



A double-blind randomized controlled trial investigating a time-lapse algorithm for selecting Day 5 blastocysts for transfer

Aisling Ahlström ^{1,2,*}, Kersti Lundin^{2,3}, Anna-Karin Lind^{1,4}, Kristina Gunnarsson¹, Göran Westlander¹, Hannah Park³, Anna Thurin-Kjellberg^{2,3}, Steinunn A. Thorsteinsdottir⁵, Snorri Einarsson⁵, Mari Åström⁶, Kristina Löfdahl⁶, Judith Menezes⁷, Susanne Callender⁸, Cina Nyberg⁸, Jens Winerdal⁹, Camilla Stenfelt⁹, Brit-Randi Jonassen¹⁰, Nan Oldereid¹⁰, Lisa Nolte¹¹, Malin Sundler¹¹, and Thorir Hardarson^{1,12}

¹Livio Gothenburg, Gothenburg, Sweden ²Department of Obstetrics and Gynecology, Institute of Clinical Sciences, Sahlgrenska Academy, Gothenburg University, Gothenburg, Sweden ³Reproductive Medicine, Sahlgrenska University Hospital, Gothenburg, Sweden ⁴Livio Falun, Falun, Sweden ⁵Livio Reykjavik, Reykjavik, Iceland ⁶Livio Umeå, Norrlands Universitetssjukhus, Umeå, Sweden ⁷Reproductive Medicine, Karolinska University Hospital, Stockholm, Sweden ⁸Livio Kungsholmen, Stockholm, Sweden ⁹Livio Gärdet, Stockholm, Sweden ¹⁰Livio Oslo, Oslo, Norway ¹¹Livio Malmö, Malmö, Sweden ¹²Vitrolife Sweden AB, Göteborg, Sweden

*Correspondence address. Livio Gothenburg, Carlandersparken 24, 402 29 Gothenburg, Sweden. E-mail: aisling.ahlstrom@livio.se

 <https://orcid.org/0000-0002-5074-912X>

Submitted on November 08, 2021; resubmitted on December 12, 2021; editorial decision on January 18, 2022

STUDY QUESTION: Can use of a commercially available time-lapse algorithm for Day 5 blastocyst selection improve pregnancy rates compared with morphology alone?

SUMMARY ANSWER: The use of a time-lapse selection model to choose blastocysts for fresh single embryo transfer on Day 5 did not improve ongoing pregnancy rate compared to morphology alone.

WHAT IS KNOWN ALREADY: Evidence from time-lapse monitoring suggests correlations between timing of key developmental events and embryo viability. No good quality evidence exists to support improved pregnancy rates following time-lapse selection.

STUDY DESIGN, SIZE, DURATION: A prospective multicenter randomized controlled trial including 776 randomized patients was performed between 2018 and 2021. Patients with at least two good quality blastocysts on Day 5 were allocated by a computer randomization program in a proportion of 1:1 into either the control group, whereby single blastocysts were selected for transfer by morphology alone, or the intervention group whereby final selection was decided by a commercially available time-lapse model. The embryologists at the time of blastocyst morphological scoring were blinded to which study group the patients would be randomized, and the physician and patients were blind to which group they were allocated until after the primary outcome was known. The primary outcome was number of ongoing pregnancies in the two groups.

PARTICIPANTS/MATERIALS, SETTING, METHODS: From 10 Nordic IVF clinics, 776 patients with a minimum of two good quality blastocysts on Day 5 (D5) were randomized into one of the two study groups. A commercial time-lapse model decided the final selection of blastocysts for 387 patients in the intervention (time-lapse) group, and blastocysts with the highest morphological score were transferred for 389 patients in the control group. Only single embryo transfers in fresh cycles were performed.

MAIN RESULTS AND THE ROLE OF CHANCE: In the full analysis set, the ongoing pregnancy rate for the time-lapse group was 47.4% (175/369) and 48.1% (181/376) in the control group. No statistically significant difference was found between the two groups: mean difference -0.7% (95% CI $-8.2, 6.7, P=0.90$). Pregnancy rate (60.2% versus 59.0%, mean difference 1.1% , 95% CI $-6.2, 8.4, P=0.81$) and early pregnancy loss (21.2% versus 18.5%, mean difference 2.7% , 95% CI $-5.2, 10.6, P=0.55$) were the same for the

time-lapse and the control group. Subgroup analyses showed that patient and treatment characteristics did not significantly affect the commercial time-lapse model D5 performance. In the time-lapse group, the choice of best blastocyst changed on 42% of occasions (154/369, 95% CI 36.9, 47.2) after the algorithm was applied, and this rate was similar for most treatment clinics.

