the intervention will be rated grey **unless safety concerns have been identified in which case SCAAC may decide to rate the add-on red.**
	On balance, the findings **from moderate/high quality evidence shows that this add-on has no effect on the treatment outcome.**
	**There are potential safety concerns** and/or, on balance, the **findings from moderate/high quality evidence shows** that **this add-on may reduce treatment effectiveness.**

4.4. The five-category rating system was also approved by the Authority to be applied to additional outcomes, such as miscarriage, and outcomes for specific patient groups, such as male-factor infertility, in addition to live births.

5. Recommendations

5.1. The committee is asked to:

- consider the quality of evidence for each treatment add-on based on the findings from an independent assessor;
- agree and recommend ratings for each treatment add-on based on the outcome of live birth rate for the general population; and
- agree and recommend ratings for each additional outcome and population relevant to specific treatment add-ons.

6. Considerations and recommendations for each treatment add-on

Artificial egg activation using calcium ionophore

6.1. Artificial egg activation using calcium ionophore is an authorised process for use only in suitable patients. When the SCAAC considered the use of calcium ionophore for egg activation as an authorised process, they highlighted the theoretical risks relating to embryo viability (eg premature activation and triploid embryos). Given the theoretical risks of using calcium ionophore, centres using it are expected to do so only in selected patients, such as those with Phospholipase C zeta (PLCz) deficiency. Centres are expected to document their rationale for using Calcium Ionophore for individual cases. As with all treatments and processes, centres should ensure that patients are fully informed about the efficacy and potential risks and that validation is carried out.

6.2. Artificial egg activation calcium ionophore was introduced to the HFEA's traffic light rated list of add-ons in February 2017 and was assigned an amber traffic light rating by the Committee. No changes have been made to this traffic light rating since then. However, in the February 2023 review the independent reviewer stated there was 1 moderate/high quality study whereas in this review there were no moderate/high quality studies. The independent reviewer commented 'Caglar Aytac 2015 study was initially assessed as a moderate/high quality study. This was inconsistent with the standard applied regarding reporting of a concealed allocation process (an aspect consistently found to be one of the most serious risks of bias).

6.3. The panel recommended the review for this add-on to include outcomes relating to embryo formation and early development in addition to live birth rate. Outcomes for patients with failed fertilisation in previous ICSI cycles were requested in addition to outcomes for the general population.

- The committee is asked to consider whether the ratings for all outcomes should be GREY given the reviewer's comment 'no moderate/high quality study' was identified.

Current rating	Expert review July 2023 (current)

Live birth rate for most fertility patients

- GREY for live birth rate for most fertility patients
- GREY for embryo formation and early development for most fertility patients
- GREY for live birth rate for patients with failed fertilisation in previous ICSI treatments
- GREY for embryo formation and early development for patients with failed fertilisation in previous ICSI treatments

6.4. This review included 10 RCT's and 7 NRSI's for artificial egg activation using calcium ionophore.

6.5. The Association of Reproductive and Clinical Scientists (ARCS) and British Fertility Society (BFS) are currently working on developing professional guidelines on best practice use of artificial egg activation. Once published, the Executive may bring a discussion back to the SCAAC to consider whether artificial egg activation should remain on our treatment add-ons list given that it is an authorised process (with specific conditions), once published professional guidelines provide best practice recommendations.

Assisted hatching

6.6. Assisted hatching was introduced to the HFEA's traffic light rated list of add-ons in February 2017 and was assigned a red traffic light rating by the Committee. No changes have been made to this traffic light rating since then.

- The panel did not make any recommendations to rate populations or outcomes other than live birth rate for the general population, however the committee is asked to consider whether it is more appropriate to separate out ratings for the use of assisted hatching in frozen oocytes, frozen embryos, fresh embryos and blastocysts in order to provide more useful information for patients.
- If the committee recommend that we should continue to rate this add-on for live birth rate in the general population only then the committee is asked to consider whether the rating should be GREY given the independent reviewer's comment that 'only one moderate/high quality study was identified for live birth rate and no safety concerns were raised across a large number of studies.' The reviewer also noted that 'there are multiple studies of unclear risk of allocation bias that, on the whole, favour assisted hatching in terms of clinical pregnancy rate. Cochrane's review concludes unproven for live birth and contains many older RCTs in addition to those in this review.'

Current rating	Expert review July 2023 (current)
	

Live birth rate for most fertility patients	• GREY for live birth rate for most fertility patients

6.7. This review included 24 RCT's and 3 NRSI's for assisted hatching.

Elective freeze all cycles

6.8. Elective freeze all cycles was introduced to the HFEA's traffic light rated list of add-ons in February 2017 and was assigned an amber traffic light rating by the Committee. No changes have been made to this traffic light rating since then.

6.9. The panel recommended the review for this add-on to include ovarian hyperstimulation syndrome (OHSS) outcomes, obstetric/neonatal outcomes, and time to birth, in addition to live birth rate. Outcomes for patients at increased risk of OHSS were requested in addition to outcomes for the general population.

- The Committee is asked to consider whether OHSS outcomes for the general population should be rated GREEN given the reviewer's comment that 'on balance, there is consistent evidence.'
- The Committee is asked to consider whether obstetric/neonatal outcomes in the general population should be rated GREY given the reviewer's comment that 'studies were underpowered.'
- The Committee is asked to consider whether live birth rate for populations at increased risk of OHSS should be rated GREY given the reviewer's comment 'only 2 moderate/high quality studies' were identified.
- The Committee is asked to consider whether OHSS for populations at increased risk of OHSS should be rated GREEN given the reviewer's comment that there are 'only 2 studies but this is consistent with the general population.' This has changed from the previous recommendation of GREY/GREEN at the February 2023 SCAAC. The independent reviewer also commented that 'Chen 2016 study is comfortably the largest prospective study in this group providing what could be considered definitive evidence on its own. Most statisticians would argue that the best evidence of a treatment effect in a subgroup is given by the estimated effect in the whole population. Here, we have evidence in the subgroup that could arguably stand alone (as with Miller 2019 in PICSI) but it is also consistent with the evidence for this outcome in the population.' The independent reviewer noted they could have recommended GREEN for this outcome in the February 2023 report and that it is hard to argue against this rating now.
- The Committee is asked to consider whether obstetric/neonatal outcomes for populations at increased risk of OHSS should be rated GREY given the reviewer's comment that there is 'only 1 RCT reporting.'

Current rating	Expert review July 2023 (current)
	

Live birth rate for most fertility patients		• AMBER Live birth rate for most fertility patients
		• GREEN for OHSS outcomes for most fertility patients • GREEN for OHSS outcomes for populations at increased risk of OHSS
		• GREY for obstetric/neonatal outcomes for most fertility patients • GREY for obstetric/neonatal outcomes for populations at increased risk of OHSS • GREY for live birth for populations at increased risk of OHSS

6.10. This review included 17 RCT's and 8 NRSI's for elective freeze all cycles.

Endometrial receptivity array (ERA)

6.11. Endometrial receptivity array (ERA) was introduced to the HFEA's traffic light rated list of add-ons in June 2021 and was assigned a red traffic light rating by the Committee. No changes have been made to this traffic light rating since then.

- This previous rating was given due to the independent reviewer's comment that there was only one high quality study identified but there were safety concerns raised by Cozzolino 2022 study. The Committee is asked to consider whether a RED rating should continue to be given for this add-on.

Current rating	Expert review July 2023 (current)
 Live birth rate for most fertility patients	 • RED for live birth rate for most fertility patients

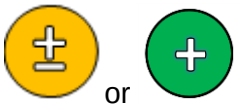
6.12. This review included 2 RCT's and 3 NRSI's for ERA.

Endometrial scratching

6.13. Endometrial scratching was introduced to the HFEA's traffic light rated list of add-ons in February 2017 and was assigned an amber traffic light rating by the Committee. No changes have been made to this traffic light rating since then.

6.14. The panel recommended the review for this add-on to include outcomes for patients with recurrent implantation failure in addition to outcomes for the general population.

- The Committee is asked to consider whether the rating for live birth rate in most fertility patients should be AMBER or GREEN given the independent reviewer's comment that 'the more recent evidence reviewed does not materially affect the previous review but the terminology of the grading has changed. There is consistent high quality evidence of an effect size ranging from zero to a few percentage points. That is, excluding detriment. Meta-analysis is inconclusive at the standard 95% confidence level but "on balance" there is evidence for a small beneficial effect in terms of live birth.

Current rating	Expert review July 2023 (current)
Live birth rate for most fertility patients	 or  • AMBER/GREEN for live birth rate for most fertility patients. • GREY for live birth rate for patients with recurrent implantation failure

6.15. This review included 41 RCT's and 6 NRSI's for endometrial scratching.

Hyaluronate enriched medium (eg EmbryoGlue)

6.16. Hyaluronate enriched medium was introduced to the HFEA's traffic light rated list of add-ons in February 2017 and was assigned an amber traffic light rating by the Committee. No changes have been made to this traffic light rating since then.

- The Committee are asked to consider whether the rating for live birth rate in the general population should be rated GREEN given the independent reviewer's comment that there are 'at least three moderate/high quality studies with broadly consistent results.'

Current rating	Expert review July 2023 (current)
Live birth rate for most fertility patients	 • GREEN for live birth rate for most fertility patients

6.17. This review included 10 RCT's for hyaluronate enriched medium,

Intracytoplasmic morphologic sperm injection (IMSI)

6.18. IMSI was introduced to the HFEA's traffic light rated list of add-ons in October 2018 and was assigned a red traffic light rating by the Committee. No changes have been made to this traffic light rating since then.

6.19. The panel recommended the review for this add-on to include outcomes for patients with male-factor infertility in addition to outcomes for the general population.

- The Committee are asked to consider whether the rating for live birth rate in the general population should be rated GREY given the independent reviewer's comment that 'only one moderate/high quality study was identified with no safety concerns.'
- The Committee is asked to consider whether the rating for male factor in the general population should be GREY given the independent reviewer's comment that 'only one moderate/high quality study was identified with no safety concerns.'

Current rating	Expert review July 2023 (current)
 Live birth rate for most fertility patients	 • GREY for live birth rate for most fertility patients • GREY for live birth rate for male-factor infertility patients

6.20. This review included 15 RCT's and 4 NRSI's for IMSI.

Intrauterine culture

6.21. Intrauterine culture was introduced to the HFEA's traffic light rated list of add-ons in February 2017 and was assigned a red traffic light rating by the Committee. Only one published study has been identified for this add-on and no safety concerns have been raised. This results in a grey rating as per the evidence decision tree as there is an insufficient number of publications.

- The Committee are asked to consider whether the rating for live birth rate in the general population should be rated GREY given the independent reviewer's comment that 'no further studies have been identified since the last review in October 2021'

Current rating	Expert review July 2023 (current)
 Live birth rate for most fertility patients	 • GREY for live birth rate for most fertility patients

6.22. This review included 1 NRSI for intrauterine culture.

Physiological intracytoplasmic sperm injection (PICSI)

6.23. PICSI was introduced to the HFEA's traffic light rated list of add-ons as in October 2018 and was assigned a red traffic light rating by the Committee. No changes have been made to this traffic light rating since then.

6.24. The panel recommended the review for this add-on to include miscarriage rate, in addition to live birth rate. Outcomes for patients with male-factor infertility and older women were requested in addition to outcomes for the general population.

- The Committee is asked to consider whether live birth rate and miscarriage for the general population should be rated as GREY given the independent reviewer's comment that 'only one moderate/high quality study with no safety concerns was identified.' The reviewer also noted that 'as for endometrial scratch there is high quality evidence that any effect on live birth is no more than a few percentage points. Here this is based on a single definitive study but it would seem implausible that future studies will be funded sufficient to materially affect the conclusion that PICSI leads to fewer miscarriages and similar live birth rate. A randomised trial with 90% power to detect a difference in live birth rates between 25% and 27% would require in excess of 20,000 participants. The Committee could consider whether GREEN for miscarriage and BLACK for live birth may be more informative summary information for patients despite not strictly fitting the current definitions for these grades.'
- The Committee is asked to consider whether miscarriage and live birth rate for male factor infertility populations should be rated GREY given the independent reviewer's comment that 'only one moderate/high quality study with no safety concerns has been identified.' The independent reviewer also noted that 'Miller 2019 comprised 95% of participants with male factor. See comment above for general population.'
- The Committee is asked to consider whether live birth rate and miscarriage in older women should be rated GREY given the independent reviewer's comment that 'Only one moderate/high quality study with no safety concerns' was identified.' The independent reviewer noted that 'given consistency with the general population, the Committee could consider grading GREEN for miscarriage. The potential effect on live birth is less clear given the larger estimate and much wider confidence intervals.'

Current rating	Expert review July 2023 (current)
 Live birth rate for most fertility patients	 • GREY for all outcomes for the general population. The Committee could consider whether GREEN for miscarriage and BLACK for live birth may be more informative. • GREY for all outcomes for patients with male factor infertility

- GREY for all outcomes for older women. Given consistency with general population, the committee could consider grading GREEN for miscarriage.

6.25. This review included 13 RCT's and 4 NRSI's for PICSI.

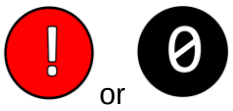
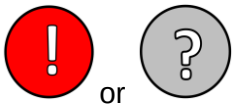
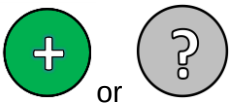
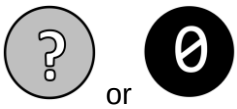
Pre-implantation genetic testing for aneuploidy (PGT-A)

6.26. PGT-A for day five embryos was introduced to the HFEA's traffic light rated list of add-ons in February 2017 and was assigned an amber traffic light rating by the Committee, this rating was changed to a red traffic light by the Committee in October 2019.

6.27. The panel recommended the review for this add-on to include miscarriage rate and time to birth, in addition to live birth rate. Outcomes for older women were requested in addition to outcomes for the general population.

- The Committee is asked to consider whether live birth rate for the general population should be rated as RED/BLACK given the independent reviewer's comment that the Yan 2021 study looks definitive that the add-on reduces treatment effectiveness, but there is only one study of this so it could be argued another way.
- The Committee is asked to consider whether the time to birth for the general population should be rated as RED/GREY given the same comment above regarding the Yan 2021 study.
- The Committee is asked to consider whether the miscarriage rate for the general population should be rated as GREEN given the independent reviewers comment that 'several moderate/high quality studies, with consistent findings were identified.'
- The Committee is asked to consider whether the miscarriage rate for older women should be rated as GREEN/GREY given the independent reviewer's comment that although there is only one high quality study available, it is consistent with studies in the general population. This has changed from the recommendation GREY made at the February 2023 meeting. The independent reviewer commented that in the Verpoest 2018 study 'PGT-A was by polar body biopsy six to nine hours after ICSI, which may constitute a different class of intervention.'. The committee should consider whether this study should be included, in which case the rating would be GREEN. The evidence is not 'stand-alone' within the subgroup but is consistent with the effect within the full population. If the committee consider this a different intervention, then there is no moderate/high quality evidence specific to the subgroup and it may be preferable to rate as GREY.
- The Committee is asked to consider whether live birth in older women should be rated as BLACK/GREY given the independent reviewer's comment regarding the one eligible study above.
- The Committee is asked to consider whether time to birth in older women should be rated as RED/GREY given the independent reviewer's comment regarding the one eligible study above. This recommendation has changed from

GREY in the February 2023 review. The independent reviewer commented that this is the same issue regarding Verpoest 2018 study above and it is consistent with the effect in the general population – 'If ineligible there is no evidence specific to the subgroup. If eligible, the committee may prefer consistency with the population rating.' In this case, the committee should decide whether to rate RED/GREY depending on whether they feel that Yan 2021 study should be considered definitive.

Current rating	Expert review July 2023 (current)
 Live birth rate for most fertility patients	     • RED/BLACK for live birth rate for most fertility patients • RED/GREY for time to birth for most fertility patients • RED/GREY for time to birth for older women • GREEN for miscarriage for most fertility patients • GREEN/GREY for miscarriage in older women • GREY/BLACK for live birth rate for older women

6.28. This review included 18 RCT's and 3 NRSI's for PGT-A.

Immunological tests and treatments for infertility

6.29. Immunological tests and treatments for infertility was introduced to the HFEA's traffic light rated list of add-ons as an umbrella term covering all immunological tests and treatments for infertility treatments in February 2017 and was assigned a red traffic light rating by the Committee. No changes have been made to this traffic light rating since then.

6.30. At the **October 2020** SCAAC meeting it was proposed that immunological tests and treatments for infertility be broken down by treatment type and an individual traffic light rating be allocated to each type.

6.31. The independent reviewer did not make an overarching recommendation for immunological tests and treatments for infertility. The committee is asked to consider the independent reviewer's recommendations on intralipids, intravenous immunoglobulin and steroids (glucocorticoids), and recommend an overall rating for the group if appropriate.

6.32. The panel recommended the review for this add-on to include miscarriage rate in addition to live birth rate. Outcomes for patients undergoing immunological testing, such as natural killer cell blood tests, were requested in addition to outcomes for the general population.

- For intralipids the Committee is asked to consider whether live birth rate and miscarriage rate in the general population should be rated as GREY/RED given the independent reviewer's comment that 'No moderate/high quality studies' were identified. The reviewer noted that 'there is a question over whether the Committee considers the safety concerns raised over congenital malformations justify the red rating.'
- For intralipids the Committee is asked to consider whether live birth rate and miscarriage rate in populations with immunological testing should be rated as GREY/RED for all outcomes given the independent reviewer's comment that 'no moderate/high quality studies and no safety concerns specific to this sub-population were identified but the safety concern raised above regarding congenital malformations may need to be considered.'
- For IV immunoglobulin the Committee is asked to consider whether live birth rate and miscarriage rate in the general population should be rated as AMBER given the independent reviewer's comment that '3 RCTs providing moderate quality evidence' were identified. The reviewer also noted that although 'not conflicting, the results were too imprecise to determine effectiveness at this stage.'
- For IV immunoglobulins the Committee is asked to consider whether live birth rate and miscarriage rate in populations with immunological testing should be rated as GREY given the independent reviewer's rating that 'no moderate/high quality studies and no safety concerns' were identified.
- For steroids (glucocorticoids) the Committee is asked to consider whether live birth rate and miscarriage rate in the general population should be rated as GREY given the independent reviewer's comment that there is 'insufficient evidence from moderate/high quality studies, and no safety concerns' identified.
- For steroids (glucocorticoids) the Committee is asked to consider whether live birth rate and miscarriage rate in populations with immunological testing should be rated as GREY given the independent reviewer's comment as above for the general population.

Current rating	**Expert review July 2023 (current)**

Live birth rate for most fertility patients for all immunological tests and treatments	The exert reviewer was not asked to recommend a rating for live birth rate for most fertility patients for all immunological tests and treatments.
Intralipids	 or  • GREY/RED for live birth rate for most fertility patients • GREY/RED for miscarriage rate for most fertility patients • GREY/RED for live birth rate for populations with immunological testing • GREY/RED for miscarriage rate for populations with immunological testing
Intravenous immunoglobulin	 • AMBER for live birth rate for most fertility patients • AMBER for miscarriage rate for most fertility patients  • GREY for live birth rate for populations with immunological testing • GREY for miscarriage rate for populations with immunological testing
Steroids (glucocorticoids)	 • GREY for live birth rate for most fertility patients • GREY for miscarriage rate for most fertility patients • GREY for live birth rate for populations with immunological testing • GREY for miscarriage rate for populations with immunological testing

6.33. This review included 4 RCT's and 1 NRSI for intralipids.

6.34. This review included 5 RCT's and 3 NRSI's for intravenous immunoglobulin.

6.35. This review included 10 RCT's and 6 NRSI's for steroids (glucocorticoids).

Time-lapse imaging and incubation

6.36. Time-lapse incubation and imaging was introduced to the HFEA's traffic light rated list of add-ons in February 2017 and was assigned an amber traffic light rating by the Committee. No changes have been made to this traffic light rating since then.

