Time-lapse in the IVF-lab: how should we assess potential benefit?

S. Armstrong¹,*, A. Vail², S. Mastenbroek³, V. Jordan¹, and C. Farquhar¹

¹Department of Obstetrics and Gynaecology, University of Auckland, Auckland, New Zealand ²Manchester Academic Health Science Centre, Salford Royal Hospital, Centre for Biostatistics, University of Manchester, Manchester, UK ³Center for Reproductive Medicine, University of Amsterdam, Academic Medical Center, Amsterdam, The Netherlands

*Correspondence address. E-mail: s.armstrong@auckland.ac.nz

Submitted on July 27, 2014; resubmitted on July 27, 2014; accepted on August 21, 2014

ABSTRACT: Time-lapse imaging of embryos has been widely introduced to fertility laboratories worldwide with the aim of identifying the best quality embryos to transfer that will ultimately improve IVF success rates. In this opinion paper, we explore the lack of evidence of benefit of this novel intervention, analyse the methodological flaws of current studies, offer ideal study designs that assess the various features of time-lapse imaging, and discuss forthcoming studies. In particular, we emphasize the ethical aspects of hastily adopting a costly technology without current high level evidence of improved live birth rates, safety and cost effectiveness.

Key words: IVF / time-lapse / live birth / IVF outcome / embryo incubation

Introduction

Time-lapse imaging of embryos has arrived. Over the past 5 years a number of different time-lapse systems have been introduced to fertility laboratories with the aim of identifying better quality embryos that would ultimately improve the success rates of IVF (Meseguer et al., 2012).

Time-lapse monitoring systems (TLS) take digital images of embryos at set time intervals. The system can either be installed into an existing embryo incubator or can exist as a combined time-lapse incubation system. The images are compiled using specialist software to create a time-lapse sequence of embryo development, thus negating the need for the embryologist to remove embryos from the incubator for morphological assessment. Some TLS also utilize computer-assisted assessment of developmental milestones of embryos, also known as morphokinetic parameters, to offer a semi-quantitative process of embryo evaluation (Conaghan et al., 2013). These cell-tracking software algorithms have evolved as a non-invasive, non-subjective way of attempting to improve the selection of embryos with the highest implantation potential.

There are a number of TLS on the market developed by various manufacturers, including Embryoscope® (Fertilitech), Primo Vision (Vitrolife) and Eeva (Auxogyn, Inc.). Despite the technology being novel, numerous fertility clinics worldwide already have adopted TLS, often charging patients an additional fee (from several hundreds to well over one thousand U.S. dollar) for its use. The media have enthusiastically reported on TLS based on preliminary scientific publications (Campbell et al., 2013; Devlin, 2013; Sample, 2013). Claims of tripling of IVF success rates by TLS, leading to live birth rates of 78%, helped to fuel the hype around the technology, without acknowledging limitations in study design of the available scientific publications for establishing the effectiveness of the novel intervention (Campbell et al., 2013; Devlin, 2013; Sample, 2013). The industry behind TLS has largely driven the widespread adoption of the technology citing ‘improved success rates’, the advantage of ‘bringing the latest technology to patients’ and ‘adding value to the treatment cycle’ (FertiliTech).

How should we test new interventions in fertility treatments?

Fertility is no different from any other clinical specialty in that new interventions, such as drugs and devices, should be evaluated by RCTs to establish their safety and clinical and cost effectiveness. Selecting an embryo using a TLS is a complex intervention and there are recognized ways of assessing such technologies (Campbell et al., 2000). Countering the potential benefits outlined in the introduction, TLS involve exposing embryos to light during image acquisition at pre-determined intervals, often as frequently as every 5 min. Although the total dose of UV radiation is likely to be low, there is a potential for harm. Moreover, TLS are more costly than conventional incubators. It is therefore incumbent on proponents to demonstrate both safety of the intervention and sufficient clinical benefit to justify the additional cost. In this context, demonstration only of equivalence or non-inferiority would not justify uptake outside of the experimental setting.
How should time-lapse systems be tested?