LIMITATIONS, REASONS FOR CAUTION: During 2020, the patient recruitment rate slowed down at participating clinics owing to coronavirus disease-19 restrictions, so the target sample size was not achieved as planned and it was decided to stop the trial prematurely. The study only investigated embryo selection at the blastocyst stage on D5 in fresh IVF transfer cycles. In addition, only blastocysts of good morphological quality were considered for transfer, limiting the number of embryos for selection in both groups: also, it could be argued that this manual preselection of blastocysts limits the theoretical selection power of time-lapse, as well as restricting the results mainly to a good prognosis patient group. Most patients were aimed for blastocyst stage transfer when a minimum of five zygotes were available for extended culture. Finally, the primary clinical outcome evaluated was pregnancy to only 6–8 weeks.

WIDER IMPLICATIONS OF THE FINDINGS: The study suggests that time-lapse selection with a commercially available time-lapse model does not increase chance of ongoing pregnancy after single blastocyst transfer on Day 5 compared to morphology alone.

STUDY FUNDING/COMPETING INTEREST(S): The study was financed by a grant from the Swedish state under the ALF-agreement between the Swedish government and the county councils (ALFGBG-723141). Vitrolife supported the study with embryo culture dishes and culture media. During the study period, T.H. changed his employment from Livio AB to Vitrolife AB. All other authors have no conflicts of interests to disclose.

TRIAL REGISTRATION NUMBER: ClinicalTrials.gov registration number NCT03445923.

TRIAL REGISTRATION DATE: 26 February 2018.

DATE OF FIRST PATIENT'S ENROLMENT: 11 June 2018.

Keywords: time-lapse / blastocyst / single embryo transfer / IVF / embryo selection

Introduction

Internationally, several embryos are often transferred after IVF in order to achieve satisfactory birth rates. The main side-effect of this strategy is an unacceptably high rate of multiple pregnancies with the risk of serious consequences for the children (Beral *et al.*, 1990; Craft, 1990; Gissler *et al.*, 1995; Westergaard *et al.*, 1999). The reason for transferring more than one embryo is the lack of an ideal method for selecting the best embryo. With many clinics striving to implement transfer of single embryos, additional transfers with cryopreserved embryos derived from the same fresh IVF cycle are often required to achieve a live birth (Fernandez-Shaw *et al.*, 2015; De Vos *et al.*, 2016; De Croo *et al.*, 2019). The success of these frozen-thawed embryo transfer cycles suggests that embryos with live birth potential are present within the same cohort but were not chosen first. As such, selection methods, which can ensure that the best embryos are chosen first, would promote single embryo transfer (SET), possibly shorten time to pregnancy and minimize problems associated with multiple births.

Currently, non-invasive embryo selection methods are based on developmental rate and morphological grading evaluated at distinct time points by light microscopy (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011). This methodology is limited by the ability of a small number of observations to provide a complete picture of the dynamic process of embryo development and by observer subjectivity, with considerable inter- and intra-observer variation when grading embryos (Arce *et al.*, 2006; Paternot *et al.*, 2011). In comparison, time-lapse (TL) monitoring systems, which capture continuous images over time, make it possible to observe morphological events and define more precisely the timings of development while maintaining embryos within a stable environment. A large number of studies report correlations between timings of key events and implantation potential or surrogate outcomes, such as

good blastocyst development and aneuploidy (reviewed in Kirkegaard *et al.* (2015)). Predictive models using these TL parameters to improve discrimination between viable and non-viable embryos have shown promising results in retrospective studies (Meseguer *et al.*, 2011, 2012; Rubio *et al.*, 2012; Conaghan *et al.*, 2013). However, almost all prospective studies either have been underpowered, non-randomized or have a high risk of bias (Goodman *et al.*, 2016; Kaser *et al.*, 2017; Yang *et al.*, 2018; Armstrong *et al.*, 2019; Kovacs *et al.*, 2019). To date, only one adequately powered randomized controlled trial (RCT) has been published (Rubio *et al.*, 2014). The authors concluded that the use of TL increased the chance of a pregnancy significantly. However, the design of the study did not allow for a distinction between the possible benefit of the improved culture conditions and the hierarchical selection model used.

The aim of this RCT was to determine if, when performing embryo selection and transfer on Day 5, a TL algorithm *per se* can enhance the prediction of the embryo's reproductive potential compared to embryo morphology alone.

Materials and methods

Study design

A double-blinded RCT was conducted at 10 Nordic IVF clinics between June 2018 and March 2021. All participants provided written informed consent. The study was approved by the Swedish Ethical Review Authority (Dnr 954-17), the National Bioethics committee in Iceland (VSN-18-132) and the Norwegian National Research Ethics Committee (No. 2018/126/REK sør-øst). The study was also registered in the ClinicalTrials.gov database (registration number NCT03445923).

