Adjuncts in the IVF laboratory: where is the evidence for ‘add-on’ interventions?

Running title: Adjuncts in IVF

Key words: IVF, adjuncts, sperm DNA fragmentation, embryo glue, adherence compounds, assisted hatching, PGS, mitochondria load, time lapse imaging, RCT, live birth rate

Joyce Harper¹, Emily Jackson², Karen Sermon³, Robert John Aitken⁴, Stephen Harbottle⁵, Edgar Mocanu⁶, Thorir Hardarson⁷, Raj Mathur⁸, Stephane Viville⁹, Andy Vail¹⁰, Kersti Lundin¹¹

¹Embryology, IVF and Reproductive Genetics, Institute for Women’s Health, University College London, UK

²Law Department, London School of Economics and Political Science, UK

³Research Group Reproduction and Genetics, Vrije Universiteit Brussel, Laarbeeklaan 101, 1090 Brussels, Belgium

⁴Priority Research Centre for Reproductive Science, University of Newcastle, NSW, Australia

⁵Cambridge IVF, Cambridge University Hospitals NHS Trust, Cambridge, UK

⁶Rotunda Hospital and RCSI, Parnell Square, Dublin 1, Ireland

⁷Fertilitetscentrum, Carlanders Hospital, 402 29, Gothenburg, Sweden

⁸Department of Reproductive Medicine, Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester M13 9WL, UK
Abstract

Globally, IVF patients are routinely offered and charged for a selection of adjunct treatments and tests or ‘add-ons’ that they are told may improve their chance of a live birth, despite there being no clinical evidence supporting the efficacy of the add on. Any new IVF technology claiming to improve live birth rates (LBR) should, in most cases, first be tested in an appropriate animal model, then in clinical trials, to ensure safety, and finally in a randomized controlled trial (RCT) to provide high quality evidence that the procedure is safe and effective. Only then should the technique be considered as ‘routine’ and only when applied to the similar patient population as those studied in the RCT. Even then, further paediatric and long-term follow up studies will need to be undertaken to examine the long-term safety of the procedure. Alarmingly, there are currently numerous examples where adjunct treatments are used in the absence of evidence-based medicine and often at an additional fee. In some cases, when RCTs have shown the technique to be ineffective, it is eventually withdrawn from the clinic. In this paper, we discuss some of the adjunct treatments currently being offered globally in IVF laboratories including embryo glue and adherence compounds, sperm DNA fragmentation, time lapse imaging, preimplantation genetic screening, mitochondria DNA load measurement and assisted hatching and examine the evidence for their safety and efficacy in increasing LBRs. We conclude that robust studies are needed to confirm the safety and efficacy of any adjunct treatment or test before they are offered routinely to IVF patients.
Introduction

IVF is a globally adopted technique supporting an extremely lucrative medical industry which has revolutionized human reproduction by offering hope of a family where none existed before. Patients routinely pay large sums of money for treatment and many are willing to try anything that might help them improve their chances of having a baby.

The vast majority of IVF clinics want to help their patients achieve this objective as much as possible which may involve undertaking unproven procedures and tests supported by anecdotal, low quality or unpublished evidence. In the last decade, a plethora of adjuncts or ‘add-ons’ have been introduced, many without any robust evidence that they increase the chances of a live birth or have any tangible benefit in terms of the health and wellbeing of the offspring (Nardo et al., 2009, Datta et al., 2015, Spencer et al., 2016, Harper and Brison, 2013).