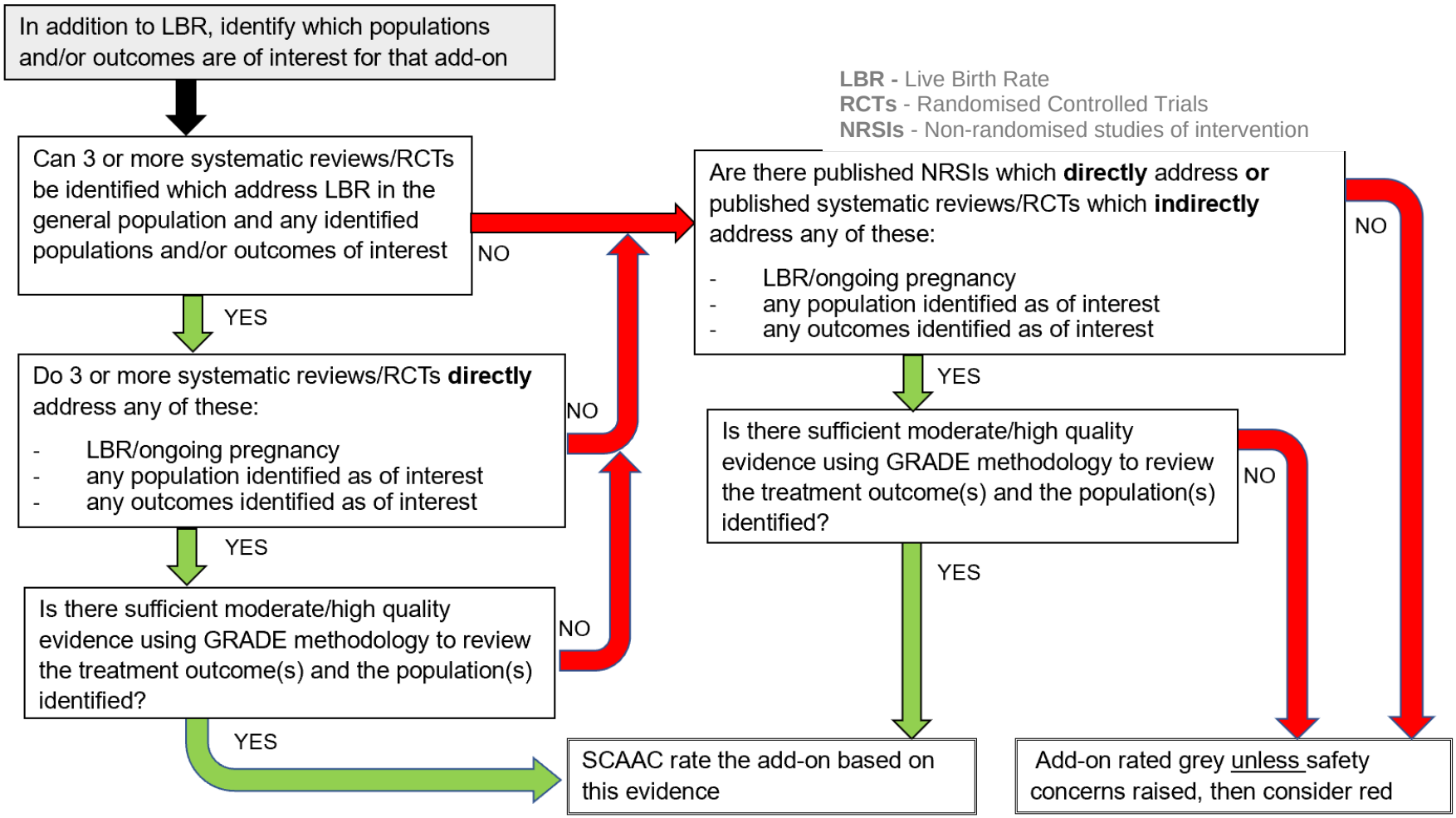
6.37. The panel recommended the review for this add-on to report any differences in outcomes for the use of manual annotation of time lapse images by an embryologist vs automated annotation of time lapse images using a computer software or artificial intelligence.

- The Committee is asked to consider whether live birth rate for the general population should be rated as BLACK for automated and manual annotation, given the independent reviewer's comment that '5 moderate/high quality studies with consistent results' were identified.

Current rating	Expert review July 2023 (current)
Live birth rate for most fertility patients	• BLACK for live birth rate for most fertility patients when using automated annotation • BLACK for live birth rate for most fertility patients when using manual annotation

6.38. This review included 17 RCT's and 2 NRSI's for time-lapse incubation and imaging.

Annex A. Evidence decision tree for rating add-ons



[1]GRADE is an approach for grading the quality of evidence and the strength of recommendations. It was developed by the Grading of Recommendations, Assessment, Development and Evaluation Working Group.

Annex B. References of reviewed studies

Bold indicates studies added for the July 2023 update.

Adjunct	Study	DOI/reference
Artificial Egg Activation		
General	**Nasr-Esfahani 2007**	**10.1016/j.fertnstert.2007.10.047**
	Borges 2009	**10.1016/j.fertnstert.2008.04.046**
	Liu 2011	**10.1017/S0967199411000530**
	Ebner 2012	**10.1016/j.fertnstert.2012.07.1134**
	Liu 2014	**10.1089/cell.2013.0081**
	Eftiikhar 2013	**IRCT2012112610328N1**
	Caglar Aytac 2015	10.1016/j.fertnstert.2015.07.1163
	Fawzy 2018	10.1093/humrep/dey258
	Shebl 2021	**10.1007/s10815-021-02338-3**
	Yin 2022	**10.1007/s00404-021-06329-8**
Failed fertilisation	Meerschaut 2012	10.1093/humrep/des097
	Montag 2012	**10.1016/j.rbmo.2012.02.002**
	Ebner 2015	**10.1016/j.rbmo.2014.11.012**
	Darwish 2015	**10.1016/j.rbmo.2015.08.012**
	Aydinuraz 2016	10.1080/14647273.2016.1240374
	Hao 2016	**10.3760/cma.j.issn.0376-2491.2016.43.010**
	Li 2019	**10.1016/j.rbmo.2019.03.216**
Assisted Hatching: Stored	Balaban 2006	10.1093/humrep/del097
	Ge 2008froz	RBMO 2008;16(4):589-96.
	Valojerdi 2010	10.1016/j.rbmo.2009.11.002
	Fang 2010	10.1016/j.fertnstert.2009.08.014
	Debrock 2011	**10.1093/humrep/der161**
	Figueria 2012	**10.1016/j.ejogrb.2012.05.022**
	Ren 2013	**10.1007/s10815-013-9984-2**
	Wan 2014	**10.1016/j.rbmo.2014.01.006**
	Wang 2016	10.3892/br.2016.716
	Knudtson 2016	F&S 2016;106(3) Suppl:e141
	Safari 2017	10.1016/j.repbio.2017.05.003
	Elnahas 2017	**10.1016/j.mefs.2017.05.006**
	Kirienko 2019	**10.1016/j.rbmo.2019.06.003**
Fresh	Sagoskin 2007	10.1016/j.fertnstert.2006.07.1498
	Ge 2008fresh	RBMO 2008;16(4):589-96.
	Balakier 2009	10.1016/j.fertnstert.2008.07.1729
	Hagemann 2010	10.1016/j.fertnstert.2009.01.116
	Kutlu 2010young	10.1007/s10815-010-9431-6
	Kutlu 2010old	10.1007/s10815-010-9431-6
	Razi 2013	Iran J reprod Med 2013;11(12):1021-6.
	González-Ortega 2015	**Ginecol Obstet Mex 2015;83:670-9.**
	Shi 2016	10.1177/1933719116641764
	Chang 2016	F&S 2016;106(3) Suppl:e314
	Nada 2018	10.1007/s00404-017-4604-5

	Abulsoud 2019	**Int J Pharm Rev Res 2019;56(1):112-6**
	Fawzy 2020	**10.1093/humrep/deaa160**
	Zhang 2022	**10.3389/fendo.2022.927834**
Embryo Glue	Morbeck 2007	NCT005882250
	Mahani 2007	EMHJ 2007;13(4):876-80.
	Friedler 2007	10.1093/humrep/dem220
	Korosec 2007	RBM0 2007;15(6):701-7.
	Hazlett 2008	10.1016/j.fertnstert.2007.05.063
	Urman 2008	10.1016/j.fertnstert.2007.07.1294
	Fancsovits 2015	10.1007/s00404-014-3541-9
	Singh 2015	10.4103/0974-1208.170398
	Kleijkers 2016	**10.1093/humrep/dew156**
	Yung 2021	**10.1016/j.fertnstert.2021.02.015**
Endometrial Receptivity	Simón 2020	10.1016/j.rbmo.2020.06.002
	Cohen 2020	**10.1080/19396368.2020.1824032**
	Cozzolino 2020	**10.1007/s10815-020-01948-7**
	Cozzolino 2022	**10.1016/j.fertnstert.2022.07.007**
	Doyle 2022	**10.1001/jama.2022.20438**
Endometrial Scratching	Raziel 2007	10.1016/j.fertnstert.2006.05.062
General	**Zhou 2008**	**10.1016/j.fertnstert.2007.05.064**
	Karimzadeh 2009	10.1111/j.1479-828X.2009.01076
	Narvekar 2010	10.4103/0974-1208.63116
	Abdelhamid 2012	10.1007/s00404-013-2785-0
	Nastri2013	**10.1002/uog.12539**
	Gibreel 2013	10.1111/j.1447-0756.2012.02016.x
	Parsanezhad 2013	IRCT:2012082510657NI
	Zarei 2014	IRCT:2012070810210NI
	Wadhwa 2015	J Hum Reprod Sci 2015;8(3):151-8.
	El Khayat 2015	10.1016/j/ejogrb.2015.08.025
	Gibreel 2015	**10.3109/09513590.2014.994603**
	Maged 2016	10.1177/1933719115602776
	Bahaa Eldin 2016	**10.1177/1933719116638191**
	Goel 2017	10.1007/s10815-017-0949-8
	Mak 2017	10.1016/j.rbmo.2017.04.004
	Aleyamma 2017	10.1016/j.ejogrb.2017.05.005
	Helmy 2017	10.1002/ijgo.12178
	Senocak 2017	10.1016/j.jogoh.2017.09.003
	Ashrafi 2017	10.1111/jog.13401
	Maged 2018	10.1002/ijgo.12355
	Hilton 2019	**10.1007/s00404-019-05044-9**
	Eskew 2019	**10.1007/s10815-018-1356-5**
	Frantz 2019	10.1093/humrep/dey334
	Lensen 2019	10.1056/NEJMoa1808737
	Olesen 2019	10.1016/j.fertnstert.2019.08.010
	Mackens 2020	**10.1093/humrep/deaa018**
	Tang 2020	**10.1111/jog.14193**
	Berntsen 2020	10.1016/j.ejogrb.2020.06.034
	Ghuman 2020	10.1016/j.ejogrb.2020.08.010

	Rodriguez 2020	10.1007/s43032-020-00204-8
	van Hoogenhuijze 2021	10.1093/humrep/deaa268
	Metwally 2021	10.1093/humrep/deab041
	Yavangi 2021	10.18502/ijrm.v19i5.9255
	Glanville 2022	**10.1016/j.rbmo.2021.10.008**
	Izquierdo 2022	**10.1016/j.jogoh.2022.102335**
	Madhuri 2022	**10.1016/j.ejogrb.2021.10.028**
	Metwally 2022	**10.3310/JNZT9406**
	Wong 2022	**10.1016/j.fertnstert.2021.12.009**
Implantation Failure	**Baum 2012**	**10.3109/09513590.2011.650750**
	Zhang 2014	**10.1007/s00404-014-3382-6**
	Zhang 2015	**10.1007/s11655-014-1843-1**
	Bord 2015	**10.1007/s00404-015-3954-0**
	Siristatidis 2017	**10.1080/09513590.2016.1255325**
	Gürgan 2019	**10.1016/j.rbmo.2019.02.014**
	Tumanyan 2019	**10.1080/09513590.2019.1632085**
	Aghajanpour 2021	**10.1016/j.jri.2021.103426**
Freeze All: General	Aflatoonian 2010	10.1007/s10815-010-9412-9
	Shapiro 2011a	10.1016/j.fertnstert.2011.05.050
	Shapiro 2011b	10.1016/j.fertnstert.2011.02.059
	Shapiro 2015	**10.1016/j.fertnstert.2015.07.1141**
	Magdi 2017	10.1016/j.fertnstert.2017.04.020
	Shi 2018	10.1056/NEJMoa1705334
	Le 2018	**10.1093/humrep/dey253**
	Vuong 2018	10.1056/NEJMoa1703768
	Vuong 2019	**10.1016/j.rbmo.2018.12.012**
	Wei 2019	10.1016/S0140-6736(18)32843-5
	Stormlund 2020	10.1136/bmj.m2519
	Simón 2020	10.1016/j.rbmo.2020.06.002
	Boynukalin 2020	**10.1371/journal.pone.0234481**
	Li 2021	**10.3389/fendo.2021.730059**
	Wong 2021	10.1093/humrep/deaa305
	Maheshwari 2022	**10.1093/humrep/deab279**
	Maheshwari 2022a	**10.3310/AEFU1104**
OHSS risk	**Chen 2016**	**10.1056/NEJMoa1513873**
	Ye 2018	**10.1186/s12958-018-0373-7**
	Deng 2019	**10.1007/s11596-019-2031-5**
	Shrem 2019	**10.1016/j.rbmo.2019.04.014**
	Santos-Ribeiro 2020	10.1093/humrep/deaa226
	Deepika 2021	**10.5935/1518-0557.20200028**
	Huang 2021	**10.1038/s41598-021-02227-w**
	Vuong 2021	**10.1007/s10815-021-02180-7**
IMSI: General	**Balaban 2011**	**10.1016/j.rbmo.2010.11.003**
	Figueira 2011	**10.1016/j.fertnstert.2010.11.018**
	Setti 2012	**10.1016/j.rbmo.2012.01.007**
	Setti 2013	10.1016/j.ejogrb.2013.09.006
	Marci 2013	10.1186/1742-4755-10-16
	Cassuto 2014	10.1016/j.rbmo.2013.08.013

Male Factor	**Setti 2014**	**10.1016/j.ejogrb.2014.10.008**
	Antinori 2008	**10.1016/s1472-6483(10)60150-2**
	Knez 2011	**10.1186/1477-7827-9-123**
	Setti 2011	**10.1016/j.fertnstert.2011.03.003**
	Knez 2012	**10.1016/j.rbmo.2012.03.011**
	Check 2013	**Clin Exp Obstet Gyn 2013;40(1):15-7.**
	De Vos 2013	10.1093/humrep/des435
	Leandri 2013	10.1111/j.2047-2927.2013.00104.x
	Kim 2014	10.5653/cerm.2014.41.1.9
	Sifer 2014	**10.1016/j.ejogrb.2014.07.017**
	La Sala 2015	10.1186/s12958-015-0096-y
	Mangoli 2019	10.1111/and.13340
	Mangoli 2020	**10.1007/s10815-020-01910-7**
Immunological testing	Singh 2019	10.1016/j.ejogrb.2019.06.007
	Al-Zebeidi 2019	10.1080/09513590.2019.1631280
	Dakhly 2016	10.1016/j.ijgo.2016.06.026
	Meng 2016	**10.1007/s00404-015-3922-8**
	Rogenhofer 2021	**10.1111/aji.13506**
Intrauterine culture	Blockeel 2009	10.1093/humrep/dep005
IV Immunoglobulin		
General	Stephensen 2010	10.1093/humrep/deq179
	Christiansen 2014	10.1111/1471-0528.13192
	Jørgensen 2020	**10.1016/j.jri.2020.103128**
Immunological testing	**Dendrinos 2009**	**10.1016/j.ijgo.2008.11.010**
	Yamada 2015	**10.1016/j.jri.2015.01.008**
	Lee 2016	**10.1111/aji.12442**
	Meng 2016	**10.1007/s00404-015-3922-8**
	Ahmadi 2017	**10.1016/j.imlet.2017.10.003**
MACS	Romany 2014	10.1016/j.fertnstert.2014.09.001
	Troya 2015	10.5935/1518-0557.20150015
	Romany 2017	10.1007/s10815-016-0838-6
	Ziarati 2018	10.1080/14647273.2018.1424354
PGT-A (Day 3)	Mastenbroek 2007	NEJM 2007;357:9-17.
	Hardarson 2008	10.1093/humrep/den217
	Staessen 2008	10.1093/humrep/den367
	Blockeel 2008	RBMO 2008;17(6):848-54.
	Meyer 2009	10.1016/j.fertnstert.2008.02.162
	Schoolcraft 2009	10.1016/j.fertnstert.2008.05.029
	Sher 2009	10.1016/j.fertnstert.2008.11.029
	Debrock 2010	10.1016/jfertnstert.2008.10.072
	Ikuma 2015	10.1371/journal.pone.0129958
	Rubio 2017	10.1016/j.fertnstert.2017.03.011
PGT-A (Day 5) General	Yang 2012	Molec Cytogen 2012;5:24
	Forman 2013	10.1016/j.fertnstert.2013.02.056
	Scott 2013	10.1016/j.fertnstert.2013.04.035
	Ozgur 2019	10.1007/s10815-018-01399-1
	Munné 2019	10.1016/j.fertnstert.2019.07.1346
	Cimadomo 2019	**10.1093/humrep/dez078**
	Yan 2021	**10.1056/NEJMoa2103613**