The potential advantages of TLS can be divided into two distinct categories. First is the ability of the system to accumulate multiple, detailed, time-lapse images of embryos which can be utilized either by cell-tracking software algorithms or by embryologists undertaking morphological assessment, to select the highest quality embryo for transfer. Second is the effect of improved culture conditions, whereby human handling is minimized, air temperature and gas compositions are kept stable, and embryos are not exposed to bench-top light microscopy. Designing a study to assess any overall benefit of TLS over current conventional care does not require separation of these two potential effects. However, in the event that an overall benefit is identified, the ability to distinguish between these two potential effects would aid efficient implementation. Below are three suggested trial designs that address these disparate issues.

Trial design one: the aim of this trial will be to assess the culture conditions of TLS. Embryos will be cultured in either the TLS or conventional incubator (CI) but assessment will be by routine morphological criteria in both arms.

As an early phase design concerned only with preliminary safety assessment of the TLS culture method it would be statistically efficient to randomize embryos. The drawback of this design is that it precludes interpretation of clinical outcomes. Randomization of women will be essential to assess any clinical benefit of the culture conditions of TLS. With either design it is essential to account for the clustering of embryos in the sample size calculation and statistical analysis when comparing morphological criteria between treatment arms.

Trial design two: the aim of this trial will be to establish whether time-lapse cell-tracking algorithms are better at selecting ‘top quality’ embryos more likely to result in live birth than conventional morphological assessment undertaken by an embryologist. The trial would need to randomize women (rather than embryos) to either TLS utilizing time-lapse cell-tracking algorithms, or TLS with conventional assessment of morphological parameters from TLS images, to select which embryo to transfer. Live birth and adverse events would be the primary outcomes. The identical culture conditions in both arms of this trial will help determine whether it is the time-lapse algorithms or rather the stable culture conditions that have the greatest impact on embryo development and live birth.

Apart from the different algorithms used between different TLS, the complexities in this study design also arise from the various definitions of what ‘conventional morphological assessment’ is. There is consensus on morphological parameters at any given developmental stage that are important for an embryologist when selecting embryos for transfer, as outlined by the Istanbul consensus workshop on embryo assessment (Balaban et al., 2011). However, the parameters that are actually being used differ between labs, just as the number of assessments that are performed (e.g. assessment every day or only one assessment just before transfer). Some even use algorithms, not unlike the algorithms used in TLS, using the parameters obtained from only one assessment a day (van Loendersloot et al., 2014). Therefore trial reports need extensive description of protocols used, so this can be taken into account in future reviews and meta-analyses.

Trial design three: the aim of this trial will be to establish whether TLS, including its cell-tracking algorithms and stable culture environment, results in improved cumulative live birth rates over conventional incubation and morphological selection. The trial would randomize women to have all of their embryos incubated within either the TLS, utilizing the cell-tracking algorithms, or the CI utilizing conventional light microscopy for embryo assessment. The primary outcomes would be live birth and adverse events. Given the additional cost associated with TLS, economic evaluation would be undertaken as part of this trial. An RCT that included a childhood follow-up assessment for morbidity would be very valuable to assess potential longer-term outcomes from time-lapse imaging technology.

Also in this case it is essential to be aware of possible differences in protocols between trials. There are differences between TLS systems in methodology and culture conditions that could have an effect on embryo quality, and with that live birth rates. Systems cannot be considered the same in this respect. But the difficulty also lies in the control group. For example, what incubator is being used? And is the microscopy performed on a regular bench in an open environment without controlled conditions or in a heated, gas-controlled cabinet that is positioned close to the incubator? This is important to know as a positive effect of TLS relative to a ‘stable environment’ may be much less likely.

Systematic reviews of RCTs of time-lapse systems

Our Cochrane group has recently assessed the published studies for inclusion in a systematic review on time-lapse imaging of embryos in assisted reproduction technology (ART) (Armstrong et al., 2014). A comprehensive search strategy was used to capture as many RCTs of time-lapse imaging for human ART as possible in electronic databases. This search identified four reports of randomized trials from two separate studies (Table I) (Maria et al., 2010; Cruz et al., 2011; Ingerslev et al., 2011; Kirkegaard et al., 2012). Both of these studies utilized the same time-lapse imaging technology; Embryoscope. The trial by Kirkegaard et al. (2012) acknowledges part-funding by FertilitTech, the manufacturers of Embryoscope.