The requirement that patients give informed consent to adjunct treatment in IVF, whilst necessary, may be insufficient to eliminate the over-selling or mis-selling of adjunct therapies for which the evidence of efficacy is poor or non-existent. In the UK, the Human Fertilisation and Embryology (HFEA) Act 1990, as amended, requires patients to ‘be provided with such relevant information as is proper’ before embarking on treatment (Human Fertilisation and Embryology Act 1990, sch. 3, para. 3(1)(b)). Patients should also be provided with ‘a personalised cost treatment plan’ (HFEA Code of Practice, para 4.3). Before an adjunct treatment is offered, the legislation requires clinics to provide open and honest information about the existence of robust evidence to support the particular intervention, along with information about costs. However, the ‘therapeutic illusion’ (Casarett, 2016), which commonly involves ‘unjustified enthusiasm for treatment on the part of both patients and doctors’ (Thomas, 1978), may mean that patients are not necessarily put off by low success rates or underpowered trial data, especially when simplistic explanations for reproductive failure circulate online and in the popular press. For example, a clinician might explain that the studies of immune therapy in assisted reproduction treatments to date have been poorly designed and that larger RCTs are necessary (Nardo et al, 2015), while patients might read newspaper articles with headlines such as ‘The killer cells that robbed me of four babies’ (Barber, 2011).
All IVF clinics need to consider the safety and efficacy of new technologies before introducing them and beginning to charge patients. In most cases, this should include preliminary work on animal models, followed by studies on human embryos donated for research and finally well designed RCTs with a follow up of all children born from the procedure (Harper et al., 2012). If such preliminary studies are not published, it is possible that technology bringing no clinical benefit or even leading to adverse health outcomes may be introduced.

There are several key factors affecting the validity and usefulness of any RCT performed for IVF. Validity can be assessed through risk of bias (Higgins et al., 2011) whereas the usefulness depends on the definition of the patient cohorts, the interventions compared, the primary outcome and the number of participants. Typically, demonstrating a clinical benefit will require many more participants than that required to demonstrate physiological effects.

Wilkinson et al (2016) analysed 142 IVF RCTs published in 2013 and 2014. They found that no consistent outcome measure was used. They suggest that initiatives to standardize outcome such as LBR or cumulative LBR should be encouraged. Trials using implantation rate or clinical pregnancy rate (CPR) as outcome measures are only appropriate for preliminary studies. After any technique is brought into routine clinical practice, follow up longitudinal studies should be undertaken to ensure the safety and efficacy of the intervention.

In 2009, the Policy and Practice Committee of the British Fertility Society reported on medical adjuncts in IVF and concluded that “there is a need for good clinical trials in many of the areas surrounding medical adjuncts in IVF to resolve the empirical/evidence divide” (Nardo et al., 2009). Datta et al. (2015) reported on clinical and laboratory adjuncts and tests in IVF and stated that properly powered RCTs are more valuable than a meta-analysis of a number of small heterogeneous RCTs. Spencer et al. (2016) carried out an audit of UK based IVF clinic web sites and found that many were offering patients a large number of unproven adjuncts at additional cost.

In this paper we describe some of the adjunct IVF laboratory treatments and tests that are currently being offered globally (Table 1), often at a substantial cost for the
patient. We describe the techniques and discuss the evidence for their safety and whether they increase LBR. The majority of the adjunct treatments listed here are included in the HFEA’s recent addition to their website where they list evidence for ‘add-on’ treatments (due to go live in February 2017). Laboratory adjuncts that we could have included but were not due to space limitations are; intracytoplasmic sperm injection (ICSI) for non male infertility patients (Grimstad et al., 2016), intracytoplasmic morphologically selected sperm injection (IMSI), in-vitro maturation, artificial oocyte activation, augmentation of mitochondria, intrauterine culture and elective freeze all embryos strategies.

**Embryo glue and adherence compounds**

The use of fibrin sealants to reduce ectopic pregnancy rate and increase LBRs was first proposed by Feichtinger (1990) and the same author published further supportive data in 1992. Despite this early promise, treatment using fibrin sealants never demonstrated reliable significant improvement in clinical outcomes and more recently, the focus has shifted to the use of a specific embryo transfer medium enriched with the glycoprotein hyaluronan (HA). It is well reported that HA is naturally present in the female reproductive tract and endometrium and forms a viscous solution which could enhance the embryo transfer (ET) process and prohibit embryo expulsion (Bontekoe et al., 2014).