	De Munck 2022	**10.1371/journal.pone.0267241**
	Idárraga 2022	**10.5935/1518-0557.20210085**
Older women	**Ubaldi 2017**	**10.1016/j.fertnstert.2017.03.007**
	Verpoest 2018	**10.1093/humrep/dey262**
PICSI	**Parmegiani 2012**	**10.1016/j.fertnstert.2012.05.043**
	Worrilow 2013	10.1093/humrep/des417
	Majumdar 2013	10.1007/s10815-013-0108-9
	Mokanszki 2014	10.3109/19396368.2014.948102
	Troya 2015	10.5935/1518-0557.20150015
	Lohinova 2017	PMID: 29099693
	Erberelli 2017	10.5935/1518-0557.20170002
	Korosi 2017	**PMID: 28724183**
	Avalos-Durán 2018	**10.5935/1518-0557.20180027**
	Miller 2019	10.1016/S0140-6736(18)32989-1
	Hasanen 2020	**10.1007/s10815-020-01913-4**
	Novoselsky 2021	**10.1111/andr.12982**
	Hozyen 2022	**10.1007/s43032-021-00642-y**
Steroids (DHEA)	Wiser 2010	10.1093/humrep/deq220
	Kara 2014	10.1016/j.ejogrb.2013.11.008
	Yeung 2014	10.1016/j.fertnstert.2014.03.044
	Tartagni 2015a	PMID: 24867068
	Tartagni 2015	10.1186/s12958-015-0014-3
	Narkwichean 2017	10.1016/j.ejogrb.2017.09.006
	Wang 2022	10.1111/1471-0528.17045
Steroids (Glucocorticoids)	**Fawzy 2008**	**10.1007/s00404-007-0527-x**
General	Fawzy 2013	10.1007/s00404-013-3020-8
	Gomaa 2014	**10.1007/s00404-014-3262-0**
	Taiyeb 2017	10.1007/s12020-017-1446-7
	Yeganeh 2017	10.1080/01443615.2017.1346593
	Kaye 2017	**10.1016/j.fertnstert.2017.04.003**
	Milardi 2017	**10.1111/andr.12300**
	Siristatidis 2018	**10.1080/09513590.2017.1380182**
	Liu 2018	10.1111/cen.13824
	Thalluri 2022	**10.1093/humrep/deac142**
Immunological testing	**Turi 2010**	**10.1016/j.clinthera.2011.01.010**
	Tang 2013	**10.1093/humrep/det117**
	Fan 2016	10.1111/aji.12559
	Huang 2021	**10.1016/j.jri.2020.103245**
	Gao 2021	**10.1177/09612033211055816**
	Zhou 2022	**10.1186/s12884-022-04532-2**
Time Lapse (i)	Kirkegaard 2012	10.1007/s10815-012-9750-x
	Van Blerkom 2014	**10.1016/j.rbmo.2013.11.012**
	Park 2015	**10.1093/humrep/deu316**
	Wu 2016	**10.1186/s12958-016-0181-x**
	Barberet 2018	**10.1016/j.fertnstert.2017.10.008**
	Chen 2020	10.1093/humrep/deaa268
	Guo 2022	**10.3389/fphys.2021.794601**
(ii)	Goodman 2016	10.1016/j.fertnstert.2015.10.013

	Kaser 2017	**10.1093/humrep/dex231**
	Alhelou 2018	10.1016/j.repbio.2017.12.003
	Kovacs 2019	10.1016/j.ejogrb.2018.12.011
	Ahlstrom 2022	**10.1093/humrep/deac020**
(iii)	Kahraman 2013	10.1177/205891581200300204
	Rubio 2014	10.1016/j.fertnstert.2014.07.738
	Insua 2017	10.1016/j.fertnstert.2017.06.031
	Yang 2018	10.1093/humrep/dey047
	Meng 2022	**10.1016/j.fertnstert.2022.02.015**
	Zhang 2022	**10.1016/j.rbmo.2022.06.017**
	Guo 2022	**10.3389/fphys.2021.794601**

Traffic Light System for Treatment Add-ons

Andy Vail, June 2023

INTRODUCTION

The HFEA website provides patients with digestible information on treatment add-ons in the form of a rating system. The purpose of this report is to inform the Scientific and Clinical Advances Advisory Committee's (SCAAC) deliberations on updating this information. In particular, this update extends the ratings system to five categories, supplements sparse evidence from randomised trials with additional data and considers outcomes other than live birth. Further development of the literature search procedure has also identified some earlier studies previous not reviewed.

The aim of the work reported below is to critically appraise, interpret and summarise, for consideration by the HFEA, the reports of identified studies. Given the time constraints to report in time for the July meeting of SCAAAC, some prioritisation has been necessary. Where three or more randomised trials of at least moderate quality have reported clinical results, or sufficient randomised studies have reported clinical results to categorise the evidence as moderate or high quality under GRADE criteria, there is no further consideration of non-randomised evidence.

METHOD

Dina Halai, Scientific Policy Manager, provided references and hyperlinks to identified studies for consideration, categorised by add-on, study design and population under study. The earliest, newly incorporated papers were published in 2007. I screened systematic reviews to ensure that all randomised studies were considered in the current assessment.

Critical review of studies included assessment of risk of bias from allocation method, blinding, selective reporting, unexplained attrition, unplanned interim analysis and other miscellaneous errors in the design, conduct or reporting of results. To classify a randomised trial as providing moderate/high quality evidence I have applied the default classification of the Cochrane Gynaecology and Fertility review group. Specifically, for a study to be considered in this category it must describe an adequately concealed randomisation process to prevent selection bias. It must also not be identified as at high risk of bias in other regards ('unclear' is acceptable) other than where blinding is unrealistic. Where HFEA specifically requested results for a sub-population of interest, I have presented first the studies addressing the general population and then studies addressing the specific sub-populations. The extent to which interpretation of sparse results for a sub-population should borrow from the broader information available is addressed on a case-by-case basis.

To calculate odds ratios, published results were re-calculated applying the intention to treat (ITT) principle and using two-sided confidence intervals. As these were being interpreted as indicative rather than inferential, no technical adjustments were applied for multiple testing, covariate adjustment or planned interim analyses. Odds ratios were calculated for the latest clinical outcome presented. That is, live birth rate was first choice, followed by ongoing, clinical, unspecified or biochemical pregnancy. An odds ratio greater than 1.0 for these outcomes implies benefit of the add-on under study. Additional outcomes, particularly those relating to safety such as OHSS incidence and miscarriage, are reported where these are a particular aim of the add-on or have been requested by HFEA. An odds ratio greater than 1.0 for adverse outcomes implies detriment of the add-on under study.

RESULTS

1. Artificial egg activation

The previous review in 2019 included four studies: two within-patient designs on sibling oocytes and two RCTs of patients that each suggested promise but studied quite different populations and were dogged by methodological issues. The current search identified a total of 37 primary research studies and four systematic reviews. Searching of these reviews identified one further randomised study for consideration.

Several studies randomised oocytes from each retrieval cycle to be subjected to artificial activation or not. I have referred to these as 'within-patient' studies. By design, these studies do not contribute to the information for clinical outcomes. In principle, they can provide valuable information on pre-transfer, developmental outcomes. For statistically valid inference, investigators need to have accounted for both the 'clustering' of multiple oocytes allocated to each intervention arm from each retrieval cycle and the 'pairing' of the two clusters from each retrieval cycle. In practice, such analyses were rare.

1 (i) *General population*

Nasr-Esfahani 2007 reported randomly assigning oocytes within retrieval cycles of 87 couples with severe male factor (teratozoospermia) and at least four mature oocytes. Oocytes were cultured in G1 medium with or without activation using 10µM ionomycin for 10 minutes. Assessment of fertilisation and embryo scoring were conducted blind to treatment allocation. There was no description of the randomisation process and the study design did not allow for comparison of clinical outcomes. Analysis of development outcomes failed to account for the paired design. Mean fertilisation rates were reportedly higher with activation, and the mean percentage of embryos considered high quality was similar.

Borges 2009 randomised 204 couples with severe male factor(azoospermia) to culture in G1-V3-Plus medium with or without activation using 5µM calcium ionophore A23187 for 30 mins. There was no description of the randomisation process and no discussion of blinding. Up to four embryos were transferred for each participant. The design was stratified by the method of sperm extraction into three groups: those with obstructive azoospermia undergoing percutaneous aspiration (PESA) or testicular aspiration (TESA) and those with a non-obstructive diagnosis undergoing TESA. Clinical pregnancy was reported only as a percentage. For those undergoing TESA it was possible to recalculate the number of participants achieving clinical pregnancy: OR=1.2 (0.51 to 3.0). Reported fertilisation rates were similar overall. The reported percentages of high quality embryos were similar for patients with non-obstructive azoospermia, whereas for those with an obstructive diagnosis rates were lower with activation in the TESA stratum and higher in the PESA stratum. It is unclear how these rates were calculated and analysed.

Liu 2011 conducted a non-clinical study with what appears to have been considered 'waste product'. From previous ICSI cycles they took oocytes that had failed to develop (germinal vesicle or metaphase I). These had been vitrified, thawed and then matured for 24-36 hours, with 204 oocytes maturing to be subject to ICSI using donor sperm. They then describe randomly assigning these to either standard cleavage medium or activation for 6 minutes in 7% ethanol prior to standard cleavage medium. There is no detail to assess the allocation but it appears to have been done regardless of sibling status. The number of women who provided the oocytes is not reported and there appears to have been no intention to transfer any resulting embryos. Reported fertilisation

rates were similar between groups. Only the activated oocytes led to any high quality embryos (n=12 from 104), or high quality blastocysts (n=4).

Ebner 2012 prospectively recruited 66 couples undergoing ICSI with severe male factor and "sufficient" number of oocytes. All were treated with calcium ionophore immediately following ICSI. Unfortunately, there are several methodological issues with this study that preclude statistical interpretation. The presentation and analysis do not account for the multiple cycles per participant. Comparison is made with multiple historic cycles of the same participants. Comparison also fails to account for the inherent matching and is almost guaranteed to show 'benefit' given that regression to the mean, Hawthorne and placebo biases all favour the intervention. The authors reported higher fertilisation and blastocyst formation in the studied cycles, with 26 (39%) participants achieving live birth.

Eftekhar 2013 randomised 38 couples with male factor (teratozoospermia) to culture in GIVF-Plus medium with or without activation using 5µM calcium ionophore A23187 for five minutes. There was no description of a concealment process and explicitly no blinding. All participants received the allocated intervention and analysis was presented by intention to treat. Ongoing pregnancy was higher in the activated arm: OR=2.5 (0.51 to 12) under a policy of transferring up to three embryos. Participants in the intervention arm had an average of 1.8 (0 to 3.5) more fertilised oocytes and 1.8 (-0 to 3.7) more embryos than controls. It should be noted however that they also had an average of 1.5 more oocytes retrieved prior to intervention differences.

Liu 2014 conducted a non-clinical study with what appears to have been considered 'waste product'. From 102 previous IVF and ICSI cycles they took 179 metaphase II oocytes with normal morphology that had failed to fertilise. They then describe randomly assigning these to activation with either 5µM calcium ionophore A23187 for five minutes or strontium chloride for 20 minutes. There is no detail to assess the allocation but it appears to have been done regardless of sibling status. Presentation and analysis also ignored all aspects of clustering. Activation was reported to be higher in the calcium ionophore arm (104 versus 66 oocytes) as was blastocyst formation (four versus one).

Caglar Aytac 2015 randomised 296 couples with diminished ovarian reserve but normal sperm parameters and no previous fertilisation failure. It appears that the allocation process was not concealed and there was no blinding. Transfer was more common in the activation group (68% vs 56%) and there were more pregnancies per transfer, leading to a higher ongoing pregnancy rate: OR=1.9 (0.80 to 4.4). Mean fertilisation rate was higher in the intervention arm: 5.3 (4.4 to 6.2) percentage points. Note that this figure was strongly statistically significant even if reported standard deviations were erroneously standard errors, yet the comparison was described only as 'not significant' in the paper. The mean number of high quality (Grade A or B) embryos per participant was almost identical.

Yang 2015: only available as an abstract so not assessed in this round.

Fawzy 2018 randomised 443 participants evenly between three groups: two active arms using either strontium chloride or calcymicin and a control. Participants had either a diagnosis of male factor infertility (61%) or at least two previous cycles with <30% fertilisation rate (6% total failure). Several methodological issues raise caution. In particular, early randomisation (day 21 of previous cycle) may have resulted in opportunity for selection bias. It is noteworthy that participants in the active arms had both more oocytes retrieved and more mature oocytes than those in the control arm. The trial also finished early following an interim analysis of the data but with no clear specification of any statistical stopping rule applied. The numbers of transfers and of embryos per transfer were similar across groups. The results however show clinical advantage for artificial activation in both active

arms: live birth OR = 3.0 (1.6 to 4.5) and 2.2 (1.2 to 4.0) for strontium chloride and calcymicin respectively. Reported rates of fertilisation and of blastocyst formation from fertilised oocytes were higher in both activated arms.

Shebl 2021 presented a within-patient, sibling oocyte design in 78 couples undergoing ICSI with either a history of <50% fertilisation (n=47) or severe male factor (n=31). Activation was by ionophore (calcimycin) for 15 minutes within 10 minutes of ICSI. All embryos were then cultured in a time-lapse system to allow comparison of morphokinetics. Unfortunately, there was no description of how selection took place so major bias cannot be ruled out. Their analyses of embryo formation and early development did calculate a value per person and then recognise the inherent pairing of the design. Fertilisation and utilisation rates were both significantly higher under activation. Time to appearance of two pronuclei (t2PNa) was reduced by 0.74 (0.28 to 1.25) hours. Other developmental times and occurrence of irregular cleavages did not differ between arms. Interpretation of clinical outcomes is unreliable as there was no description of how selection was undertaken between equal quality embryos in different treatment arms. However, 74 transfers took place using embryos from a single arm (all bar one were elective single embryo transfers) resulting in 22 live births from activated embryos and 11 from control embryos.

Yin 2022 presented a within-patient, sibling oocyte design in 140 couples identified through previous ICSI cycle failure due to either zero (n=66) or <30% (n=74) good quality embryo rate calculated for patients who had a normal fertilisation rate calculated from at least 5 mature metaphase II oocytes. Although the selection of embryos is described as 'random' this appears unlikely: no detail is provided and the 'spare' from an odd number was always allocated to the active arm. Activation was achieved by 10 minutes in ionomycin solution one hour following ICSI. Unfortunately, the inherent matching was ignored in both presentation and analysis of the data. Interpretation of clinical outcomes is unreliable as there was no description of how selection was undertaken between equal quality embryos in different treatment arms. However, 84 transfers took place using embryos from a single arm (all bar one were elective single embryo transfers) resulting in 32 live births. The mean numbers of fertilised oocytes and of day 3 good quality embryos were both higher in the activation arm.

Current rating amber.
Recommendation: GREY for all outcomes [No moderate/high quality study, no safety concerns]

1 (ii) *Failed fertilisation in previous ICSI cycle*

Meerschaut 2012 presented a within-patient design on sibling oocytes from 14 couples with normal sperm but failed or low fertilisation in a previous ICSI cycle. They did not specify allocation method so there is substantial scope for selection bias. Failure to present or analyse the data in a way that recognised the inherent matching precludes statistical interpretation. Ignoring matching, more embryos (74% vs 44%) were fertilised in the 'activation' arm. The nature of the sibling-oocyte design does not allow interpretation of the clinical outcomes.

Montag 2012 prospectively recruited 89 couples undergoing ICSI with previous failed fertilisation (Group 1); fertilisation between 1 and 29% (Group 2); or fertilisation between 30 and 50% (Group 3). All were treated with 10 µmol/l calcium ionophore A23187 for 15 minutes immediately following ICSI. This study was by the same team as Ebner 2012 (reviewed under 1(i) above) and unfortunately shared the same methodological issues. Live births were achieved by 19% of participants in Group 1, 37% in Group 2 and 25% in Group 3. The authors reported fertilisation rates ranging from 42% to 56% in each group. Although the comparison with previous failed cycle is clearly problematic, the uncontrolled cohort demonstrates that successful treatment is possible in this population.

Ebner 2015 largely repeated the study of Montag 2012 from the same team. They prospectively recruited 101 couples undergoing ICSI following previous fertilisation 'problems': failed fertilisation (n=15); fertilisation between 1 and 30% (n=52); fertilisation between 31 and 50% (n=34). All were treated with calcium ionophore A23187 for 15 minutes immediately following ICSI. Although analyses recognised the pairing of participants from index and previous cycle, the major methodological issues from Montag 2012 also apply to this study. There were 35 clinical pregnancies and 28 of these progressed to live birth, including seven twin deliveries. The authors reported substantially 48% fertilisation and 10% cultured to blastocyst from 884 metaphase II oocytes. Only one participant had total fertilisation failure and the remaining 100 all progressed to embryo transfer.

Darwish 2015 undertook a similar but far smaller 'preliminary' study. They prospectively recruited four couples whose previous ICSI cycle was incomplete due to 2PN arrest. The same statistical issues apply to interpretation of the data. All four participants progressed to embryo transfer with a total of eleven embryos transferred. Only one had a positive pregnancy test and this resulted in a healthy twin delivery at term from three transferred embryos. Fertilisation rate was 68% of metaphase II oocytes.

Aydinuraz 2016 presented a within-patient, sibling oocyte design in 21 couples with terato-zoospermia and a low fertilisation rate in the previous cycle. Unfortunately, their presentation and all analyses ignored the matching of the design, precluding statistical interpretation of their data. However, it is clear that only 13 of the 21 couples produced at least one top quality embryo from artificially activated oocytes, whereas 20 achieved this from conventionally cultured oocytes. There was a similar mean number of fertilised oocytes per participant (3.8 versus 3.4) but fewer Grade A embryos from the activated arm (1.3 versus 1.8).

Hao 2016 presented a within-patient design of eleven couples with low (<30%) previous fertilisation, no high quality day 3 embryo or round-headed sperm. Activation was by 10 µmol/l calcium ionophore A23187 for 15 minutes. There was no description of a concealment process and no discussion of blinding. From presented tables, it was possible to obtain the data for the six participants with low previous fertilisation and re-analyse respecting the paired design. Three of these participants had twin deliveries from double embryo transfers. The activated arm produced a mean of 1.8 (-0.8 to 4.4) more fertilised oocytes and 0.5 (-0.1 to 1.1) more good quality embryos.

Li 2019 presented a within-patient, sibling oocyte design in 50 couples identified through previous ICSI cycle failure (15 total fertilisation failure; 18 low fertilisation; 17 severe teratozoospermia). An independent embryologist divided oocytes into groups that were either activated using two 5-minute spells in ionomycin solution or subjected to 'simulated manipulation' by rinsing at comparable times. There is no suggestion that the selection process was randomised. If transferable embryos were achieved from both arms for a participant, the control embryos were preferentially selected. This design prevents interpretation of the clinical outcomes. Unfortunately the development arms were almost exclusively reported per oocyte rather than per participant and the inherent matching was ignored in both presentation and analysis of the data. The authors report higher proportions of transferable day 3 embryos (44% versus 27%) and of fertilisation (50% versus 34%) in the active arm.

Current rating amber.
Recommendation: GREY for all outcomes [No moderate/high quality studies, no safety concerns]

2. Assisted hatching

The previous review included 14 RCTs and three other designs considering a range of techniques for assistance (laser thinning or creation of hole by laser or chemically) in various settings (fresh, frozen-thaw and vitrified; oocytes, embryos, blastocysts). Results were conflicting but no study was deemed of moderate/high quality. Ten additional studies are considered below.

Debrock 2011 randomised couples with frozen/vitrified embryos at the time of thaw/warming. Assisted hatching was by modified quarter laser-assisted zona thinning. This study was unusual in re-randomising couples who wished to participate repeatedly. In total, they randomised 647 treatments for at least 438 different participants. Allocation used a sealed envelope process for concealment. Live birth rate was slightly lower in the intervention arm: OR=0.89 (0.55 to 1.4).

Figueira 2012 reported results from a trial of 60 participants receiving vitrified donor oocytes. Assisted hatching was enabled by laser drilling of a 30µm hole. The allocation process was not reported in sufficient detail to assess risk of bias and the average of more than two embryos transferred at a time may have implications for generalisability to current UK practice. Clinical pregnancy rate was slightly higher in the intervention arm: OR=1.5 (0.54 to 4.4).