Both of these studies utilized ‘trial design one’ above. They were designed to address the important initial question of whether there were safety concerns within the preimplantation phase of embryo development that should preclude clinical evaluation of the technology. Embryos from each participant were distributed between the competing incubators. The primary outcome assessment, implicit from the reported sample size calculations, was of development and grading of the embryos.

There are technical flaws in the design, reporting and interpretation of both studies. For example, neither accounted in design or analysis for clustering induced by multiple embryos from each participant, both concluded equivalence apparently on the basis of absence of a statistically significant difference, and both sought to draw inference from implantation and pregnancy outcomes that, by design, occurred after the randomized comparison was completed (Altman and Bland, 1995; Vail and Gardener, 2003).

Neither group of authors made explicit a threshold value at which their stated conclusion of laboratory equivalence would have been altered. Kirkegaard et al. (2012) appropriately described their study as ‘a necessary step’ preliminary to a study that could compare live birth rate and
would require a significantly larger number of participants’. The flaws in both studies mean that they failed to rule out inferiority of the TLS culture conditions.

Trials databases were searched to establish ongoing RCTs utilizing TLS for embryo selection (Table II). Five studies were found, run by the following principal investigators: Lundin, Meseguer, Kovacs, Desai and Pai (Kovacs, 2012; Meseguer, 2012; Pai, 2012; Lundin, 2013; Desai, 2014). Lundin and Kovacs have utilized ‘trial design one’ above but, unlike Cruz and Kirkegaard, randomized women rather than embryos. This design allows assessment of clinical as well as laboratory outcomes. With a combined sample size of around 400 participants these are unlikely to establish realistic differences in clinical outcomes.

Desai has utilized ‘trial design two’ above, where both the intervention arm and the control arm of the study involve incubation in a TLS. The intervention arm embryos are selected by TLS cell-tracking algorithms, whereas the control arm embryos are selected using morphological assessment of TLS images by embryologists. In principle this design allows assessment of the added value of TLS imaging over standard morphological assessment. In practice, the target enrolment of just 300 participants has little statistical power to detect realistic differences in the stated primary outcome of clinical pregnancy.

The two remaining studies have utilized ‘trial design three’ above, comparing the introduction of TLS incubation, including exploitation of the time-lapse cell-tracking algorithms, with conventional incubation and standard morphological assessment. In Meseguer’s study the primary outcome is ongoing pregnancy, and on correspondence with the author, there are plans to also examine live birth as the trial matures, whereas Pai has proportion of good quality embryos as the primary outcome and clinical pregnancy as the secondary. The combined recruitment target of nearly 1000 participants would, even if within a single trial, require a substantial difference in clinical pregnancy rates to claim statistical power at 80%.

A systematic review of clinical outcomes following selection of embryos with TLS has recently been undertaken by Kaser and Racowsky (2014). Included studies were retrospective or prospective cohort studies, most without a control arm, thus making comparisons unreliable. Kaser and Racowsky (2014) conclude that there are currently no high quality data to support the clinical use of TLS and that it should remain an experimental strategy subject to institutional review and approval (Kaser and Racowsky, 2014). The authors outline an ideal study, similar to our ‘trial design two’, but neglect to mention the importance of also running trials that utilize CI as a control in order to compare the clinical outcomes, safety and cost effectiveness of TLS in its entirety against CI.

Conceptually, if studies using ‘trial design two’ find equivalent live birth rates after IVF, but trials using trial design three find a benefit, then we can discard time-lapse and focus on the improved culture conditions to achieve progress in this field of work.

What do these studies tell us?

Time-lapse systems for selecting embryos have been introduced on the basis of evidence from studies designed to assess equivalence and limited safety prior to embryo transfer without any evaluation of clinical benefit to subfertile couples or cost effectiveness. The only identified published studies chose to randomize oocytes or embryos and not women. Randomizing oocytes or embryos to receive either the intervention (TLS) or control (CI), rather than randomizing the woman herself, means that it is not possible to evaluate the pregnancy outcomes per woman. Therefore the drawing of clinical or policy conclusions on the potential advantage of time-lapse imaging of embryos is currently not possible.