The published data surrounding the use of adherence compounds are highly varied in quality and robustness of study design and as a result, the use of HA supplemented media for ET is still regarded as controversial (Bontekoe et al., 2014).

The latest Cochrane review of 3898 participants from 17 RCTs demonstrated moderate quality evidence for an improvement in CPR and LBR, with an associated increase in multiple pregnancy rate, when transfer medium was supplemented with HA (Bontekoe et al., 2014). The authors concluded that further high-quality studies were required, in particular where an elective single embryo transfer (eSET) procedure was performed, in part to alleviate concerns over the reported increase in the multiple pregnancy rate. A more recent RCT by Fancsovits et al., (2015) looked at 581 cycles and did not show a benefit in implantation rate, CPR or LBR, but found a higher birth weight in the HA group.
The reported increase in multiple pregnancy rate is suggestive of a need for clinics considering the use of a HA supplemented ET medium not only to re-evaluate their eSET policy and closely monitor their multiple pregnancy rate but also to ensure that patients are aware, not only of the possible increased chance of pregnancy, but also of the increased chance of multiple pregnancy when they are considering the number of embryos they wish to transfer.

The published evidence may be suggestive of a beneficial effect of the use of HA supplemented ET media. Before robust conclusions can be drawn, however, further RCTs are needed to evaluate the efficacy of HA as an adherence compound during ET with respect to eSET and the possibility to reduce the multiple pregnancy rate.

**Sperm DNA fragmentation**

Many clinics offer all their patients a sperm DNA fragmentation test. The assays include TUNEL (Terminal deoxynucleotidyl transferase dUTP Nick-End Labeling,) Comet, SCD (Sperm Chromatin Dispersal) assay, SCSA (Sperm Chromatin Structure Assay) and 8OHdG test (Shamsi et al., 2011). There are clear differences between assays in terms of the type of DNA damage being measured and their relative sensitivity (Smith et al., 2013). However, no particular assay has yet emerged as being of greater diagnostic value than any other. Ultimately the purpose of such an assay is to indicate which treatments may be contra-indicated for, or beneficial to, patients. This requires both diagnostic accuracy for the assay and evidence of effectiveness for the treatment(s). If for example the purpose of the assay is to determine whether antioxidant therapy is appropriate for the male partner then the measurement of 8OHdG is of paramount importance and robust assays to assess this base adduct need to be developed and optimized (Muratori et al., 2015).

Three recent meta-analyses looked at measuring sperm DNA fragmentation in patients undergoing IVF and ICSI. Osman et al. (2015) performed a meta-analysis of six studies and found that, overall, men with low sperm DNA fragmentation had a higher LBR than those with high DNA fragmentation, but that the evidence was not sufficient to support this when ICSI was used. They concluded that further RCTs are needed to examine the role of ICSI versus IVF for men with high DNA fragmentation. Simon et al., (2015) looked at 8068 treatment cycles where DNA damage was
measured using all four assays and found a modest but statistically significant association of DNA damage with CPR following IVF and/or ICSI. They found that the data varied depending on the assay used. Cissen et al (2016) performed a systematic review and meta-analysis looking at the prognostic value of sperm DNA damage measurement, including 30 out of 658 studies. They concluded that current tests have limited capacity to predict either the chance of conception after ART or which treatment method to choose, and that for now there is insufficient evidence to recommend sperm DNA testing.

The Practice Committee of the ASRM has concluded that ‘current methods for assessing sperm DNA integrity do not reliably predict treatment outcomes and cannot be recommended routinely for clinical use’ (Pfeifer et al., 2014).

However a recent Cochrane report observed that low quality evidence suggests that antioxidant therapy in the male might increase CPR and LBR in patients where the spermatozoa are suffering from oxidative stress (Showell et al., 2014). In this context, accurate assessment of 8OHdG levels could be of value in selecting a valid patient population. An RCT investigating the hypothesis that antioxidants can reverse oxidative DNA damage in spermatozoa is therefore urgently needed to address this possibility.