Ren 2013 randomised 160 participants aged under 40 undergoing vitrified-warmed blastocyst cycles. Both treatment arms involved laser drilling of a 50µm hole, either located near or opposite to the inner cell mass. The allocation process was not reported in sufficient detail to assess risk of bias and the average of more than two embryos transferred at a time may have implications for generalisability to current UK practice. Live birth rate was slightly lower in those allocated to drilling near the inner cell mass: OR=0.81 (0.43 to 1.5).

Wan 2014 randomised 203 highly selected participants. Low grade, cleavage stage embryos were allowed to develop to blastocysts and then vitrified if high or fair quality. These were then offered to patients who had exhausted, through fresh and vitrified cycles, all cleavage stage embryos that had been assessed as high or fair grade. At this stage participants were enrolled and apparently randomised for use of assisted hatching. Unfortunately, there was no information on which to assess the risk of allocation bias. Assisted hatching was enabled by use of a laser to open 25% of the zona pellucida. Reported results for live birth slightly favoured the intervention arm: OR=1.6 (0.88 to 2.9).

González-Ortega 2015 randomised 303 participants with poor prognosis. Assisted hatching was before an hour before embryo transfer in fresh cycles by laser thinning (‘quarter technique’). Correspondence with authors of the current Cochrane review on this topic suggests this was a high quality study with concealed allocation and blinding of clinicians and participants. Suspicion may be raised that the publication came five years after recruitment completed and was in a low impact journal despite clear results. A mean of 2.5 embryos per transfer also raises questions regarding applicability to the UK setting. Reported results favoured the intervention arm for clinical pregnancy: OR=2.7 (1.6 to 4.6).

Elnahas 2017 randomised 160 participants aged under 40 years and with no history of recurrent implantation failure. Assisted hatching was by laser thinning of one eighth of the surface area of good and excellent quality, day 3, cryopreserved embryos. Clinicians performing embryo transfer were blinded to the intervention but the description of allocation is too vague to assess risk of bias. The clinical pregnancy rate was higher in the intervention arm: OR=1.6 (0.81 to 3.1).

Abulsoud 2019 randomised 130 participants aged over 38 years with at least one previous, failed ICSI cycle. Assisted hatching was by laser thinning of one quarter of the surface area for fresh, day 3

embryos. Clinicians performing embryo transfer were blinded to the intervention but the description of allocation is too vague to assess risk of bias. The clinical pregnancy rate was higher in the intervention arm: OR= 2.5 (1.1 to 5.5). Although the methods described use very similar wording to Elnahas 2017 and the studies are both from Cairo, there is no overlap in the authorship of the two papers and the populations described are different.

Kirienko 2019 randomised 419 participants with broad eligibility criteria. Assisted hatching was by mechanical removal of the zona pellucida from vitrified-warmed blastocysts assessed as high grade at the time of vitrification. Unfortunately, there was no information on which to assess the risk of allocation bias. The ongoing pregnancy rate was similar between groups: OR= 0.94 (0.63 to 1.4).

Fawzy 2020 randomised 966 participants who were undergoing a first or second cycle of ICSI. Assistance entailed a laser pulse to open the zona pelucida of all metaphase II oocytes to facilitate ICSI. This appears to have been a methodologically strong study. Clinical results for ongoing pregnancy favoured the control arm: OR= 0.79 (0.61 to 1.0).

Zhang 2022 conducted an early-phase sibling-embryo study in participants undergoing their first IVF cycle who had more than two highly fragmented day-3 embryos. Sibling embryos were randomised between laser thinning and laser opening of the zona pellucida on day 4, with vitrification of all viable and good quality blastocysts on day 5 or 6. No detail was given to assess risk of bias in the allocation process but analysis correctly accounted for sibling status. No marked differences were identified in blastocyst assessments.

Current rating red.
Recommendation: GREY [Only one moderate/high quality study for LBR. No safety concerns raised across large number of studies]. Note: there are multiple studies of unclear risk of allocation bias that, on the whole, favour assisted hatching in terms of clinical pregnancy rate. Cochrane review concludes unproven for live birth and contains many older RCTs in addition to those included here.

3. Embryo glue (Hyaluronate-enriched culture medium)

The previous review in 2021 covered eleven studies including nine RCTs with a total of over 3000 participants. Most were of poor quality with high risk of bias. However, the largest and methodologically strongest study, Urman 2008, found significantly increased live birth rate when using embryo glue in fresh embryo transfers at day 3 or day 5: OR = 1.5 (1.2 to 1.9). Five additional studies were identified including three conference abstracts. A further trial register entry was identified (NCT00588250) but this was retrospective registration of Morbeck 2007, a previously reviewed trial.

Drew 2014: only available as an abstract so not assessed in this round.

Kleijkers 2016 randomised 836 participants who were undergoing either a first IVF/ICSI cycle or their first following previous success. Rather than 'embryo glue' as such, allocation was for culture throughout in G5, a medium containing hyaluronan, or HTF, a medium without this component. This was a well-designed and well-reported study comparing cumulative outcome to 1 year of follow-up. Live birth was higher with the G5 medium: OR=1.3 (0.98 to 1.7).

Kandari 2019: only available as an abstract so not assessed in this round.

Yung 2021, randomised 550 couples who had had an unsuccessful or cancelled fresh cycle to use of embryo glue in the subsequent frozen transfer. Like Urman, this study was of moderate/high

quality. They reported similar live birth rates in the two groups: OR=0.98 (0.67 to 1.4). They also reported very similar pregnancy losses, twin rates and obstetric outcomes. A clear difference from Urman was the use of frozen rather than fresh transfers. Other differences are likely to have occurred in standard care over the intervening period.

Sellers 2022: only available as an abstract so not assessed in this round.

Current rating amber.
Recommendation: GREEN (At least three moderate/high quality studies with broadly consistent results).

4. Endometrial receptivity analysis

The previous review considered only Simón 2020. This was a single, 3-arm randomised trial comparing 'personalised embryo transfer' based on ERA with two different control groups: elective frozen embryo transfer and fresh embryo transfer. Participants had not suffered previous recurrent implantation failure or miscarriages. The study suffered from a number of methodological issues, in particular from poor protocol adherence with more than 40% of participants not receiving the allocated intervention.

The current review identified four further papers.

Cohen 2020 reported results in a cohort of 97 patients with a history of implantation failure. All underwent ERA assessment. Those assessed to be 'receptive' underwent embryo transfer on the corresponding day of the subsequent cycle. Those assessed to be 'not receptive' were offered a choice on the recommended day of the subsequent cycle between embryo transfer or repeated ERA. Four participants did not progress to personalised embryo transfer, two because the biopsy was considered insufficient. One of the 14 who opted for repeat ERA was assessed to be 'not receptive' for a second time. Denominators presented for clinical outcomes differ without adequate explanation, but six miscarriages and three live births were observed among the 93 women undergoing a first personalised embryo transfer. Also reported by the authors was very low concordance between assessment of receptivity using ERA versus conventional histological dating in 35 women undergoing both: kappa -0.18 (-0.5 to 0.14).

Cozzolino 2020 reported a retrospective cohort analysis of 2110 patients with history of recurrent implantation failure in at least three consecutive cycles during which neither ERA not PGT-A had been used. Patients with abnormal karyotype and various known potential aetiologies were excluded from consideration. This was a very poorly reported study. It is not clear what criteria were used to decide on use of ERA, PGT-A, both or neither. It is also unclear how patients with multiple cycles using different methods were classified into just one of these four categories. There were 3000 analysed cycles of treatment. It also appears that an 'improper' cohort approach may have been used, in which patients were eligible for consideration only if their treatment cycle resulted in a transfer. This may not be a major source of bias for assessment of ERA as it would not be anticipated that the result of the ERA intervention would prevent transfer. However, PGT-A may do so, so any correlation between selection for the two approaches may have indirectly led to bias. Ongoing pregnancy rates were very similar between the 126 patients categorised as receiving ERA and those not: OR= 0.99 (0.69 to 1.4).

Cozzolino 2022 similarly reported a retrospective analysis of 5372 patients with a previous failed embryo transfer, excluding any who had taken part in Simón 2020 (above). This appears again to

have only included patients who progressed to receive a transfer. Two of the authors are noted as "inventors of the endometrial receptivity array patent". All results were presented with participants divided according to receipt or not of donated oocytes, use of PGT and whether standard (non-ERA) cycles used fresh or frozen embryo transfer. Live birth rates were considerably lower with ERA: OR=0.51 (0.41 to 0.62). Cumulative live birth rates were also lower.

Doyle 2022 studied participants undergoing scheduled transfer of a frozen, single, euploid blastocyst. They excluded participants with recurrent implantation failure or recurrent miscarriage. All potential participants underwent ERA assessment and only those with informative results proceeded to randomisation. Comparison was between timing based on the ERA protocol and standard timing of transfer, with ERA results only divulged in the intervention arm. This was a methodologically strong study with arguably a small, built-in advantage to ERA given the exclusion of those with uninformative ERA results. Live birth rate was lower in the intervention arm: OR=0.85 (0.63 to 1.2). Note also that Richter 2023 (doi:10.1093/humrep/dead083) in an Opinion article for Human Reproduction reports further breakdown of the Doyle 2022 data. This shows that those in the intervention arm whose ERA indicated standard timing experienced similar success rates as those allocated to the standard timing arm, with a correspondingly lower success rate where ERA recommended non-standard timing.

Current rating red.
Recommendation: RED (Only one high quality study but safety concerns raised by Cozzolino 2022.

5. Endometrial scratching

The previous review considered 27 studies reporting outcomes for a total of more than 6000 participants. Results for natural/IUI cycles were consistently positive but tended to be from early, small studies at questionable risk of bias. More recently, several large and well-designed studies had reported results for IVF/ICSI cycles with odds ratios for live birth or ongoing pregnancy consistently between 1.0 and 1.4, suggesting possible benefit of a few percentage points but not reaching statistical significance.

The current review identifies twenty further papers, including eight specifically in participants with recurrent implantation failure (RIF).

5 (i) *General population*

Zhou 2008 randomly selected 60 from 121 participants to receive a single procedure between day 5 and 22 of the stimulation cycle. Eligible participants had a regular cycle length, responded well to stimulation and had endometrium "diagnosed irregular echo" by ultrasound. The procedure involved scratching until the echos disappeared followed by scraping if the echoes had been strong. There was inadequate description to assess the risk of bias. Live birth rate was higher with the scratch procedure: OR=2.4 (1.1 to 5.3).

Nastri 2013 allocated 158 participants to a single procedure 7-14 days preceding controlled ovarian stimulation. The study appears to be biased to an unpredictable extent by planned repeated analyses conducted without consideration of cumulative error. It stopped after the fourth such analysis on what appeared at face value to be a significant finding in favour of the scratch procedure: live birth OR=2.4 (1.2 to 4.8).

Gibreel 2015 allocated 387 participants with at least one previous IVF cycle failure but not poor response to stimulation and no known uterine factors. Those in the intervention arm receive two

procedures, two to three days apart, starting on day 21 of the preceding cycle. The control group received a sham procedure with blinding of the participants attempted. This appears to be a good study but there is a question over the allocation process, which is described as using both a random number table and a tombola system. Allocation concealment is therefore unclear and could not have been verifiable. Live birth rate was higher with the scratch procedure: OR=1.4 (0.97 to 2.2).

Bahaa Eldin 2016 allocated 349 participants undergoing IUI for unexplained or mild male factor infertility to receive either a scratch procedure on day 5-7 of the controlled ovarian hyperstimulation cycle with prophylactic antibiotic or just the antibiotic. Timing and process of the randomisation procedure was unclear. Follow-up only extended to diagnosis of clinical pregnancy. This outcome clearly favoured the scratch procedure: OR=2.8 (1.4 to 5.6).

Hilton 2019 allocated 51 participants undergoing a first or second IVF cycle to receive either a scratch procedure 5-10 days before the start of stimulation or no procedure. The study was not blinded but reported a securely concealed randomisation process. Unfortunately, the study finished earlier than planned due to issues with recruitment. It is therefore smaller than intended but should not be biased by the early termination. Live birth rate was higher with the scratch procedure: OR=2.0 (0.66 to 6.3).

Eskew 2019 allocated 100 participants with broad eligibility criteria to receive either a scratch or sham procedure in the luteal phase of the preceding cycle. The study reported a secure randomisation process. Unfortunately they experienced poor recruitment and stopped at the halfway stage on the basis of an unplanned futility analysis, which may well induce bias. Live birth rate was lower with the scratch procedure: OR=0.52 (0.23 to 1.2).

Mackens 2020 allocated 200 participants to a scratch procedure on day 6-8 of the ovarian stimulation cycle for fresh ART transfer. This was a well-designed study that stopped after the second planned interim analysis due to safety concerns regarding miscarriage. Results show higher numbers of clinical pregnancies in the intervention arm with more miscarriages leading to slightly lower live birth rate: OR=0.84 (0.47 to 1.5).

Tang 2020 allocated 220 participants undergoing frozen thawed embryo transfer to a scratch procedure or not on day 3 of the preceding cycle. They do not describe a secure randomisation process and there was no attempt at blinding. Live birth rate was higher with the scratch procedure: OR=1.8 (1.0 to 3.1).

Glanville 2022 allocated 117 participants with polycystic ovary syndrome to a scratch procedure on day 1-12 of the cycle preceding three consecutive cycles of planned ovarian induction. This was a well-designed study but struggled to recruit. The authors acknowledge the resulting imprecision. Live birth was higher after the first cycle but cumulatively lower after the third: OR=0.72 (0.30 to 1.8).

Izquierdo 2022 published a follow-up of the previously reviewed trial, Rodriguez 2020. They report detailed follow-up information on up to four subsequent treatment cycles over the 12 months following the planned randomised comparison. Subsequent attempts, and whether or not each involved a preceding endometrial scratch procedure, were at the discretion of treating clinicians and the participants. They report a total of 120 live births in the initially allocated scratch participants and 114 in the control arm but it is not clear even how this intention to treat perspective should be interpreted.

Madhuri 2022 reported 168 participants with previously failed IUI cycles. They randomised to scratch on day 9 preceding the first of up to three planned cycles of ovarian stimulation for IUI. This was a well-designed study but too small to give a precise result. Live birth was higher after each cycle with ultimate OR=2.2 (0.90 to 5.6). There were just two miscarriages, both in the active arm, and no multiple pregnancies.

Metwally 2022 is the detailed HTA Monograph describing the study previously reviewed as Metwally 2020.

Wong 2022 allocated 220 participants with unexplained infertility planning up to three natural cycles. They randomised to scratch on day 1-12 of the first cycle. As with Glanville 2022 above (same study team) this was a well-designed study but fell short of its initial recruitment target. Live birth was higher after each cycle with ultimate OR=1.4 (0.51 to 3.8).

The current Cochrane review performs formal meta-analysis on the studies meeting the criteria for moderate/high quality evidence: 8 studies (4402 participants); I^2 = 15%; OR=1.1 (0.98 to 1.3). "*This suggests that if the chance of live birth with IVF is usually 27%, then the chance when using endometrial injury would be somewhere between <27% and 32%*".

Current rating amber.
Recommendation: AMBER/GREEN [The more recent evidence reviewed above does not materially affect the previous review but the terminology of the grading has changed. There is consistent, high quality evidence of an effect size ranging from zero to a few percentage points. That is, excluding detriment. Meta-analysis is inconclusive at the standard 95% confidence level but "on balance" there is evidence for a small beneficial effect in terms of live birth. The Committee needs to balance this against cost, inconvenience and pain of the procedure].

5 (ii) *Recurrent Implantation Failure (RIF)*

Baum 2012 randomised 36 participants with recurrent implantation failure to scratch procedures on days 9-12 and 21-24 of the cycle preceding a planned fresh transfer, IVF cycle. The randomisation process was not clearly described. All four live births and five of the six pregnancies occurred in the control group, who underwent a sham procedure.

Zhang 2014 reported a retrospective study that included 55 participants who had received either endometrial scratch or intracavitary physiotherapy. Unfortunately these were 'improper' cohorts, defined by having gone on to receive embryo transfer in the following cycle rather than by receipt of the intervention itself, rendering the results uninterpretable. On face value those who had undergone the scratch procedure had marginally higher clinical and ongoing pregnancy rates.

Zhang 2015 reported a randomised comparison that included 55 participants who had received "hysteroscopic examination and mechanical stimulation" and 57 receiving conventional transfers. Eligible participants had recurrent implantation failure and adequate quality frozen-thawed embryos for transfer. Much concerning the design is unclear, including the timing and process of allocation, making assessment of the results challenging. Clinical results reported for the hysteroscopy group were substantially better than those for control participants. A third group undergoing Chinese medicine prior to embryo transfer had results similar to those of the hysteroscopy group.

Bord 2015 reported a retrospective analysis of 854 cycles in patients with recurrent implantation failure. Unfortunately these cycles were in 183 (or 184) patients and the presented analyses are

invalid as they reverse the risk factors and clinical outcome. It is not possible from the paper to determine either the number of patients undergoing the scratch procedure or the success rates.

Siristatidis 2017 initiated a randomised trial in patients with recurrent implantation failure defined as at least two failed transfers each of at least two good quality embryos. Unfortunately, they found randomisation to be impractical "early after the initiation" of the study. It is not clear exactly why this was the case nor whether and, if so, how recruitment continued after this point. The final data suggested a strong benefit of the scratch procedure in terms of live birth, with low miscarriage and multiple pregnancy rates in both arms.

Gürgan 2019 randomised 305 participants with recurrent implantation failure to receive scratch on day 10-12 of the cycle preceding scheduled IVF treatment. The study is at unclear risk of bias given a lack of information on the timing and process of randomisation. Presented analyses excluded more than 20% of randomised participants. However, assuming unsuccessful outcome in excluded participants allows calculation of an 'intention to treat' effect of live birth as OR=2.1 (1.1 to 4.2).

Tumanyan 2019 reported a comparison of 62 patients with recurrent implantation failure scheduled for IVF. It is unclear whether the study was retrospective or prospective, with inclusion criteria including a stipulation that patients had to undergo consecutive fresh and frozen/thaw cycles to be eligible. Results purported to strongly favour those undergoing a scratch procedure on day 20-22 of the preceding cycle.

Aghajanpour 2021 randomised just 20 participants to scratch procedure on day 9-11 of the cycle preceding IVF treatment. Their focus was on molecular changes and therefore all participants underwent biopsy on day 19-21 of the same cycle. Clinical pregnancy, miscarriage and live birth rates were all unsurprisingly similar given the small numbers and intervention in each arm.
It is also worth noting that the largest high-quality study, Lensen 2019, explicitly considered whether any potential effect of the scratch procedure differed for the subpopulation of participants with a history of recurrent implantation failure. They found no evidence that this was the case.

Recommendation: As above for general population [No moderate/high quality studies explicitly for the sub-population. No evidence that effect for this subpopulation differs from the general].

6. Freeze all

The previous review considered 11 studies including several moderate/high quality RCTs. These are included below alongside additional studies, categorised as requested with the additional consideration of outcomes including ovarian hyperstimulation syndrome (OHSS), time to birth and obstetric outcomes.

6 (i) *General population*

Aflatoonian 2010 described good trial methods but was retracted following "results of an investigation" due to "serious methodological flaws". Clearly results cannot be relied upon.