Eligibility criteria

Couples undergoing standard IVF or ICSI were eligible for enrollment. Patients were excluded if they were undergoing double embryo transfer, preimplantation genetic testing treatment, embryo transfer on any other day than Day 5 or freeze all was planned. Patients were also excluded if their treatment cycle used vitrified-warmed oocytes. Only one cycle per patient was randomized in the study.

Randomization and blinding

Patients meeting the inclusion criteria and with at least two good quality blastocysts (GQB) on Day 5 were randomized. Blastocyst quality was graded according to Gardner and Schoolcraft (1999), and a GQB was defined as an embryo with a grade ≥ 3 BB. Randomization into either the control or TL group in a 1:1 ratio was performed on the day of embryo transfer by a computerized randomization program. This program balanced allocation by way of minimization (1:1) taking into account female age, number of GQE available on Day 5, number of previous fresh IVF transfer cycles, indication for treatment, BMI, smoking, method of fertilization and number of oocytes retrieved. Embryologists performed morphological scoring of all available embryos in a cohort and then identified the best embryo according to the Gardner score that would be transferred based on morphology alone. If two or more GQB were present, a randomization was then performed. As such, embryologists were blinded to group allocation during morphological evaluation and embryo preselection. The physician and patients were blinded to which group they were allocated until after the primary endpoint was known.

IVF treatment

Clinical protocol

All patients underwent controlled ovarian stimulation according to routine methods in each clinic. Both GnRH antagonist and agonist treatment protocols were included. Oocytes were retrieved $\sim 36 \pm 2$ h after hCG administration (Ovitrelle 6500 IU, Merck, Darmstadt, Germany).

Lutinus vaginal tablets (Ferring, Copenhagen, Denmark) three times daily, Cyclogest vaginal tablets (Gedeon Richter, Hungary) twice daily or Crinone vaginal gel (Merck Serono) once or twice daily were given as luteal support, starting the day after oocyte retrieval.

Laboratory protocol

All oocytes were rinsed in G-Gamete or G-MOPS (Vitrolife, Gothenburg, Sweden) and placed in G-IVF medium (Vitrolife) for 2–5 h before fertilization by conventional methods for standard insemination or ICSI.

Immediately after ICSI or for inseminated oocytes after denudation on Day 1, embryos were placed into pre-equilibrated culture dishes (Embryoslides, Vitrolife) containing single step GTL media (Vitrolife) overlaid with paraffin oil (Ovoil, Vitrolife). Embryos were cultured up to the blastocyst stage in an Embryoscope time lapse incubator (Vitrolife) at 37°C, 5% CO₂ and under reduced oxygen tension (5–6% O₂). Images of each embryo were captured every 10 min. Fertilization was evaluated using the images captured by the Embryoscope at 16–18 h post-insemination. In general, patients with ≥ 5 zygotes on Day 1 were eligible for blastocyst culture aiming for Day 5 embryo transfer. After fertilization check, no further evaluations were performed, and embryos remained in culture until the day of transfer.

Embryo selection strategies

On Day 5 (115 ± 2 h post-ICSI/insemination), embryos were assessed under an inverted microscope according to Gardner and Schoolcraft's (1999) grading system. Routinely, blastocysts were graded by at least two embryologists in order to reduce potential interobserver variations in embryo grading. An embryo with a grade ≥ 3 BB was considered a GQB. A preliminary decision was then made on which embryo to select for transfer. If two or more GQB were available, patients were randomized into either the control or the TL group.

For patients randomized into the control group, the preliminary embryo selection decision was maintained. No information available from TL imaging was used for embryo assessment or selection in the control group.

For patients randomized into the TL group, the GQB with the highest Day 5 KIDScore™ D5 was chosen for embryo transfer. To ensure reliable scores were attained, the morphokinetic variables; time to two, three, four and five cells, time to start of blastulation (tSB) and morphological inner-cell mass (ICM) and trophectoderm (TE) scores were evaluated using the TL images captured up to 120 h post-insemination. Morphokinetic variables were annotated according to the guidelines proposed by Ciray et al. (2014). During the study period the algorithm software was updated and three versions of the KIDScore™ D5 model (v1, v2, v3) (Vitrolife) were used. Fresh SET on Day 5 was performed and Embryo glue (Vitrolife) was used as transfer media. Supernumerary GQBs were vitrified on Day 5 or Day 6.