Moral obligation

The lack of evidence of benefit of TLS would not matter if the companies who manufacture them, and the clinics who utilize them, had not promoted and sold to their patients the technology as improving pregnancy outcomes (FertiliTech; Levack et al., 2006). History has taught us that the hasty adoption of a novel intervention based on limited available literature, such as that of preimplantation genetic screening (PGS) 15 years ago, can lead to a huge number of women paying more for a less effective treatment (Mastenbroek and Repping, 2014). Concern exists that the widespread introduction of novel technologies, such as TLS, may be due to lack of awareness of the true value of the available data, or that other commercial motives are at play (Mastenbroek and Repping, 2014).

Let us not forget that time-lapse imaging of embryos is not novel per se. The first application of time-lapse photography in embryo research stems from 1929, when WH Lewis undertook a study imaging rabbit embryos (Lewis and Gregory, 1929). Recent hype surrounding the technology

---

<table>
<thead>
<tr>
<th>Study reference</th>
<th>Patients/population</th>
<th>Intervention</th>
<th>Comparison</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cruz (2011)</td>
<td>478 embryos at 17–19 h from donor oocytes from women aged 18–35 years being replaced into 60 women undergoing assisted reproduction techniques (ART) aged 32–45 years</td>
<td>Incubation in a time-lapse monitoring system (TLS: Embryoscope)</td>
<td>Incubation in a conventional incubator</td>
<td>Blastocyst rates</td>
</tr>
<tr>
<td>Kirkegaard et al. (2012)</td>
<td>676 oocytes from 59 women undergoing their 2nd or 3rd cycle of ART, aged &lt;37 years</td>
<td>Initial incubation to 20 h in a conventional incubator followed by incubation in a TLS (Embryoscope)</td>
<td>Incubation in a conventional incubator</td>
<td>4 cell embryos on Day 2</td>
</tr>
</tbody>
</table>

Proportion of 7–8 cell embryos on Day 3

Proportion of blastocysts on Day 5

Implantation rate (fetal heart at 8 weeks after embryo transfer)
<table>
<thead>
<tr>
<th>Study title</th>
<th>Year registered</th>
<th>Identifier</th>
<th>Principle investigator</th>
<th>Sponsor</th>
<th>Location</th>
<th>Design</th>
<th>Status</th>
<th>Primary outcome measure</th>
<th>Stated purpose</th>
<th>Usefulness of trial design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessing quality of human embryos cultured in a closed system compared with embryos cultured in a conventional incubator</td>
<td>2013</td>
<td>ISRCTN13111817</td>
<td>Prof Kersti Lundin</td>
<td>Unisense FertilTech</td>
<td>Sweden</td>
<td>RCT</td>
<td>Completed, awaiting publication</td>
<td>Number of cleaved embryos of high quality</td>
<td>To determine whether incubation in a TLS utilizing time-lapse images to assess morphology improves the rate of high quality cleaved embryos when compared with conventional incubation without time-lapse.</td>
<td>Utilizes elements of ‘trial design one’ methodology, relying on conventional morphological scoring of images. Primary outcome is proportion of high quality embryos, therefore acting as a trial to determine effectiveness of culture conditions over TLS imaging.</td>
</tr>
<tr>
<td>Embryo Selection by Time-Lapse Monitoring for Single Embryo Transfer</td>
<td>2012</td>
<td>NCT01694641</td>
<td>Peter Kovacs, MD</td>
<td>Kaali Institute IVF Center</td>
<td>Hungary</td>
<td>RCT</td>
<td>Recruiting</td>
<td>Intrauterine pregnancy with heartbeat</td>
<td>To study whether TL monitoring is superior in terms of pregnancy rates to traditional embryo observation when a single blastocyst is selected for transfer</td>
<td>Utilizes elements of ‘trial design one’ methodology, relying on conventional morphological scoring of images, therefore acting as a trial to determine effectiveness of culture conditions over TLS imaging.</td>
</tr>
<tr>
<td>Use of Time-Lapse Morphological Kinetics in the Selection of Blastocysts</td>
<td>2014</td>
<td>NCT02081859</td>
<td>Nina Desai, PhD</td>
<td>The Cleveland Clinic</td>
<td>USA</td>
<td>RCT</td>
<td>Recruiting</td>
<td>Clinical pregnancy rate</td>
<td>To determine if data obtained from the embryoscope is helpful in determining which blastocysts to transfer</td>
<td>Utilizes ‘trial design two’ methodology to address whether there is any difference in pregnancy rates between embryos selected by conventional morphological assessment and time-lapse cell-tracking algorithms.</td>
</tr>
<tr>
<td>Clinical Validation of Embryo Culture and Selection by Morphokinetic Analysis</td>
<td>2012</td>
<td>NCT01549262</td>
<td>Marcos Meseguer, PhD</td>
<td>Instituto Valenciano de Infertilidad, IVI Valencia</td>
<td>Spain</td>
<td>RCT</td>
<td>Completed, awaiting publication</td>
<td>Ongoing pregnancy rate</td>
<td>To determine whether the multivariable time-lapse model for embryo selection (Meseguer et al., 2012) together with undisturbed controlled conditions obtained by a time-lapse incubator affects pregnancy rate</td>
<td>Utilizes ‘trial design three’ methodology to establish whether ongoing pregnancy is influenced by TLS embryo assessment as well as culture conditions.</td>
</tr>
</tbody>
</table>
has arisen from the development by industry of new machines that allow easy implementation of the technology. The pharmaceutical and devices industry employ sophisticated marketing campaigns and one must merely glance at TLS suppliers’ list of conference attendances to recognize that a costly public relations offensive is underway.