Shapiro 2011a and Shapiro 2011b compared freezing of all oocytes followed by blastocyst transfer with fresh blastocyst transfer, selecting the best one or two for transfer in each case. The difference was in eligibility criteria, reporting 'normal responders' (8 to 15 antral follicles) in 2011a and 'high responders' (>15 antral follicles) in 2011b. Each used an insecure method of allocation concealment and blinding would not have been possible. Both stopped early on planned interim analyses: the first for efficacy and the second due to unacceptably high multiple conception rate. They did not

report OHSS explicitly but one fresh cycle in 2011a and two in 2011b were “cancelled for medical reasons”. None was cancelled in the corresponding intervention arms. Both reported statistically non-significant higher rates of 10-week pregnancy with the freeze-all policy: OR=1.9 (0.95 to 3.7) and 1.5 (0.74 to 3.2) respectively. Neither reported later outcomes.

Shapiro 2015 randomised 140 couples undergoing IVF with women aged up to 40 years and at least eight antral follicles. They compared freezing of all embryos at the two pronuclear (2PN) stage with freezing of all at the blastocyst stage. They used the same insecure method of allocation described above. They reported higher ongoing pregnancy at 10 weeks for the 2PN arm: 1.3 (0.68 to 2.6). This was despite observing marginally more miscarriages in the 2PN arm. They did not report later outcomes.

Magdi 2017 studied 171 couples undergoing ICSI following unexplained, recurrent implantation failure in at least three previous ICSI cycles with fresh embryo transfer. Unfortunately, allocation was by alternation rather than randomisation, leaving high risk of selection bias. It should also be noted that the high number of embryos transferred in each cycle (>2 in each trial arm) may also limit applicability to the UK setting. They did not report OHSS explicitly and it is possible that cancelled cycles were omitted from the report, which would explain an imbalance in reported group size despite having allocated by alternation. Results for ongoing pregnancy were promising even after adjustment of the report for an intention to treat approach: OR = 2.2 (1.1 to 4.2). Later outcomes were not reported.

Shi 2018 randomised over 2000 good prognosis couples to a fresh or freeze-all strategy for day 2 or day 3 embryos. This was a well-designed and reported study. OHSS was lower in the intervention arm: OR=0.31 (0.13 to 0.74). Live birth was quite similar between arms: OR=0.94 (0.80,1.1). There were no statistically significant differences in reported obstetric outcomes (gestational diabetes and hypertension, pre-eclampsia, preterm delivery) or in neonatal outcomes (birthweight, congenital anomalies, neonatal death).

Le 2018 present a cost effectiveness analysis based on the comparison and data presented by Vuong 2018 (immediately above). They obtained costs for 704 couples. Assuming that those lost to follow-up or declining to provide data were not atypical, the authors estimated that costs were higher on average in the intervention arm. Higher direct medical costs were driven by the additional freezing and thawing entailed. Direct non-medical and indirect costs were similar between arms. Given the similar chance of success observed, it follows that it is unlikely that the freeze-all strategy could be cost effective for this population.

Vuong 2018 randomised nearly 800 good prognosis couples to a fresh or freeze-all strategy for day 3 embryos. This was a well-designed and reported study but it is worth noting that the standard policy was for double embryo transfer. OHSS was only a little lower in the intervention arm: OR=0.75 (0.17 to 3.4). Live birth was quite similar between arms: OR= 1.1 (0.82 to 1.5). Median time to pregnancy was delayed by 1.4 months in the intervention arm. Most obstetric outcomes were similar but the authors noted a lower proportion being small for gestational age and correspondingly higher mean birthweight in the intervention arm. Vuong 2019 adds some secondary, mechanistic exploration but no additional information for this review.

Wei 2019 randomised 1650 good prognosis couples to a fresh or freeze-all strategy for blastocysts. This was a well-designed and reported study using single blastocyst transfer. OHSS was lower in the intervention arm: OR=0.44 (0.14 to 1.4). Live birth was higher: OR= 1.6 (1.3 to 2.0). Time to pregnancy or live birth was not reported but with only one transfer cycle per participant must have been later by design. Most obstetric outcomes were similar but the authors noted a higher

proportion with pre-eclampsia and higher proportion being large for gestational age in the intervention arm.

Stormlund 2020 randomised 460 good prognosis couples to a fresh or freeze-all strategy for blastocysts. They randomised early to incorporate the opportunity to reduce risk of OHSS by using a gonadotrophin releasing hormone agonist to trigger final oocyte maturation. This was a pragmatic comparison using a conventional trigger for fresh transfer but allowing those at high risk of OHSS to delay until a frozen cycle. This was a well-designed and reported study using single blastocyst transfer. There was only one case of OHSS. This occurred in the control arm and required hospital admission. Live birth was quite similar between arms: OR= 0.90 (0.60 to 1.4). Most obstetric outcomes were similar but the authors noted a lower proportion with pre-term delivery and higher mean birthweight in the intervention arm. Further outcomes are promised but not reported.

Simón 2020 was intended as a study of ERA (see 4 above) but the two 'control' groups provide a comparison of elective freeze-all with fresh transfer in 310 low risk women scheduled for blastocyst transfer. This was a poorly designed study in that early randomisation allowed substantial protocol non-adherence, with 40% of participants not receiving their allocated intervention. Under 'per protocol' analysis, OHSS occurred in just one participant who was in the control arm. Success rates were lower in the frozen transfer group: live birth OR (95% CI) = 0.71 (0.45 to 1.1); and cumulative birth: OR=0.95 (0.61 to 1.5). Under per protocol analysis, obstetric complications were rare and similar between arms. There was one neonatal death in the intervention arm and slightly higher mean birthweight in both singletons and twins.

Boynukalin 2020 reported retrospective analysis of all patients undergoing a single blastocyst transfer after elective freeze-all versus all those undergoing a similar transfer after rejecting the offer of elective freeze-all. As well as being subject to clear selection bias through the patient preference design, this study makes the mistake of defining the cohort by those reaching a later stage (single blastocyst transfer) that could have been affected by the preceding decision. They report much lower rates of moderate/severe OHSS in the elective freeze-all arm: OR= 0.04 (0.01 to 0.12). Live birth rates were higher in the freeze-all arm: OR=1.5 (1.3 to 1.8) for first transfer and for cumulative live birth. Obstetric complications were similar between groups. Birthweights were higher in the freeze-all arm.

Li 2021 randomised 360 couples who were about to undergo endometrial preparation for their first frozen transfer in a freeze-all cycle. Their comparison was between preparation methods: down-regulation ovulation-induction using HMG versus a modified natural cycle approach. The study was at risk of bias due to insecure concealment of the allocation process. Cycles were cancelled for five participants in the down-regulation arm and none in the conventional arm to prevent OHSS. Despite this, higher average number of embryos per transfer and higher average quality of embryos in the conventional arm, the ongoing pregnancy rate was higher in the down-regulation arm: OR=2.2 (1.5 to 3.4). The paper did not report obstetric and neonatal outcomes outcomes.

Wong 2021 randomised 204 couples with any indication, regardless of available numbers of follicles or embryos, undergoing their first treatment cycle. They randomised before the start of down-regulation and compared a policy of cryopreservation of all embryos on day 6 with a strategy of fresh single blastocyst transfer on day 5 followed by cryopreservation of all surplus embryos on day 6. This was a well-designed and reported study. There were three cases of OHSS requiring hospitalisation, all in the control arm: OR= 0 (0 to 2.4). Success rates were much lower in the freeze-all arm: live birth OR (95% CI) = 0.27 (0.11 to 0.66); and cumulative birth to 12 months OR=0.54 (0.28 to 1.1). They did not report detailed obstetric outcomes. Time to ongoing pregnancy was reported with a statistically significant log-rank test (p=0.02) favouring the control arm. The authors report no

evidence of a difference in birthweights or other neonatal outcomes, and confirm that there were no congenital abnormalities in either arm.

Maheshwari 2022 randomised 619 couples between freeze-all and fresh transfer strategies if they were undergoing a first, second or third cycle of IVF treatment with their own gametes and had no clinical indication for elective freeze-all. Fuller details are presented in Maheshwari 2022a and used here. Unfortunately, this study suffered from poor recruitment and from very high non-adherence with the freeze-all strategy: 31% received a fresh transfer despite their allocation. Moderate/severe OHSS was lower in the freeze-all arm: OR=0.27 (0.10 to 0.73). Live birth was also lower: OR= 0.76 (0.54 to 1.1). This conclusion was similar under re-analyses using different strategies (per protocol, as treated and compliance-adjusted). There was no evidence of differences in obstetric outcomes such as gestational diabetes or hypertension. Birthweights and rate of congenital anomalies were also similar between trial arms.

Current rating amber.

Recommendation: AMBER for live birth [Conflicting findings from 4 moderate/high quality studies]
GREEN for OHSS [On balance, consistent evidence]
GREY for obstetric/neonatal outcomes [Studies underpowered for these]

6 (ii) *Populations at increased risk of OHSS*

Chen 2016 randomised 1508 couples undergoing a first IVF cycle for PCOS at Day 3 if they had embryos and were deemed low risk for OHSS. Couples either underwent Day 3 freezing of all embryos or Day 3 transfer of (usually two) embryos. This was a well-designed and clearly reported study. Live birth rate was higher in the freeze-all arm: OR=1.3 (1.1 to 1.6). There were two stillbirths in the freeze-all arm and five neonatal deaths. The odds ratio for 'take-home baby' was as for live birth above. Moderate/severe OHSS was markedly lower: OR=0.18 (0.09 to 0.35). Mean birthweight of singletons was higher (3511g vs 3349g) whereas that for twins was similar (2480g). There were similar numbers of cases of gestational hypertension (10 vs 5) and congenital anomalies (24 vs 17).

Rahav Koren 2018: only available as an abstract so not assessed in this round.

Ye 2018 reported a retrospective comparison of 110 patients receiving each of two freeze-all strategies for women at high risk of OHSS defined by PCOS. One group received progestin-primed ovarian stimulation using a lower dose of hMG with 50mg clomiphene citrate. The other received standard stimulation. There was one case of moderate/severe OHSS in each arm. Cumulative live birth was higher in the standard stimulation arm: OR= 0.59 (0.33 to 1.1). They did not report obstetric outcomes. Birthweights were similar between arms.

Deng 2019 reported a retrospective cohort of 21 patients at high risk of OHSS defined by having at least 30 follicles of at least 11mm diameter or pre-trigger peak oestradiol of >10,000pg/mL. All patients were undergoing ovarian stimulation with a GnRH antagonist protocol and all received a second dose of GnRHa 12 hours after the first and again at 0.25mg/day for three days following oocyte retrieval. There were 15 (71%) cases of mild OHSS but none progressed to moderate/severe. No comparison was made and no outcomes were reported regarding subsequent transfers.

Shrem 2019 reported a retrospective cohort of 480 patients at high risk of OHSS defined by having PCOS, antral follicle count>8, or 18 follicles >10mm diameter. All patients underwent a GnRH antagonist protocol, GnRH agonist trigger and a freeze-all strategy. In addition to those receiving the standard trigger, one group received 0.5mg/day oral cabergoline for 7 days and one received this plus 5 days of GnRH antagonist from the day of oocyte retrieval. As with Deng 2019 (immediately

above), this was more a study of how to prevent OHSS and contained no comparison of the freeze-all strategy. There were no cases of severe OHSS. Mild/Moderate OHSS was reported for 80 (38%), 48 (29%) and 19 (18%) of patients in each of the three groups. No outcomes were reported regarding subsequent transfers.

Santos-Ribeiro 2020 randomised 209 couples at high risk of OHSS defined by their high response to ovarian stimulation. They were allocated to a fresh or freeze-all strategy for either day 3 or day 5 transfer using the same pre-defined criteria in each arm. This was a well-designed and reported study using single or double embryo transfer. There were nine cases of moderate/severe OHSS, all in the control arm: OR= 0 (0 to 0.49). Live birth rate was very similar between arms: OR= 1.1 (0.61 to 1.8). Cumulative live birth to 24 months was also similar. Time to pregnancy was slightly reduced in the control arm: HR= 0.92 (0.68,1.2) but with very similar trajectories after the second month. That is, similar patterns with a one-cycle lag with the freeze-all strategy. They did not report detailed obstetric or neonatal outcomes.

Deepika 2021 presented follow-up data from an earlier randomised trial of 210 participants undergoing a first treatment cycle scheduled for a freeze-all strategy. Randomisation was between trigger using GnRH agonist versus conventional hCG. Sixteen (8%) participants had ceased follow-up prior to this follow-on report. Moderate/Severe OHSS was the primary outcome in the original. This occurred in none of the GnRH agonist arm and 38 of the hCG arm: OR=0 (0 to 0.07). Using the GnRH agonist trigger, live birth was higher in the first cycle: OR=1.5 (0.81 to 3.0). This was also the case for cumulative live birth measured up to three transfers: OR=2.2 (1.2 to 3.8). The paper did not report obstetric and neonatal outcomes.

Huang 2021 presented a retrospective analysis of 333 couples with PCOS undergoing their first IVF cycle using a freeze-all strategy. They reported results for 160 couples using GnRH antagonist to prevent premature LH surge prior to a change in their routine practice and 173 couples after switching from GnRH antagonist to dydrogesterone for this purpose. They observed no cases of OHSS in either group. Live birth rate was similar in the two groups: OR=1.0 (0.68 to 1.6). The paper did not report obstetric and neonatal outcomes.

Vuong 2021 randomised 40 couples undergoing in vitro maturation at high risk of OHSS due to high antral follicle count, including those with PCOS. Randomisation was to fresh embryo transfer or a freeze-all strategy for day 3 embryos, with all but one couple receiving two embryos at transfer. This was a well-designed and reported study. There were no cases of OHSS and no cases of either gestational diabetes or hypertension. Live birth was higher in the freeze-all arm: OR=6 (1.5 to 25) but it should be noted that this was based on small numbers. Time to live birth for those delivering after the first cycle was unsurprisingly a median of 43 days less in the fresh transfer arm. Birthweights were similar but too sparse for meaningful comparison and there were no congenital abnormalities.

Current rating amber.

Recommendation: GREY for live birth [Only 2 moderate/high quality studies]
GREEN for OHSS [Only 2 studies but consistent with general population]
GREY for obstetric/neonatal outcomes [Only one RCT reporting]

7. IMSI

The previous review considered eight studies including just one randomised trial that provided moderate/high quality evidence, Setti 2013. They studied IMSI in couples consisting of a woman aged over 37 years and a fertile man, with the hypothesis that older eggs may be less able to repair

DNA damage, and found improved ongoing pregnancy rate: OR= 4.1 (1.2 to 15). Further studies are considered below categorised as requested.

7 (i) *General population*

Balaban 2011 randomised 168 couples to undergo IMSI or conventional ICSI. No eligibility criteria are specified but the table of participant characteristics shows that nearly half had male factor infertility. This was a randomised study but at unclear risk of bias through concealment of allocation. Live birth rate was slightly higher in the intervention arm: OR= 1.3 (0.68 to 2.3). The odds ratio was similar in the subgroup of participants with male factor.

Figueira 2011 randomised 120 couples to undergo IMSI or conventional ICSI. Participants were undergoing IVF-PGS cycles due to "advanced maternal age" and had at least six oocytes. Severe male factor cases were excluded. This was a randomised study but at unclear risk of bias through concealment of allocation. Clinical pregnancy rate was slightly higher in the intervention arm: OR= 1.3 (0.64 to 2.7).

Setti 2012 randomised 160 couples to undergo IMSI or conventional ICSI. Participants were undergoing IVF-PGS cycles due to "advanced maternal age" and had at least six oocytes. Severe male factor cases were excluded. This was a randomised study but at unclear risk of bias through concealment of allocation. It comes from the same team as Figueira 2011 with the same start date for recruitment and the same eligibility criteria so appears to be an extension. Clinical pregnancy rate remained slightly higher in the intervention arm: OR= 1.4 (0.76 to 2.6).

Setti 2014 undertook a review of comparative studies, without regard to study design. They identified a few studies as randomised controlled trials between 2008 and 2011 that have not been reviewed here. From the summary information presented these studies appear to favour IMSI over ICSI for the outcome of 'pregnancy rate'. However, the risk of bias inherent in these studies is not clear and nor can it be assumed that the trial authors and reviewers have correctly analysed by randomised participants rather than by numbers of treatment cycles.

Current rating red.
Recommendation: GREY [Only one moderate/high quality study, no safety concerns]

7 (ii) *Male factor*

Antinori 2008 randomised 446 couples with male factor (oligoasthenoteratozoospermia) and no known female factor and a female partner aged 35 or younger to IMSI or conventional ICSI. This was a methodologically strong study. They reported higher ongoing pregnancy beyond 24 weeks: OR=1.9 (1.2 to 3.0).

Knez 2011 randomised 57 couples with male factor to IMSI or conventional ICSI and excluded women over 42 years or with PCOS or endometriosis. Although reporting a concealed randomisation process there the imbalance in group size (20 vs 37) is unlikely to occur through chance. The clinical pregnancy rate was much higher in the IMSI arm: OR=3.8 (0.80 to 18).

Mahmoud 2011: only available as an abstract so not assessed in this round.

Setti 2011 randomised 500 couples with isolated male factor who had at least six oocytes. There was insufficient detail to assess risk of bias. The clinical pregnancy rate was very similar between groups: OR=1.0 (0.71 to 1.5).

Knez 2012 randomised 122 couples with male factor (isolated teratozoospermia) and at least six mature oocytes to receive IMSI or conventional ICSI with a policy of transferring up to two blastocysts. The methods of this study were poorly reported making it hard to assess risk of the most common biases. In particular there was no information on the method or timing of randomisation and no explanation for the imbalance in sample size between arms. The only reported clinical outcome was clinical pregnancy rate, which was higher in the IMSI arm: OR=3.2 (1.5 to 7.0).

Check 2013 randomised 24 couples with male factor (DFI>30%) and women under 40 years who had three previous, failed embryo transfers. The methods of this study were poorly reported making it hard to assess risk of the most common biases. The live birth rate was similar in the two groups with a wide confidence interval ruling out very little: OR=1.0 (0.16 to 6.3).

Sifer 2014 studied 91 couples with no more than 2 previous failed ICSI cycles where the man had severe teratozoospemia. All underwent IMSI using fresh sperm with the strategy of transferring up to two embryos at day 2 or 3. Their study groups were defined by the availability of sperm using the Vanderzwalmen criteria: Grade I & II available or having to use Grade III or IV. Both clinical pregnancy and live birth were marginally higher in the second group. However, interpretation is unclear given the potential confounding inherent in this design. It may be that the grading is not relevant to viability or that, for example, those with higher grade sperm may have been in couples with poorer female prognosis.

Mangoli 2020 randomised 95 couples with male factor, primary infertility where the woman was considered healthy and had at least six mature oocytes. They compared IMSI with ICSI under a policy of transferring two day-3 embryos. There was no description of the randomisation process to allow assessment of risk of bias. Note that this study was concurrent in the same centre as Mangoli 2019 (see previous review). The difference here is that these participants had at least 3 years of primary infertility, which was listed as an exclusion criterion in the earlier paper. Forty couples in each arm underwent transfer and live birth was more frequent in the IMSI arm: OR=1.5 (0.56 to 3.9).

Current rating red.
Recommendation: GREY [One moderate/high quality study, no safety concerns]

8. Intralipids

The previous review considered three RCTs that were each at high risk of bias but consistently supported the use of intralipids. These are included below alongside additional studies, categorised as requested with the additional consideration of miscarriage rates.

8 (i) *General population*

El Khayat 2015: only available as an abstract so not assessed in this round.