**Time lapse imaging**

Taking pictures over time and reviewing them as a film, also known as time-lapse imaging (TL), is a technique that has been used for a century. Indeed, the first time TL imaging was reported as a tool to visualize early embryonic development was in 1929 (Lewis and Gregory, 1929). In that report, a remarkably detailed description of hamster embryonic development was described and the authors went on to speculate whether the observed timings in cleavage rate could predict "embryonic potential". More than 50 years later, human embryos were filmed using TL technology during their first three days of development (Eriksson et al. 1981). The next significant breakthrough was the work by Payne et al. (1997) who used TL imaging to describe the first events during fertilization, thus providing insight into how diverse and dynamic early embryonic development can be.
The first attempt to meaningfully use the unique information from different embryo cleavage timings and/or cleavage patterns was performed by Meseguer and colleagues in 2011 based on data from 247 embryos known to have implanted (Meseguer et al., 2011). The latest prediction model was published by Petersen et al. (2016) but still requires extensive prospective testing and validation.

The usefulness of TL imaging in human IVF has been well debated. Among the proposed benefits that have been put forward are "not missing important events during culture", quality control, teaching applications, more information to the patient and, of course, an increase in LBR.

Rubio et al. (2014) conducted the largest RCT to date that included 843 patients randomized mainly on day 3 but also on day 5. They reported a 9.7% increase in CPR compared to traditional culture and morphology assessments alone. This effect was diluted in the Cochrane review that also included two smaller trials under the intention to treat principle. The authors concluded that “there is insufficient evidence of differences in live birth, miscarriage, stillbirth or clinical pregnancy to choose between [TL imaging] and conventional incubation” (Armstrong et al., 2015). However, more refined models are being continually developed as more data are being collected world-wide.

TL imaging serves so many other functions in the laboratory that its introduction will not be held back. It may be unthinkable in 5-10 years to still only be observing embryos by manually taking them out and looking at them. TL imaging is a tool which confers a number of practical benefits to the IVF laboratory. The future challenge for TL imaging is to find the best role in the IVF laboratory and to reduce implementation and consumable costs.

More RCTs are needed to distinguish whether there are clinical benefits of embryo selection algorithms based on TL information leading to an increase in LBR and whether there are benefits from uninterrupted embryo culture (Armstrong et al., 2014).

**Pre-implantation genetic screening**
When in the 1990s, several studies demonstrated that cleavage stage embryos showed a high level of aneuploidy (Coonen et al. 1994, Munne et al., 1995), it was postulated that selection against these aneuploid embryos would improve LBRs. It was surprising that not only were meiotic abnormalities originating in the oocyte found, but also abnormalities occurring postzygotically. As a consequence, many embryos were mosaics, containing both normal and aneuploid cells, or several different lines of aneuploid cells. Thousands of IVF cycles were performed with preimplantation genetic screening (PGS), by biopsying one cell at day 3 and performing fluorescent in-situ hybridization (FISH) for five chromosomes. Eleven RCTs later, PGS was shown not to increase CPR or LBR and, in some cases, to decrease LBR (Harper et al., 2010, Geraedts and Sermon, 2016). It was realized that PGS at day 3 was not effective because of the limited accuracy of FISH, the limited number of cells available for biopsy, and because at day 3, cleavage stage embryos are at a peak of chromosomal abnormality/mosaicism.

With the advent of new technology allowing comprehensive chromosome screening (CCS) of day 5 biopsied trophectoderm cells, PGS is now actively marketed as increasing implantation rates, and consequently decreasing time to pregnancy, recurrent miscarriages and repeated implantation failure (Sermon et al., 2016).