From an ethical standpoint, the charging of patients, often between 10 and 20% of the cost of an IVF cycle, to use TLS is questionable given the paucity of evidence on live birth outcomes. The economic burden to society must also be considered, especially if TLS is adopted as an intervention available through public health systems. Can governments afford such a costly addition, without the corresponding evidence to advocate its use? And evaluation of effectiveness is not to be dismissed as it can be argued that spending so much money on acquiring equipment to improve embryo selection is misguided, especially in the light of recent evidence that suggests that with a good cryopreservation programme, transfer of frozen embryos in a subsequent non-hyperstimulated cycle may result in similar, if not improved, cumulative pregnancy rates (Roberts et al., 2010; Mastenbroek et al., 2011; Wong et al., 2014). Therefore, in a situation where all available embryos can be cryopreserved and transferred in subsequent cycles without impairment in pregnancy rates, no embryo selection method, including TLS, will ever lead to improved live birth rates; the only parameter that may be influenced is time to pregnancy, if embryo ranking is improved.

### Conclusion

TLS has been hastily introduced to IVF laboratories worldwide without current high level evidence of improved live birth rates, safety and cost effectiveness. Couples seeking fertility treatments are eager to attempt anything that might increase their chance of a healthy child, and are vulnerable to paying for novel interventions before the evidence is available to support their use. It is therefore incumbent on the fertility profession, whose responsibility lies with the patient, not commercial interests, to maintain responsibility for the proper testing and introduction of new treatments and interventions such as TLS.

RCTs that assess both the impact of the improved culture conditions and the time-lapse imaging potential are needed to establish where the virtue of TLS truly lies. Researchers who are persuaded by the existing evidence on preliminary safety should progress to study clinical and cost effectiveness. Until phase III RCTs have been completed, where women and not oocytes or embryos are randomized, fertility clinics should decline to include this costly intervention in an assisted reproduction protocol.

### Authors’ roles

S.A. and C.F. conceived the idea for the paper. S.A. wrote the initial draft, which was edited extensively by all the authors.

### Funding

No funding was received for the writing of this paper.

### Conflict of interest

C.F. is the Chair of the Cochrane menstrual disorders and subfertility group.
References


Devlin H. New IVF technique could give 78 per cent chance of success. The Times 2013.


IVF ‘imaging’ advance triples success rate. The New Zealand Herald.