Gamaleldin 2018: only available as an abstract so not assessed in this round.

Singh 2019 studied about 100 women with recurrent implantation failure undergoing IVF. Infusions were given immediately following oocyte retrieval and again one hour after embryo transfer. This too was a poorly reported study at risk of bias from both allocation concealment and blinding. It was also conducted with a policy of transferring two or three embryos when available. The reported result was a marked increase in live birth rate with intervention: OR (95% CI) = 3.3 (1.2 to 8.8). Just

two participants, both in the intervention arm, failed to progress from clinical pregnancy to live birth.

Al-Zebeidi 2019 studied nearly 150 women with unexplained recurrent implantation failure undergoing ICSI. Infusions in this study were given at the time of embryo transfer and again at the time of pregnancy testing. This too was a poor study at risk of bias from allocation concealment and with no attempt at blinding. A double embryo transfer policy was used with three embryos allowed for older women. Again, the reported live birth result favoured intervention but this time without reaching statistical significance: OR (95% CI) = 1.4 (0.57 to 3.4). This despite more reported miscarriages: OR = 1.4 (0.57 to 3.4).

Current rating [?].
Recommendation: GREY/RED for all outcomes [No moderate/high quality studies. Question over whether committee considers the safety concerns raised over congenital malformations justify the red rating].

8 (ii) *Populations with immunological testing*

Dakhly 2016 randomised nearly 300 participants with secondary recurrent miscarriage and elevated levels of natural killer cells (>12%), who were undergoing IVF, to either IV infusion on the day of oocyte retrieval or matching placebo. Unfortunately this was a poorly reported study with scope for serious bias in the allocation and blinding processes. It was conducted with a policy of transferring two or three embryos. The reported result was a marked increase in live birth rate with intervention: OR=2.1 (1.3 to 3.5). There were also fewer miscarriages in the intervention arm: OR=0.66 (0.35 to 1.2).

Meng 2016 recruited 192 participants with recurrent miscarriage and $CD56^{+}CD16^{+}$>20%. Participants were randomised between IV intralipid and IV immunoglobulin. Each started monthly and continued through to week 12 of gestation in the event of pregnancy. Injections were continued monthly for three months then stopped for three months, with the pattern repeated for up to 24 months. There is no suggestion that interventions were blinded and too little information to judge risk of bias in the allocation process. Substantial loss to follow-up occurred after intervention was completed at 12 weeks gestation. Ongoing pregnancy rate to this point was quite similar: OR=1.2 (0.62 to 2.2). Miscarriage within this timeframe was lower in the intralipid arm: OR=0.58 (0.23 to 1.5).

Rogenhofer 2021 described a patient preference study of 12 participants with recurrent miscarriage that was unexplained other than being positive for anti-trophoblast antibodies (ATAb) activity. Ten chose to accept off-label IV infusions of intralipids from their positive pregnancy test every three weeks up to the 33rd week of gestation. The remaining two agreed to repeated monitoring. These two both miscarried a euploid fetus within the first trimester. There was one miscarriage of a fetus with trisomy 16 in the active arm. All other pregnancies continued to live birth with no neonatal malformations. The study is not of a suitable design or scale to draw statistical conclusions regarding these clinical outcomes. The focus was on ATAb activity which was noted to decrease progressively throughout pregnancy with intralipid treatment.

Recommendation: GREY/RED for all outcomes [No moderate/high quality studies, no safety concerns specific to this sub-population but may need to consider safety concern raised above]

9. IV immunoglobulin

The previous review considered two well designed but small RCTs in participants with unexplained secondary recurrent miscarriage. These are included below alongside additional studies, categorised as requested with the additional consideration of miscarriage rates.

9 (i) *General population*

Stephenson 2010 randomised 77 participants with idiopathic secondary recurrent miscarriage in a double-blind, placebo controlled trial. IVIG was delivered at a dose of 500mg/kg two to three weeks before the next anticipated menstrual period and then every four weeks for up to 6 cycles or until reaching 18 to 20 weeks gestation. This was a well-designed study but small. The size of study ruled out very little. The live birth odds ratio was 1.2 (0.47 to 2.9), consistent with the intervention more than doubling or halving the odds of success. Miscarriage was lower in the intervention group but, again, with wide confidence intervals: OR=0.38 (0.07 to 2.1).

Christiansen 2014 conducted a study of similar size in a similar patient population. The main difference was that IVIG was first given on confirmation of pregnancy by repeated biochemical testing. A total of eight infusions was given up to week 15 of gestation at a dose of approximately 25g for those up to 75kg of weight and 35g for heavier women. This was a well-designed study but small. Live birth was similar in the two arms: OR=1.2 (0.51 to 2.9). Miscarriage rates were also very similar: OR=1.1 (0.40 to 2.8).

Jørgensen 2020 reported further blood analyses from a trial by Christiansen 2002. They found that participants in the IVIG arm had markedly boosted production and release of smaller extracellular vesicles. The initial study randomised 58 women with recurrent miscarriage to IVIG or placebo from the time of positive pregnancy test. It was a methodologically strong study but too small to give a precise estimate of effectiveness. Infusions of 0.8g/kg bodyweight were given weekly from week 5 to week 10 of gestation then fortnightly through to week 20. From then to week 26 the fortnightly dose increased to 1.0g/kg. Live births and, conversely, miscarriages were identical between the two groups: OR=1.0 (0.36 to 2.8).

Current rating [?].
Recommendation: AMBER for all outcomes [3 RCTs providing moderate quality evidence. Not 'conflicting' but results too imprecise to determine effectiveness at this stage].
9 (ii) *Populations with immunological testing*

Dendrinos 2009 randomised 85 couples with recurrent miscarriage and positive antiphospholipid antibodies. Couples received IVIG from the date of positive pregnancy test every four weeks through to week 32, or low molecular weight heparin and low dose aspirin. The paper did not describe concealment of the randomisation process or blinding of participants. Live birth was much lower in the IVIG arm - OR=0.30 (0.12 to 0.73) – owing to a higher miscarriage rate and two intrauterine deaths.

Cohen 2015: only available as an abstract so not assessed in this round.

Yamada 2015 conducted a prospective, single group study of 14 women with unexplained recurrent miscarriage (13 primary, one secondary) and previous failure of low dose aspirin and heparin treatment. 20g IV immunoglobulin was given on each of three days immediately following confirmation of a gestational sac. Natural killer cell status was not an eligibility criterion. Four of the 14 pregnancies resulted in healthy live birth. Eight ended in first trimester miscarriage and two in 'stillbirths' at 17 and 21 weeks gestation. Natural killer cell activity was reduced in all but three of

the participants, each of whose pregnancy resulted in miscarriage. No comparative data are presented.

Lee 2016 conducted a retrospective analysis of 189 women with at least two previous miscarriages who had undergone full assessment and had known follow-up data. The latter feature could bias comparisons but the extent of this concern would depend on the number of potential participants excluded from the cohort, which is not reported. Women were categorised according to known or unknown aetiology and presence or absence of a cellular immune abnormality, defined by natural killer cells or T helper cells. Those with immune abnormality were treated with IVIG 400mg/kg at week 4-6 of gestation, repeated every three weeks up to 30 weeks gestation. All patients were given standard care according to any known aetiology. In total, 111 women received IVIG with 94 live births and 17 miscarriages. This was compared with the 70 live births and 8 miscarriages observed in 78 women who were not diagnosed with immune abnormalities and therefore did not receive IVIG.

Meng 2016 recruited 192 participants with recurrent miscarriage and $CD56^{+}CD16^{+}$>20%. Participants were randomised between IV intralipid and IV immunoglobulin under the hypothesis that intralipid would achieve similar therapeutic aims with reduced adverse effect profile. The study is described more fully above under 8 (ii). In brief, treatments were started prior to pregnancy and continued through to week 12 of gestation in the event of pregnancy. The study was poorly reported and at risk of bias. Ongoing pregnancy to 12 weeks gestation was quite similar between groups: OR=1.2 (0.62 to 2.2). Miscarriage within this timeframe was lower in the intralipid arm: OR=0.58 (0.23 to 1.5).

Ahmadi 2017 reported a prospective patient preference study in 94 participants with recurrent miscarriage and abnormal flow cytometry for either natural killer or T helper cells. Participants who volunteered were given 400mg/kg IVIG after a positive pregnancy test and every 4 weeks up to 32 weeks gestation. Their outcomes were compared with a concurrent control group of those who chose not to receive IVIG. The study is subject to bias both from the allocation process and the lack of blinding. Results were strongly in favour of intervention: live birth OR=8.7 (3.1 to 24) and, conversely, miscarriage OR=0.11 (0.04 to 0.32).

Recommendation: GREY for all outcomes [No moderate/high quality studies, no safety concerns]

10. PGT-A (Blastocyst)

The previous review considered randomised trials that made subtly different comparisons in a range of settings. Previously reviewed studies are included below alongside seven additional publications, categorised as requested with the additional consideration of miscarriage rates and time to birth.

10 (i) General population

Yang 2012 randomised 112 couples undertaking a first cycle of ICSI and scheduled for elective single embryo transfer. The study was restricted to women under 35 years old. They employed assisted hatching on day 3 in both groups to facilitate PGT-A of the blastocyst and compared outcomes in the fresh transfer cycle. Yang 2012 did not report allocation concealment but attempted to blind patients to their intervention. There were only three miscarriages reported: OR= 0.49 (0.04 to 5.6). Ongoing pregnancy (beyond 20 weeks) was much higher in the PGT-A arm: OR=3.8 (1.7 to 8.3). The single cycle comparison precluded consideration of time to success.

Forman 2013 randomised 175 couples undertaking a first or second cycle of IVF who had produced at least two good quality blastocysts. They employed assisted hatching on day 3 in both groups to facilitate PGT-A of the blastocyst and compared outcomes in the first transfer cycle. They compared single transfer in the PGT-A arm with double transfer (DET) in the controls. They reported a secure randomisation process but did not attempt blinding. Miscarriage was lower in the PGT-A arm: OR=0.44 (0.17 to 1.1). Higher clinical pregnancy in the control arm meant that ongoing pregnancy rate was similar: OR= 0.83 (0.45 to1.5). The single cycle comparison precluded consideration of time to success.

Scott 2013 was from the same research team as Forman 2013 with apparently overlapping recruitment periods. They randomised 155 couples undertaking a first or second cycle of IVF who had produced at least two good quality blastocysts. They employed assisted hatching on day 3 in both groups to facilitate PGT-A of the blastocyst and compared outcomes in the first transfer cycle. They reported a secure randomisation process but did not attempt blinding. Miscarriage was lower in the PGT-A arm: OR=0.41 (0.15 to 1.1). Despite slightly higher clinical pregnancy in the control arm, live birth rate was higher as a result: OR= 6.5 (2.3 to 18). The single cycle comparison precluded consideration of time to success.

Ozgur 2019, randomised 220 couples undertaking ICSI under a freeze-all policy who had produced at least two good quality blastocysts. The study was restricted to women no more than 35 years old and compared only the first transfer cycle. Ozgur 2019 did not report allocation concealment but attempted to blind clinicians but not patients to the intervention. Miscarriage was lower in the PGT-A arm: OR=0.44 (0.15 to 1.3). This did not compensate for lower clinical pregnancy rate so live birth rate was also lower: 0.75 (0.44 to 1.3). The single cycle comparison precluded consideration of time to success.

Munné 2019 randomised 661 couples undertaking ICSI under a freeze-all policy with similar characteristics to those of Ozgur 2019. They allowed women up to 40 years old with zero to two previous failed attempts and compared only the first transfer cycle. They reported a secure randomisation process with blinding of clinical staff and patients. Miscarriage was similar in the two arms: OR=0.87 (0.51 to 1.5). Live birth results were also similar: OR (95% CI) = 0.93 (0.69 to 1.3). The single cycle comparison precluded consideration of time to success.

Cimadomo 2019 reported a retrospective study of transfers using poor quality blastocysts. Their clinic approach included PGT-A of all blastocysts regardless of morphological grade. Following ICSI, blastocysts were routinely biopsied and then vitrified for subsequent use in single transfers. Unfortunately, the study is reported entirely in terms of cycles and blastocysts so it is unclear how many couples contributed data and consequently not possible to draw any clinical conclusions. There were 2757 retrievals, 2217 of which resulted in at least one blastocyst and 724 of which culminated with live birth. The paper does demonstrate that, although prognosis is poorer, it remains possible to achieve live birth using euploid poor quality embryos with 21 such deliveries. Their analysis found no evidence of safety concerns in terms of obstetric or neonatal outcomes.

Yan 2021 randomised 1212 couples with good prognosis undertaking their first cycle. Participants with at least three good quality blastocysts were assigned to selection based on PGT-A using next-generation sequencing or conventional morphological criteria. All blastocysts were then cryopreserved before use in successive single transfers for up to 1 year. This was a high quality study using concealed randomisation but not attempting to blind clinicians or participants. After the first transfer cycle there were fewer miscarriages in the PGT-A arm: OR=0.69 (0.45 to 1.1). A similar pattern occurred in each subsequent transfer but it is not possible to discern from the reporting how many different couples experienced miscarriage. Although live birth was higher from the first

transfer, cumulative live birth was lower in the PGT-A arm: OR=0.75 (0.57 to 1.0). Time to conception resulting in live birth was also significantly longer (p=0.01 from Kaplan-Meier presented in supplement).

De Munck 2022 presented secondary analysis of a previously published sibling oocyte study. They studied 30 couples who had each produced at least ten cumulus oocyte complexes and were scheduled for PGT-A selection of blastocysts using next-generation sequencing. Oocytes were randomised to conventional IVF or ICSI but there was no description of the randomisation process in either report. The original paper concluded that euploid blastocysts were as likely using conventional IVF but the design does not allow for clinical comparisons. De Munck 2022 adds comparison between conventional IVF and ICSI for morphokinetic parameters from time-lapse imaging but nothing regarding the potential clinical benefit of using PGT-A for blastocyst selection.

Idárraga 2022 presented retrospective analyses of 54 couples who had undergone PGT-A either of day 3 embryos or of blastocysts. Their focus was on reporting results of the testing and there is no clinical comparison presented. In all, 32 couples had progressed to transfer at the time of the report and 13 of these had progressed to live birth. No first trimester miscarriages were reported.

Current rating red.

Recommendation: GREEN for miscarriage [several moderate/high quality studies, consistent]
RED/BLACK for live birth [Yan looks definitive but could argue either way]
RED/GREY for time to success [Yan looks definitive but just 1 study of this]

10 (ii) Older women

Ubaldi 2017 reported a retrospective analysis of 137 couples with women aged 44, 45 or 46 years. All participants underwent ICSI using a policy of blastocyst selection based on PGT-A with elective freeze-all. (See Cimadomo 2019 above). With 13 couples undergoing a second cycle, the cumulative live birth was 12 (9%) from 13 clinical pregnancies and just one miscarriage before 20 weeks.

Verpoest 2018 randomised 396 couples with women aged 36 to 40 years in a multi-national trial. Couples were eligible if they had no history of poor ovarian response in previous cycles, no more than two previous cycle failures and no more than two previous miscarriages of a clinical pregnancy. Participants received the standard ICSI protocol for their centre. **Please note that PGT-A was by polar body biopsy six to nine hours after ICSI, which may constitute a different class of intervention**. Participants were then followed up to record live births within 12 months. This was a well-designed, clearly reported trial. It should be noted however that outcomes of spontaneous pregnancies were not reported and that the definition of 'live birth within 12 months' could favour premature over term deliveries. Miscarriage was lower in the PGT-A arm: OR=0.45 (0.23 to 0.88). Live birth following the first transfer was similar between arms and cumulative live birth showed a similar pattern: OR=1.0 (0.66 to 1.7) for the latter. Time to live birth was presented graphically showing similar patterns between groups: log-rank test p=0.82.

Current rating red.

Recommendation:
GREEN/GREY for miscarriage [1 study if eligible, nothing to contradict effect in general population]
BLACK/GREY for live birth [1 study if eligible, nothing to contradict effect in general population]
RED/GREY for time to success [1 study if eligible, nothing to contradict general effect]

11. PICSI

The previous review considered seven studies. It was dominated by the well-designed and reported trial of Miller 2019 that ruled out any major effect of PICSI in their population of couples using own gametes and scheduled for fresh transfer on days 3 to 5. Previously reviewed studies are included below alongside six additional publications, categorised as requested with the additional consideration of miscarriage rates.

11 (i) *General population*

Parmegiani 2012 randomised 100 couples to PICSI or to 'Sperm Slow' selection. Women were aged up to 41 years and sperm counts were at least one million with 5% motility. The randomisation process was not well described so at unclear risk of bias, although baseline characteristics of the two groups were very similar. There was one less miscarriage in the PICSI arm, OR= 0.78 (0.20 to 3.1), and two more live births: OR=1.2 (0.52 to 2.8).

Majumdar 2013 studied couples undergoing their first cycle of IVF-ICSI for unexplained infertility (normal semen parameters) and excluded women over 38 years old. The study was at unclear risk of bias with missing detail on allocation method and blinding, as well as post-randomisation exclusions (no embryo transfer) whose group assignment was unreported. There were fewer miscarriages in the PICSI arm: OR=0.46 (0.11 to 1.9). Odds of clinical pregnancy were equal in the two groups, with a slight and non-significant benefit for live birth: OR=1.3 (0.62 to 2.6). If the participants not reaching embryo transfer were assigned to the intervention group, which would still remain smaller, the OR for live birth reduces to 1.1.

Troya 2015 studied unselected couples (normal semen parameters) undergoing ICSI. There was high risk of bias with no information on allocation method or blinding and no explanation for imbalanced group sizes, suggesting the possibility of unreported, post-randomisation exclusions. There were only three miscarriages between clinical and 'ongoing' pregnancy: OR=0.58 (0.05 to 6.6). No evidence was found for clinical benefit of PICSI over conventional ICSI. In 102 reported couples, pregnancy ongoing at 20 weeks gave OR=2.0 (0.85 to 4.7).

Miller 2019 was a pragmatically designed, well-conducted and well-reported trial of more than 2700 participants across 16 sites. ICSI had been recommended on the basis of semen assessment in over 95% of participants. Miscarriage rates were lower in the PICSI arm: OR=0.60 (0.43 to 0.83). The primary analysis ruled out major differences in the outcome of live birth: OR (95% CI) = 1.1 (0.95 to 1.3). Further secondary analyses considered stratification by factors identified in the earlier trials including, for example, hyaluronan sperm binding score, none of which showed evidence of differential effects.

Novoselsky Persky 2021 conducted an unusual retrospective analysis of 45 couples undergoing ICSI. All had accepted half their oocytes being "randomly assigned" to PICSI as part of routine clinical practice for staff to gain experience of the new method. There is no detail on the process of allocation and little on eligibility beyond "mainly couples with previous failure". Nearly two thirds of retrospectively identified couples had male factor infertility and women were aged from 27 to 34 years. The best embryo(s) selected for fresh transfer was deemed to come from the PICSI arm on 22 occasions, from ICSI on 13, and from a mix on 9. The remaining couple had no embryos of sufficient quality from either method.

Current rating: red.