Despite these claims, only three RCTs have been published, all of which have been criticized because of poor study design. The pilot RCT by Yang et al. (2012) included a small sample size of 45 young, good prognosis patients. Scott et al., (2013) performed an RCT on 72 good prognosis patients between the ages of 21 and 42 years who were randomized quite late, ie. if they had at least two blastocysts available for analysis. Although the authors claimed that PGS increased implantation and delivery rates, there was a fundamental methodological flaw in the study’s failure to account for the difference between the unit of randomization (patients) and unit of analysis (individual embryos). The third RCT studied 89 patients aiming to compare PGS and SET with the transfer of two embryos (Forman et al., 2013). The same methodological problem encountered by the Scott trial was introduced and even so, the wide confidence interval for pregnancy did not demonstrate a beneficial effect.
Currently, two larger RCTs are underway and the results are expected soon. The ESTEEM study recruits patients of advanced maternal age and includes analyses of polar bodies using array-CGH, while the STAR study recruits all IVF patients and uses next generation sequencing on blastocyst biopsies. Other noteworthy differences are that the ESTEEM study has an intention-to-treat analysis, while STAR includes patients with two analyzable blastocysts as in the Scott and Forman studies. Furthermore, the ESTEEM outcome is cumulative LBR, while for STAR it is ongoing pregnancy rate after one transfer, an outcome measure that has received much criticism and should be abandoned in favour of LBR (Gries, 2016).

Although these studies may serve to provide stronger evidence supporting PGS, the current RCTs do not provide sufficiently robust evidence to consider PGS as a proven and beneficial treatment.

**Mitochondrial DNA load measurement**

It has been estimated that metaphase II oocytes contain about $10^5$ mitochondrial DNA (mtDNA) copies, but since no replication of the mtDNA occurs until the blastocyst stage of embryonic development, the mtDNA molecules are divided over the cleaving cells (Fragouli and Wells, 2015). In 2015, two papers were published reporting an association between higher mtDNA level and lower implantation potential in blastocysts (Diez-Juan et al., 2015; Fragouli et al., 2015), pointing to disturbed energy provision and metabolic stress in embryos with a higher mtDNA content. While the paper of Diez-Juan et al. focused on euploid, transferred blastocysts, the other report also showed a relationship between aneuploidy of the blastocyst and a higher mtDNA load. According to both reports, euploid embryos that implanted after transfer had a mtDNA load below a data-derived threshold. Conversely, embryos that failed to implant, or that were aneuploid, showed a wide range of mtDNA load. This range overlapped with the implanting embryos at the low end, but the level of mtDNA at the high end was much higher in the non-implanting embryos. A threshold embryonic mtDNA load above which all embryos failed to implant could therefore be identified. Diez-Juan et al reported that 52% (34/65) of the embryos below the identified threshold implanted compared to an implantation rate across the whole study population of 47% (34/72). For Fragouli et al., these figures were 59% (16/27) versus 38% (16/42) respectively.
Both groups have initiated an RCT. MitoScore is marketed by the group of Diez and is currently tested in RCT NCT02662686 (clinicaltrials.gov). Mitograde™ is marketed by Reprogenetics and is being tested in RCT NCT02673125.

Currently, there is no evidence that selection through mtDNA load measurement increases LBR. Application of the technique should therefore strictly be limited to participation in either one of RCTs, and this should clearly communicated to the patient.

**Assisted hatching**

In 1990, Cohen et al., proposed that making a breach in the zona pellucida may help implantation in some patients. Assisted hatching (AH) is usually performed on day 3, 5 or 6 of embryo development using a non-contact laser, but mechanical or acidic solutions have also been used (Balaban et al., 2002). Clinics use AH for patients of advanced maternal age, smokers or patients with a raised FSH, or when transferring embryos that have been cryopreserved.

Three meta-analyses on AH have found a significant increase in CPR but no evidence for a difference in LBR. Martins et al., (2011) found a significant difference in CPR using frozen thawed embryos in unselected women and for patients with repeated IVF failure, but no evidence of benefit for subgroups of either older women or those with a good prognosis. They concluded that there were too few studies looking at LBR to draw conclusions. The Cochrane review by Carney et al. (2012) looked at 31 trials including 1992 clinical pregnancies in 5728 women. Nine of the 31 RCTs included data on LBR. There was no evidence of difference between the LBR in the AH and control groups. Li et al., (2016) looked at 36 RCTs with 6,459 participants and found that AH gave a significant increase in CPR and multiple pregnancy rate but in the 15 RCTs that looked at LBR, there was no evidence of difference between the AH and control groups.