Recommendation: GREY for all outcomes [Only 1 moderate/high quality study, no safety concern] NB As for endometrial scratch, there is high quality evidence that any effect on live birth is no more than a few percentage points. Here this is based on a single definitive study but it would seem implausible that future studies will be funded sufficient to materially affect the conclusion that PICSI leads to fewer miscarriages and similar live birth rate. A randomised trial with 90% power to detect a difference in live birth rates between 25% and 27% would require in excess of 20,000 participants. The Committee could consider whether GREEN for miscarriage and BLACK for live birth may be more informative summary information for patients despite not strictly fitting the current definitions for these grades.

11 (ii) *Male factor infertility*

Worrilow 2013 studied infertile men and excluded women over 40 years old. The design effectively comprised two separate trials: the first in couples for whom hyaluronic acid (HA) binding was greater than 65% in unprocessed semen, and the second for whom binding was less than 65%. In an otherwise well-conducted study at apparently low risk of bias, the presentation of results was confused by stratifying results by non-design features (post-processing parameters). There was sufficient information from combining text and figures to recalculate the results of the randomised comparison. Miscarriage of clinical pregnancy before classification as 'ongoing' was below 5% in all arms bar that of the routine selection arm of the low binding stratum: OR=1.4 (0.23 to 8.7) for high binding; OR=0.18 (0.4 to 0.84) for low binding. Contrary to the article title, the clinical outcomes were slightly worse in the PICSI group with no evidence of a difference. The ongoing pregnancy rate figures gave OR= 0.87 (0.52 to 1.4) for those with high binding and OR= 0.99 (0.63 to 1.6) for those with low binding.

Mokanszki 2014 presented a study of infertile men in which the proportion of HA binding determined treatment selection for the most part (PICSI if HA binding ≤60%), supplemented with cases undergoing ICSI because PICSI was contra-indicated and eight cases who were selected to undergo PICSI. It was not possible to determine numerators or denominators from reported percentages, which did not appear to be based on numbers of either women or transfers. Even if numbers were available, it is unclear how useful these results would be for the comparison of interest here.

Lohinova 2017 presented a small controlled trial of PICSI methods - 'SpermSlow' versus 'PICSI cup' - for infertile men with previous IVF failure. There was a high risk of bias with no claim of randomisation and no information regarding blinding. It was not possible to derive numerators or denominators from presented graphs of clinical outcome. Results appeared very similar with the two methods.

Erberelli 2017 reported outcomes of PICSI and ICSI for couples with 'moderate to severe' male factor. There was high risk of bias with no suggestion of randomisation or blinding. It was also unclear whether the 56 cycles reported were for 56 couples or included repeat cycles. Cycles using PICSI had twice the average number of oocytes (12 vs 6) and higher clinical pregnancy rate in this small (n=56) study. Later clinical outcomes, including miscarriage and live birth, were not reported.

Korosi 2017 reported a comparison of pre-treatment with oral supplement for subfertile men scheduled for PICSI. The pre-treated participants also had their semen incubated for 2 hours in Myo-Inositol immediately prior to selection. All participants received PICSI under the protocol. The methods state that data were excluded from analyses for men non-adherent with study medication, which clearly breaches the intention-to-treat principle. It is not explained why the active arm

remained substantially larger than the control arm. They reported no clinical pregnancies in the 13 control couples and 11 (50%) in the active arm. Of these, two miscarried, four were ongoing at the time of report and five had led to live births.

Avalos-Duran 2018 undertook a systematic review and meta-analysis of trials comparing PICSI with ICSI for infertile men in terms of live birth, miscarriage and other outcomes. They identified two small trials (Parmegiani 2010 and Castillo-Baso 2012). Neither is reviewed in this exercise. Both were of unclear risk of bias regarding allocation process and at high risk of bias for other aspects. The reviewers found no evidence or suggestion of effect for either miscarriage or live birth rates.

Hasanen 2020 randomised 413 couples on the day of autologous ICSI to selection using either PICSI or MACS. All couples had sperm DNA fragmentation, at least one million progressive motile sperm, at least five mature oocytes and women aged 18 to 35 years. Unfortunately, 17 (6%) participants were excluded post-randomisation for not having met eligibility criteria and a further 59 (14%) were omitted from presented analyses having vitrified all available embryos. The mean number of embryos per transfer was 2.3 in each arm of the trial. Ongoing pregnancy from fresh transfer was similar between arms: OR=1.1 (0.72 to 1.6). I was unable to calculate either the number of clinical pregnancies or number of miscarriages from the presented data.

Hozyen 2022 also recruited couples with sperm DNA fragmentation, at least one million progressive motile sperm, at least five COCs and women aged <37 years from the same clinic as Hasanen 2022 during the same period of time. They additionally specified the requirement to have "at least one mature oocyte developed to a blastocyst with fresh embryo transfer", although it is unclear how this could be known at the time of randomisation. They reported a four-group randomised trial comparing sperm preparation methods including PICSI alongside density gradient centrifugation (DGC), testicular sperm and MACS. PICSI had the highest clinical pregnancy and ongoing pregnancy rates. There was no description of an adequate concealment for the randomisation process. The comparison of PICSI with DGC suggested higher ongoing pregnancy rate, OR=2.0 (1.0 to 3.8) and similar miscarriage rate, OR=1.2 (0.3 to 4.5).

Recommendation: GREY for all outcomes [Only 1 moderate/high quality study, no safety concern] N.B. Miller 2019 comprised 95% participants with male factor. Please see note above for general population.

11 (iii) *Older women*

Miller 2019 (see 11 i) presented pre-planned subgroup analysis of their primary outcomes by maternal age, including a cohort of 1331 women aged at least 35 years. These data give OR=1.3 (0.97 to 1.7) for term live birth and OR=0.48 (0.30 to 0.75) for miscarriage.

Recommendation: GREY for all outcomes [Only 1 moderate/high quality study, no safety concern]. Given consistency with general population, the committee could consider grading GREEN for miscarriage. The potential effect on live birth is less clear given the larger estimate and much wider confidence intervals.

12. Steroids (glucocorticoids)

The previous review considered four RCTs and a further controlled trial that were each at risk of bias but consistently supported the use of steroids. These are included below alongside additional studies, categorised as requested with the additional consideration of miscarriage rates.

12 (i) *General population*

Fawzy 2008 randomised 180 women with at least three previous, unexplained, first or second trimester miscarriages. This was a three-arm study in which they allocated participants in early pregnancy to combined treatment (oral prednisone and progesterone for 12 weeks of gestation plus aspirin for 32 weeks), enoxaparin (20mg per day, subcutaneous injection through to term) or placebo. The authors do not report on the allocation method. They claim to have blinded participants to allocation but there is no description of how they achieved this given the different modes of delivery for different durations. They also excluded ten (6%) participants from analysis. They reported that live birth in steroid-treated participants was similar to those treated with enoxaparin [OR=1.3 (0.49 to 3.7)] and much higher than those treated with placebo: OR=6.1 (2.4 to 16). Miscarriage rates are the inverse of these figures.

Fawzy 2013 studied over 300 women with previous unexplained implantation failures. The intervention consisted of oral prednisolone 20 mg/day from the day of stimulation with 1mg/kg/day subcutaneous low molecular weight heparin (LMWH) from the day after oocyte retrieval until the day of pregnancy test (if negative) or week 8 of pregnancy. The authors reported a large increase in ongoing pregnancy but this study was unblinded and, more importantly, used entirely predictable alternation rather than randomisation to allocate participants. Results are therefore unreliable. A large benefit in terms of clinical and ongoing pregnancy rates of intervention was claimed with similar miscarriage rates.

Gomaa 2014 allocated 160 women with previous, unexplained, recurrent miscarriage. The intervention consisted of 5mg/day prednisolone in addition to the low dose aspirin and unfractionated heparin received by all participants. Recruitment occurred before seven weeks of gestation. This was a well-designed study describing concealed randomisation and placebo-based double-blinding. They excluded ten (6%) participants from analysis due to loss from follow-up. The reported effect size was extreme. Substantially more women in the intervention arm had pregnancy ongoing at 20 weeks: OR=23 (9.3 to 59). Miscarriage was the inverse of this figure.

Taiyeb 2017 studied 240 men with anti-sperm antibodies. Treatment consisted of following a course of tapering prednisolone repeated in each of three menstrual cycles prior to IVF/ICSI. There was risk of bias from both unclear allocation concealment and blinding processes and methodological issues with post-randomisation exclusions. Reconstruction of an intention to treat comparison suggested a small and non-statistically significant advantage of treatment on clinical pregnancy rate. Miscarriage rates were not reported.

Yeganeh 2017 studied over 200 women with PCOS with the aim of reducing the risk of OHSS. Intervention consisted of methylprednisolone: 1g intravenous on the days of oocyte retrieval and embryo transfer plus 16mg oral daily from the first day of stimulation through to pregnancy testing. This was another unblinded study at high risk of bias regarding allocation concealment but reported very similar clinical pregnancy rate in each group: OR= 1.2 (0.53 to 2.9). Miscarriage rates were not reported.

Kaye 2017 retrospectively analysed 876 embryo transfer procedures before and after a change in their routine practice. The earlier cohort had received prophylactic antibiotic and steroid for four days preceding the transfer. No medication was received by the later cohort. Patients from the earlier cohort were more likely to receive fresh transfer, less likely to be at blastocyst stage and, on average, received more embryos per transfer. Note that these are also 'improper' cohorts as they are defined from undergoing transfer rather than from initiation of treatment. Live birth rates were

similar: OR= 0.95 (0.73 to 1.2). Miscarriage rates were lower in the treated cohort: OR= 0.68 (0.44 to 1.0).

Milardi 2017 undertook a study of 90 men with oligozoospermia and evidence of abacterial prostate-vesiculo-epididymitis. They randomised participants to one of three doses of daily prednisone given for 1 month: 5; 12.5; 25mg. No clinical outcomes were reported with the focus on sperm parameters. These improved to some extent in the anticipated direction in all three groups. Unfortunately, the analyses were within-group rather than comparative but there was some evidence of a dose-response relationship.

Siristatidis 2018 initiated a randomised trial in patients with recurrent implantation failure defined as at least two failed transfers each of at least two good quality embryos. Unfortunately, they found randomisation to be impractical "early after the initiation" of the study. It is not clear exactly why this was the case nor whether and, if so, how recruitment continued after this point. The final data suggested higher live birth with almost identical miscarriage rates: OR=1.0 (0.14 to 7.5). It is worth noting the similar recruitment period, eligibility criteria and design difficulties to the study by the same first author reviewed under 'endometrial scratch' above.

Liu 2018 undertook an unblinded trial of 450 women undergoing their first IVF cycle with no history of recurrent miscarriage who experienced raised progesterone levels on the third or fourth day of gonadotrophin stimulation. They compared 0.75mg daily oral dexamethasone with no treatment in another unblinded study. They reported very similar live birth rates in the fresh transfer cycle: OR=1.1 (0.72 to 1.5). Miscarriage rates were also similar: OR=0.85 (0.40 to 1.8). Follow-up for two years of all frozen transfers suggested a possible advantage of intervention for the outcome of cumulative live birth: OR= 1.5 (1.0 to 2.2).

Thalluri 2022 reported a retrospective study of live births resulting from IVF/ICSI cycles. They identified 618 mothers who had received oral corticosteroids (prednisolone or dexamethasone) either during the cycle or within the first trimester. Typical indications for such treatment were recurrent implantation failure or recurrent miscarriage of presumed immune aetiology. This design does not allow for consideration of implantation or pregnancy outcomes such as miscarriage. The focus was on congenital anomalies some of which were reported to be higher in the treated group. However, the authors acknowledge that it was not possible to control for the characteristics that led to the clinical decision to treat with corticosteroids. It is therefore valuable to note the numbers of specific anomalies but not possible to distinguish to what extent these may have been a result of modifiable clinical factors such as steroid treatment.

Current rating [?].
Recommendation: GREY for all outcomes. [Insufficient evidence from moderate/high quality studies, no safety concerns].

12 (ii) *Populations with immunological testing*

Turi 2010 studied 48 women with anti-thyroid autoimmunity. They allocated participants to a tapering dose of prednisone in the month preceding ovarian stimulation or matching placebo. This appears to have been a well-conducted study with allocation concealment and blinding but was far too small to reach conclusions. Trial intervention completed before pregnancy. There were more pregnancies in the intervention arm. This allowed similar live birth rate [OR=2.1 (0.18 to 25)] despite higher miscarriage rate [OR=7.7 (0.85 to 70)] in the intervention arm, each with very wide confidence intervals.

Tang 2013 studiend 40 women with high urine natural killer cell density. They allocated participants to prednisolone or matching placebo for a total of eight weeks from study entry at around 4 to 6 weeks of gestation. This was well-conducted but was designed as a feasibility trial with small numbers and correspondingly wide confidence intervals. Trial intervention was during the first trimester of pregnancy. Live birth was higher in the intervention arm: OR=2.3 (0.63 to 8.0). Miscarriage rate was the inverse of this.

Fan 2016 studied 130 women undergoing IVF with antinuclear antibody who had experienced a previous implantation failure. Treatment consisted of prednisolone 10mg daily plus aspirin 100mg daily from 3 months before ovulation induction until clinical pregnancy. The trial was unblinded and unclear regarding allocation concealment. Results are therefore not reliable. A large benefit in terms of ongoing pregnancy was reported: OR=3.9 (1.8 to 8.5). A large benefit in terms of miscarriage was also reported: OR=0.43 (0.11 to 1.7).

Huang 2021 studied 19 women with recurrent implantation failure. They were all given prednisolone 10mg daily in the month preceding an intended natural cycle frozen embryo transfer. Treatment continued to the day of a negative pregnancy test or through to 12 weeks gestation. Four live births and one miscarriage were observed. The focus was on biomarkers of immune balance. Although not selected on the basis of immunological testing, these markers were shown to be worse at baseline than in a control group of fertile mothers and some markers improved by follow-up.

Gao 2021 undertook a retrospective study of 80 women under a freeze-all protocol who were positive for anti-nuclear antibody. Fifty had received a combined treatment of prednisone and hydrochloroquine from day 3 of the frozen transfer cycle through until the twelfth week of gestation. There was no indication as to why some patients had received the treatment and others not. There was also no comparison of these two groups at baseline. They reported a marked difference in clinical pregnancy rate [OR= 14 (4.6 to 43)]. Similar numbers of these pregnancies miscarried: OR=1.1 (0.33 to 3.6).

Zhou 2022 studied 346 women who underwent a first cycle of IVF/ICSI who were euthyroid but had tested positive for anti-thyroperoxidase or thyroglobulin antibodies. This was a retrospective study of those who had or had not received combined prednisone and aspirin treatment from the day of transfer until confirmation of pregnancy according to clinician inclination. Clinical pregnancy was slightly higher in the treated arm, but livebirth was lower: OR=0.91 (0.59 to 1.4). This was a result of higher miscarriage in the intervention arm: OR=2.0 (1.0 to 3.8). These figures refer to unadjusted effect measures but this was not a randomised study. In multifactorial analyses of clinical pregnancy and miscarriage, stratified by fresh/frozen transfer status, adjusted effect estimates were very similar.

Recommendation: GREY for all outcomes. [Insufficient evidence from moderate/high quality studies, no safety concerns].

13. Time lapse

Time lapse incubation involves two distinct processes both hypothesised to deliver clinical benefits. First, the ability to leave the embryo undisturbed during repeated assessment may be beneficial to the development process. Independently, the additional information available through time-lapse imaging may bring benefits for embryo selection. The previous review in 2021 identified studies in three broad categories evaluating effects of:
i) the environment for embryo development (one safety study and one ongoing RCT);
ii) the embryo selection process (two low quality studies reported non-significant benefits); and

iii) the combined effect of the two (4 studies at high risk of bias with contrasting results).

13 (i). *Studies of the environment*

The previous review contained just a single safety study of this question that contributed no clinical outcomes. The current review includes two new RCTs.

Van Blerkom 2014 undertook a within-sibling oocyte study of a closed "simplified culture system". They recruited 40 couples undergoing a regular IVF cycle with planned single embryo transfer who produced at least eight cumulus-oocyte complexes. These were then divided between routine and closed systems. After day 3, a blinded embryologist selected the best for transfer or one was selected at random in the case of ties. In 23 of 35 transfers, an embryo from the closed system was chosen but it is unclear how often this was decided at random. The design does not allow comparison of clinical outcomes.

Park 2015 randomised over 350 couples in a 2:1 ratio. Their focus was on embryo quality but they also reported clinical outcomes with more than 95% single embryo transfers. This was a well-designed study. They reported lower ongoing pregnancy rate with the stable environment of the time-lapse incubator [OR=0.64 (0.38 to 1.1). They also reported similar clinical pregnancy and higher miscarriage rates. The authors note that their use of day 2 transfer may have led to atypical results but the study appeared reasonable from a methodological perspective.

Wu 2016 reported both a small pilot RCT of couples (n=49) and an even smaller study (n=7) in which oocytes/embryos were alternately assigned to the time-lapse or standard incubator. Neither was methodologically strong and the pilot RCT in particular suffered from substantial post-randomisation loss to follow-up. Neither study supported the use of the time lapse system.

Barberet 2018 randomised 386 couples undergoing IVF who had at least six mature oocytes. They compared use of the Embryoscope incubator with a conventional desk-top incubator (K-Systems). This was a well-designed study in terms of concealed randomisation but did not employ blinding of embryo assessment or others. Ongoing pregnancy rate beyond week 12 was similar in the two arms: OR=1.1 (0.73 to 1.7).

Guo 2022 conducted a 3-arm trial randomising a total of 600 participants to conventional incubation, EmbryoScope using conventional morphological assessment, or Embryoscope employing the morphokinetic embryo selection using the KIDScore Day 3 algorithm. They describe a secure allocation process but did not employ blinding. Substantial bias may have been introduced by post-randomisation exclusion of 29% of participants for reasons that included developmental issues of the embryo. Figures presented in each comparison here assume failure for all unreported outcomes but no imputation would be secure. Those in the time-lapse incubator using morphological selection criteria had more live births than those in the conventional incubator: OR=1.8 (1.2 to 2.6).

13 (ii). Studies of the selection process

The previous review contained two small RCTs that randomised couples to use a selection algorithm based on time lapse data or conventional morphology. Each suggested promise of the intervention but was subject to high risk of bias. This review incorporates two additional RCTs of question (ii) above.

Kaser 2017 reported a 3-way comparison of single embryo transfer based on Eeva classification on either day 3 or day 5 versus conventional morphology on Day 5. Highest clinical pregnancy, lowest

miscarriage and highest ongoing pregnancy rates were observed in the conventional arm. This was a pilot study (n=163) that the sponsor stopped prematurely due to "funding priorities" but appeared methodologically sound in other key regards. The estimated effect for ongoing pregnancy in the combined Eeva groups versus conventional morphology was OR=0.69 (0.36 to 1.4).

Ahlstrom 2022 also studied elective single embryo transfer. 676 patients with at least two good blastocysts on day 5 were randomised between selection based on KIDScore or conventional morphology (Gardner/Schoolcraft). The study stopped earlier than intended as a result of the global pandemic but appears otherwise strong methodologically. Clinical pregnancy rate was a little lower in the time lapse group [OR=0.95 (0.72 to 1.3)] with higher early pregnancy loss: OR=1.2 (0.75 to 1.8).

Guo 2022 (see above) also contained a comparison of the selection process. This was subject to the same risks of bias outlined above. Those in the morphokinetic selection arm had similar live births to those in the conventional morphological selection arm: OR= 0.87 (0.59 to 1.3).

13 (iii). Trials of environment and selection

The previous review contained four studies of the combined question, none of which was at low risk of bias. Results of the two largest studies were starkly contrasting, with claims of both significant detriment and significant benefit. This review incorporates two additional RCTs of question (iii) above.

Meng 2022 compared time lapse incubation with day 3 KIDScore versus conventional incubation and morphology in 139 couples. This appears to have been a well-designed study but stopped early, seemingly at a planned interim review, due to the magnitude of difference observed. Live birth was markedly lower in the time lapse arm: OR=0.38 (0.19 to 0.76). Reported miscarriage was very low in both groups (n=6 total).

Zhang 2022 compared time lapse incubation with 'Geri assess' versus conventional incubation and morphology (Alpha consensus) in over 1200 couples. The study design and conduct appears methodologically strong. Usually two embryos were transferred, which may affect generalisability to UK practice. Live birth rate was similar but slightly higher in the time lapse arm: OR=1.1 (0.85 to 1.4) and cumulative live birth even more similar. Patients were generally good prognosis (e.g. aged <35yrs, first cycle).

Guo 2022 (see above) also contained a comparison under this category, reusing the arms contributing to the earlier analyses. This was subject to the same risks of bias outlined above. Those in the Embryoscope with morphokinetic selection arm had more live births than those in the conventional incubation arm: OR=1.5 (1.0 to 2.3).

Current rating amber.
Recommendation: BLACK [5 moderate/high quality studies with consistent results]

DISCUSSION

Caution is required as the assessments above are made from a methodological perspective without expertise in the clinical or scientific context.

The recommendations for rating are only intended as a starting point for committee discussion.

Some comparisons contain a range of interventions (e.g. steroids taken by the male or female partner, before or during pregnancy). Many post-hoc but biologically plausible rationales could be put forward to 'lump' or further 'split' categories presented above.

REFERENCES: Reviewed studies (Bold indicates full references added for 2023 update)

Adjunct	Study	DOI/reference
Artificial Egg Activation		
General	**Nasr-Esfahani 2007**	**10.1016/j.fertnstert.2007.10.047**
	Borges 2009	**10.1016/j.fertnstert.2008.04.046**
	Liu 2011	**10.1017/S0967199411000530**
	Ebner 2012	**10.1016/j.fertnstert.2012.07.1134**
	Liu 2014	**10.1089/cell.2013.0081**
	Eftiikhar 2013	**IRCT2012112610328N1**
	Caglar Aytac 2015	10.1016/j.fertnstert.2015.07.1163
	Yang 2015	**NLM 101093592**
	Fawzy 2018	10.1093/humrep/dey258
	Shebl 2021	**10.1007/s10815-021-02338-3**
	Yin 2022	**10.1007/s00404-021-06329-8**
Failed fertilisation	Meerschaut 2012	10.1093/humrep/des097
	Montag 2012	**10.1016/j.rbmo.2012.02.002**
	Ebner 2015	**10.1016/j.rbmo.2014.11.012**
	Darwish 2015	**10.1016/j.rbmo.2015.08.012**
	Aydinuraz 2016	10.1080/14647273.2016.1240374
	Hao 2016	**10.3760/cma.j.issn.0376-2491.2016.43.010**
	Li 2019	**10.1016/j.rbmo.2019.03.216**
Assisted Hatching: Stored	Balaban 2006	10.1093/humrep/del097
	Ge 2008froz	RBMO 2008;16(4):589-96.
	Valojerdi 2010	10.1016/j.rbmo.2009.11.002
	Fang 2010	10.1016/j.fertnstert.2009.08.014
	Debrock 2011	**10.1093/humrep/der161**
	Figueria 2012	**10.1016/j.ejogrb.2012.05.022**
	Ren 2013	**10.1007/s10815-013-9984-2**
	Wan 2014	**10.1016/j.rbmo.2014.01.006**
	Wang 2016	10.3892/br.2016.716
	Knudtson 2016	F&S 2016;106(3) Suppl:e141
	Safari 2017	10.1016/j.repbio.2017.05.003
	Elnahas 2017	**10.1016/j.mefs.2017.05.006**
	Kirienko 2019	**10.1016/j.rbmo.2019.06.003**
Fresh	Sagoskin 2007	10.1016/j.fertnstert.2006.07.1498
	Ge 2008fresh	RBMO 2008;16(4):589-96.
	Balakier 2009	10.1016/j.fertnstert.2008.07.1729
	Hagemann 2010	10.1016/j.fertnstert.2009.01.116
	Kutlu 2010young	10.1007/s10815-010-9431-6
	Kutlu 2010old	10.1007/s10815-010-9431-6
	Razi 2013	Iran J reprod Med 2013;11(12):1021-6.
	González-Ortega 2015	**Ginecol Obstet Mex 2015;83:670-9.**
	Shi 2016	10.1177/1933719116641764
	Chang 2016	F&S 2016;106(3) Suppl:e314
	Nada 2018	10.1007/s00404-017-4604-5
	Abulsoud 2019	**Int J Pharm Rev Res 2019;56(1):112-6**
	Fawzy 2020	**10.1093/humrep/deaa160**
	Zhang 2022	**10.3389/fendo.2022.927834**
Embryo Glue	Morbeck 2007	NCT005882250
	Mahani 2007	EMHJ 2007;13(4):876-80.
	Friedler 2007	10.1093/humrep/dem220
	Korosec 2007	RBM0 2007;15(6):701-7.
	Hazlett 2008	10.1016/j.fertnstert.2007.05.063

	Urman 2008	10.1016/j.fertnstert.2007.07.1294
	Dittmann-Muller 2009	Hum Reprod 2009;24 Suppl 1:167.
	Drew 2014	**10.1002/central/CN-01369787/full**
	Fancsovits 2015	10.1007/s00404-014-3541-9
	Singh 2015	10.4103/0974-1208.170398
	Kleijkers 2016	**10.1093/humrep/dew156**
	Zbořilová 2018	https://europepmc.org/abstract/med/30764616
	Kandari 2019	**10.1016/j.fertnstert.2021.02.015**
	Yung 2021	**10.1016/j.fertnstert.2021.02.015**
	Sellers 2022	**RBMO 2022;45(S1);e47**
Endometrial Receptivity	Simón 2020	10.1016/j.rbmo.2020.06.002
	Cohen 2020	**10.1080/19396368.2020.1824032**
	Cozzolino 2020	**10.1007/s10815-020-01948-7**
	Cozzolino 2022	**10.1016/j.fertnstert.2022.07.007**
	Doyle 2022	**10.1001/jama.2022.20438**
Endometrial Scratching	Raziel 2007	10.1016/j.fertnstert.2006.05.062
General	**Zhou 2008**	**10.1016/j.fertnstert.2007.05.064**
	Karimzadeh 2009	10.1111/j.1479-828X.2009.01076
	Narvekar 2010	10.4103/0974-1208.63116
	Abdelhamid 2012	10.1007/s00404-013-2785-0
	Nastri2013	**10.1002/uog.12539**
	Gibreel 2013	10.1111/j.1447-0756.2012.02016.x
	Parsanezhad 2013	IRCT:2012082510657NI
	Zarei 2014	IRCT:2012070810210NI
	Wadhwa 2015	J Hum Reprod Sci 2015;8(3):151-8.
	El Khayat 2015	10.1016/j/ejogrb.2015.08.025
	Mahey 2015	10.1016/j.fertnstert.2015.07.1163
	Gibreel 2015	**10.3109/09513590.2014.994603**
	Maged 2016	10.1177/1933719115602776
	Bahaa Eldin 2016	**10.1177/1933719116638191**
	Goel 2017	10.1007/s10815-017-0949-8
	Mak 2017	10.1016/j.rbmo.2017.04.004
	Aleyamma 2017	10.1016/j.ejogrb.2017.05.005
	Helmy 2017	10.1002/ijgo.12178
	Senocak 2017	10.1016/j.jogoh.2017.09.003
	Ashrafi 2017	10.1111/jog.13401
	Maged 2018	10.1002/ijgo.12355
	Hilton 2019	**10.1007/s00404-019-05044-9**
	Eskew 2019	**10.1007/s10815-018-1356-5**
	Frantz 2019	10.1093/humrep/dey334
	Lensen 2019	10.1056/NEJMoa1808737
	Olesen 2019	10.1016/j.fertnstert.2019.08.010
	Mackens 2020	**10.1093/humrep/deaa018**
	Tang 2020	**10.1111/jog.14193**
	Berntsen 2020	10.1016/j.ejogrb.2020.06.034
	Ghuman 2020	10.1016/j.ejogrb.2020.08.010
	Rodriguez 2020	10.1007/s43032-020-00204-8
	van Hoogenhuijze 2021	10.1093/humrep/deaa268
	Metwally 2021	10.1093/humrep/deab041
	Yavangi 2021	10.18502/ijrm.v19i5.9255
	Glanville 2022	**10.1016/j.rbmo.2021.10.008**
	Izquierdo 2022	**10.1016/j.jogoh.2022.102335**
	Madhuri 2022	**10.1016/j.ejogrb.2021.10.028**
	Metwally 2022	**10.3310/JNZT9406**
	Wong 2022	**10.1016/j.fertnstert.2021.12.009**

Implantation Failure	**Baum 2012**	**10.3109/09513590.2011.650750**
	Zhang 2014	**10.1007/s00404-014-3382-6**
	Zhang 2015	**10.1007/s11655-014-1843-1**
	Bord 2015	**10.1007/s00404-015-3954-0**
	Siristatidis 2017	**10.1080/09513590.2016.1255325**
	Gürgan 2019	**10.1016/j.rbmo.2019.02.014**
	Tumanyan 2019	**10.1080/09513590.2019.1632085**
	Aghajanpour 2021	**10.1016/j.jri.2021.103426**
Freeze All: General	Aflatoonian 2010	10.1007/s10815-010-9412-9
	Shapiro 2011a	10.1016/j.fertnstert.2011.05.050
	Shapiro 2011b	10.1016/j.fertnstert.2011.02.059
	Shapiro 2015	**10.1016/j.fertnstert.2015.07.1141**
	Magdi 2017	10.1016/j.fertnstert.2017.04.020
	Shi 2018	10.1056/NEJMoa1705334
	Le 2018	**10.1093/humrep/dey253**
	Vuong 2018	10.1056/NEJMoa1703768
	Vuong 2019	**10.1016/j.rbmo.2018.12.012**
	Wei 2019	10.1016/S0140-6736(18)32843-5
	Stormlund 2020	10.1136/bmj.m2519
	Simón 2020	10.1016/j.rbmo.2020.06.002
	Boynukalin 2020	**10.1371/journal.pone.0234481**
	Li 2021	**10.3389/fendo.2021.730059**
	Wong 2021	10.1093/humrep/deaa305
	Maheshwari 2022	**10.1093/humrep/deab279**
	Maheshwari 2022a	**10.3310/AEFU1104**
OHSS risk	**Chen 2016**	**10.1056/NEJMoa1513873**
	Rahav Koren 2018	**10.1159/000479557**
	Ye 2018	**10.1186/s12958-018-0373-7**
	Deng 2019	**10.1007/s11596-019-2031-5**
	Shrem 2019	**10.1016/j.rbmo.2019.04.014**
	Santos-Ribeiro 2020	10.1093/humrep/deaa226
	Deepika 2021	**10.5935/1518-0557.20200028**
	Huang 2021	**10.1038/s41598-021-02227-w**
	Vuong 2021	**10.1007/s10815-021-02180-7**
IMSI: General	**Balaban 2011**	**10.1016/j.rbmo.2010.11.003**
	Figueira 2011	**10.1016/j.fertnstert.2010.11.018**
	Setti 2012	**10.1016/j.rbmo.2012.01.007**
	Setti 2013	10.1016/j.ejogrb.2013.09.006
	Marci 2013	10.1186/1742-4755-10-16
	Cassuto 2014	10.1016/j.rbmo.2013.08.013
	Setti 2014	**10.1016/j.ejogrb.2014.10.008**
Male Factor	**Antinori 2008**	**10.1016/s1472-6483(10)60150-2**
	Knez 2011	**10.1186/1477-7827-9-123**
	Mahmoud 2011	**Hum Reprod 2011;26 Suppl 1:i181.**
	Setti 2011	**10.1016/j.fertnstert.2011.03.003**
	Knez 2012	**10.1016/j.rbmo.2012.03.011**
	Check 2013	**Clin Exp Obstet Gyn 2013;40(1):15-7.**
	De Vos 2013	10.1093/humrep/des435
	Leandri 2013	10.1111/j.2047-2927.2013.00104.x
	Kim 2014	10.5653/cerm.2014.41.1.9
	Sifer 2014	**10.1016/j.ejogrb.2014.07.017**
	La Sala 2015	10.1186/s12958-015-0096-y
	Mangoli 2019	10.1111/and.13340
	Mangoli 2020	**10.1007/s10815-020-01910-7**
Intralipids: General	**El-Khayat 2015**	**10.1016/j.fertnstert.2015.07.080**

	Gamaleldin 2018	**10.1002/central/CN-01911196/full**
	Singh 2019	10.1016/j.ejogrb.2019.06.007
	Al-Zebeidi 2019	10.1080/09513590.2019.1631280
Immunological testing	Dakhly 2016	10.1016/j.ijgo.2016.06.026
	Meng 2016	**10.1007/s00404-015-3922-8**
	Rogenhofer 2021	**10.1111/aji.13506**
Intrauterine culture	Blockeel 2009	10.1093/humrep/dep005
IV Immunoglobulin	Stephensen 2010	10.1093/humrep/deq179
General	Christiansen 2014	10.1111/1471-0528.13192
	Jørgensen 2020	**10.1016/j.jri.2020.103128**
Immunological testing	**Dendrinos 2009**	**10.1016/j.ijgo.2008.11.010**
	Cohen 2015	**PMID: 26380487**
	Yamada 2015	**10.1016/j.jri.2015.01.008**
	Lee 2016	**10.1111/aji.12442**
	Meng 2016	**10.1007/s00404-015-3922-8**
	Ahmadi 2017	**10.1016/j.imlet.2017.10.003**
MACS	Romany 2014	10.1016/j.fertnstert.2014.09.001
	Troya 2015	10.5935/1518-0557.20150015
	Romany 2017	10.1007/s10815-016-0838-6
	Ziarati 2018	10.1080/14647273.2018.1424354
PGT-A (Day 3)	Mastenbroek 2007	NEJM 2007;357:9-17.
	Hardarson 2008	10.1093/humrep/den217
	Staessen 2008	10.1093/humrep/den367
	Blockeel 2008	RBMO 2008;17(6):848-54.
	Meyer 2009	10.1016/j.fertnstert.2008.02.162
	Schoolcraft 2009	10.1016/j.fertnstert.2008.05.029
	Sher 2009	10.1016/j.fertnstert.2008.11.029
	Debrock 2010	10.1016/jfertnstert.2008.10.072
	Ikuma 2015	10.1371/journal.pone.0129958
	Rubio 2017	10.1016/j.fertnstert.2017.03.011
PGT-A (Day 5) General	Yang 2012	Molec Cytogen 2012;5:24
	Forman 2013	10.1016/j.fertnstert.2013.02.056
	Scott 2013	10.1016/j.fertnstert.2013.04.035
	Ozgur 2019	10.1007/s10815-018-01399-1
	Munné 2019	10.1016/j.fertnstert.2019.07.1346
	Cimadomo 2019	**10.1093/humrep/dez078**
	Yan 2021	**10.1056/NEJMoa2103613**
	De Munck 2022	**10.1371/journal.pone.0267241**
	Idárraga 2022	**10.5935/1518-0557.20210085**
Older women	**Ubaldi 2017**	**10.1016/j.fertnstert.2017.03.007**
	Verpoest 2018	**10.1093/humrep/dey262**
PICSI	**Parmegiani 2012**	**10.1016/j.fertnstert.2012.05.043**
	Worrilow 2013	10.1093/humrep/des417
	Majumdar 2013	10.1007/s10815-013-0108-9
	Mokanszki 2014	10.3109/19396368.2014.948102
	Troya 2015	10.5935/1518-0557.20150015
	Lohinova 2017	PMID: 29099693
	Erberelli 2017	10.5935/1518-0557.20170002
	Korosi 2017	**PMID: 28724183**
	Avalos-Durán 2018	**10.5935/1518-0557.20180027**
	Miller 2019	10.1016/S0140-6736(18)32989-1
	Hasanen 2020	**10.1007/s10815-020-01913-4**
	Novoselsky 2021	**10.1111/andr.12982**
	Hozyen 2022	**10.1007/s43032-021-00642-y**

Steroids (DHEA)	Wiser 2010	10.1093/humrep/deq220
	Kara 2014	10.1016/j.ejogrb.2013.11.008
	Yeung 2014	10.1016/j.fertnstert.2014.03.044
	Tartagni 2015a	PMID: 24867068
	Tartagni 2015	10.1186/s12958-015-0014-3
	Narkwichean 2017	10.1016/j.ejogrb.2017.09.006
	Wang 2022	10.1111/1471-0528.17045
Steroids (Glucocorticoids)	**Fawzy 2008**	**10.1007/s00404-007-0527-x**
General	Fawzy 2013	10.1007/s00404-013-3020-8
	Gomaa 2014	**10.1007/s00404-014-3262-0**
	Taiyeb 2017	10.1007/s12020-017-1446-7
	Yeganeh 2017	10.1080/01443615.2017.1346593
	Kaye 2017	**10.1016/j.fertnstert.2017.04.003**
	Milardi 2017	**10.1111/andr.12300**
	Siristatidis 2018	**10.1080/09513590.2017.1380182**
	Liu 2018	10.1111/cen.13824
	Thalluri 2022	**10.1093/humrep/deac142**
Immunological testing	**Turi 2010**	**10.1016/j.clinthera.2011.01.010**
	Tang 2013	**10.1093/humrep/det117**
	Fan 2016	10.1111/aji.12559
	Huang 2021	**10.1016/j.jri.2020.103245**
	Gao 2021	**10.1177/09612033211055816**
	Zhou 2022	**10.1186/s12884-022-04532-2**
Time Lapse (i)	Kirkegaard 2012	10.1007/s10815-012-9750-x
	Van Blerkom 2014	**10.1016/j.rbmo.2013.11.012**
	Park 2015	**10.1093/humrep/deu316**
	Wu 2016	**10.1186/s12958-016-0181-x**
	Barberet 2018	**10.1016/j.fertnstert.2017.10.008**
	Chen 2020	10.1093/humrep/deaa268
	Guo 2022	**10.3389/fphys.2021.794601**
(ii)	Goodman 2016	10.1016/j.fertnstert.2015.10.013
	Kaser 2017	**10.1093/humrep/dex231**
	Alhelou 2018	10.1016/j.repbio.2017.12.003
	Kovacs 2019	10.1016/j.ejogrb.2018.12.011
	Ahlstrom 2022	**10.1093/humrep/deac020**
(iii)	Kahraman 2013	10.1177/205891581200300204
	Rubio 2014	10.1016/j.fertnstert.2014.07.738
	Insua 2017	10.1016/j.fertnstert.2017.06.031
	Yang 2018	10.1093/humrep/dey047
	Meng 2022	**10.1016/j.fertnstert.2022.02.015**
	Zhang 2022	**10.1016/j.rbmo.2022.06.017**
	Guo 2022	**10.3389/fphys.2021.794601**
(Misc.)	Wang 2016	J Reprod Med 61(5):254-262