The National Institute for Clinical Excellence (NICE) guidelines (2013) state that “assisted hatching is not recommended because it has not been shown to improve pregnancy rates”.


Duty of care towards the offspring

Whilst we have considered evidence for increasing the chances of pregnancy and live birth, very few interventions in this field have considered the long term health of the child. Individual clinics and national and international data collection bodies have a duty to evaluate data surrounding the use of adjuncts in IVF and collect long term data pertaining to the health of any children born as a result of their use.

At the individual practitioner level, doctors and scientists recommending an unproven procedure to their patients must ensure that they provide comprehensive information surrounding the lack of evidence on the safety of the intervention for the resultant child. As best clinical practice dictates that professional guidelines are followed when managing patients, adjuncts that have not been proven to be beneficial should be used with caution, if used at all. Furthermore, regulatory bodies could insist that any empirical therapy prescribed must be accounted for, ideally with the establishment of clinical trials, to ensure long-term maternal and neonatal follow-up.

Among the techniques described in this paper, it is possible that some could have an impact upon the health of the embryo and the newborn. While some retrospective studies have been published, there are no RCTs on the impact of these technologies upon newborn health and child development.

The only Cochrane review on AH identified two studies out of 31 reporting on congenital anomalies and concluded that many unanswered questions remain about the perceived risks of the procedure, from embryo damage to chromosomal and congenital abnormalities (Carney et al., 2012).

Conclusion

IVF clinicians and scientists must recognise that appropriately powered, well-designed, peer-reviewed RCTs, with a LBR outcome measure which goes on to report on child health, are the gold standard of evidenced-based medicine.

Those advocating and recommending unproven procedures to their patients must ensure that they fully inform the patient of the evidence for its safety and effectiveness orally and in writing to ensure that people considering treatment using
adjunct therapies are in a position to make an informed decision. It is also important that all procedures performed, including the adjunct treatments, are well documented and followed up.

Regulators and professional bodies also have a role to play in ensuring that only suitable practices are used in the clinic.

**Authors' Roles**

All authors contributed to planning, writing and revising article. Additionally:

Joyce Harper conceived the idea for the paper and wrote the section on AH

Emily Jackson wrote the legal section

Karen Sermon wrote the sections on preimplantation genetic screening and mitochondria DNA load measurement

Robert John Aitken wrote the section on sperm DNA fragmentation

Stephen Harbottle wrote the section on embryo glue and adherence compounds

Edgar Mocanu wrote the section on duty of care towards the offspring

Thorir Hardarson wrote the section on time lapse imaging

**Funding**

No funding was used to produce this paper

**Conflict of Interests**

R John Aiken has an honorary position on the Scientific Advisory Board of CellOxcess, a NJ-based biotechnology company dedicated to disease prevention through the diagnosis and treatment of Chronic Cell Oxidative Stress.
Raj Mathur received hospitality at academic meetings from pharmaceutical companies Merck Serono, Ferring and Finox Biotech.

Andy Vail is a statistical editor of the Cochrane Gynaecology and Fertility Review Group

References


Barber, R. (2015) The killer cells that robbed me of four babies: TV star Sally Meen on her struggle to have two healthy daughters Daily Mail 2 January.


Griesinger G. Beware of the “implantation rate”! Why the outcome parameter “implantation rate” should be abandoned from infertility research. *Hum Reprod* [Internet] 2016;**31**:249–251.


HFEA code of practice:


NICE guidelines (2013) - https://www.nice.org.uk/donotdo/assisted-hatching-is-not-recommended-because-it-has-not-been-shown-toimprove-pregnancy-rates