Outcomes

The primary endpoint for this study was the percentage of ongoing pregnancies defined by the presence of a gestational sac with foetal heart activity detected by sonography at 6–8 weeks. Secondary endpoints were pregnancy (defined by a positive beta-hCG urinary test performed 16 days after embryo transfer), early pregnancy loss and agreement (Y/N) between the preliminary decisions based on morphology and the final decision when KIDScore ranking was applied (TL group only). Early pregnancy loss was defined as a positive pregnancy that did not result in an ongoing pregnancy.

Statistical analysis

The distribution of continuous variables is given as mean, SD, median, minimum, maximum, quartile 1 and quartile 3. Categorical variables are given as numbers and percentages. For comparison between the two randomized groups, Fisher's exact test was used for dichotomous variables (e.g. primary ongoing pregnancies), Fisher's non-parametric permutation test was used for continuous variables, Mantel–Haenszel's chi-square test for ordered categorical variables and Pearson's chi-square test for unordered categorical variables.

For all outcome variables exact 95% CI are given for all percentages within groups, and for between groups the mean differences with 95% CI together with relative risk (RR) with 95% CI. A sensitivity analysis of primary analysis adjusted for all optimal allocation variables was performed with multivariable logistic regression and odds ratio with 95% CI are given. This analysis was adjusted for female age, number of GQB on Day 5, number of oocyte retrievals, indication for treatment, female BMI, female smoking, fertilization method, number of oocytes and treatment clinic. All analyses were performed on the full analysis

set (FAS) population and included all randomized patients whereby the primary outcome was measured. No imputation was performed. All significance tests were two-sided and conducted at the 5% significance level. All statistical analyses were performed by using Statistical Analysis Software (SAS) (NC, USA).

Sample size calculation

Power analysis was based on the ongoing pregnancy data of D5 SET results of Year 2015 from the participating clinics. In order to detect an increase in ongoing pregnancies from 45% to 52% with 80% power with a two-sided Fisher's exact test at a significance level 0.05, a total of 828 women in each group is needed.

Results

Between June 2018 and March 2021, a total of 2254 patients were assessed for eligibility and 776 patients met the inclusion criteria and were randomized (Fig. 1). The study was stopped early owing to a slower rate of recruitment than expected: a main cause was coronavirus disease-19 (COVID-19) restrictions. Of the estimated number of patients needed, 46.9% (776/1656) were randomized.

After randomization, 31 patients did not receive the allocated intervention. In one case, the patient withdrew consent, in the remaining cases embryo transfer was cancelled and all GQB were cryopreserved.

In the FAS population ($n = 745$), blastocyst selection was either performed by morphology alone ($n = 376$) or by morphology and TL ($n = 369$). For two patients in the TL group, blastocyst selection was not performed as per protocol for technical reasons, but these two patients remained in the FAS analysis.

No significant differences were found for patient (Table I) or treatment characteristics (Table II) between the two study groups.

In the FAS analysis, the ongoing pregnancy rate was the same in the two groups: 47.4% (175/369) in the TL group and 48.1% (181/376) in the control group (mean difference -0.7% , 95% CI $-8.2, 6.7$, $P = 0.90$), RR 0.99, 95% CI 0.85, 1.15 (Table III) and OR 0.97, 95% CI 0.73, 1.30. No statistically significant differences were found for secondary outcomes: the pregnancy rate was 60.2% (222/369) in the TL group and 59.0% (222/376) in the control group (mean difference 1.1%, 95% CI $-6.2, 8.4$, $P = 0.81$), and early pregnancy loss was 21.2% (47/222) in the TL group and 18.5% (41/222) in the control group (mean difference 2.7%, 95% CI $-5.2, 10.6$, $P = 0.55$) (Table III).

When adjusting for all optimal allocation variables in a sensitivity analysis with multivariate logistic regression analyses, the OR for ongoing pregnancy was 0.91 (95% CI 0.68, 1.23, $P = 0.90$) between the TL group and the control group. The OR for pregnancy was 0.99 (95% CI 0.73, 1.34, $P = 0.94$) and 1.26 (95% CI 0.77, 2.06, $P = 0.56$) for early pregnancy loss between the TL and control group.

Pre-determined subgroup analyses were performed to compare ongoing pregnancy rates between the study groups for each treatment clinic (Supplementary Table SI), for maternal age (Supplementary Table SII) and for number of GQB available on Day 5 (Supplementary Table SIII). The comparisons showed no statistically significant differences for ongoing pregnancy rates between the study groups. From the logistic regression analysis, no significant interactions were found

between the studied baseline variables and the study group for prediction of ongoing pregnancy.

For the TL group, ongoing pregnancy rates for version 2 (46.8%, 95% CI 40.5, 53.2) and version 3 (47.8%, 95% CI 38.3, 57.4) of the KIDScore™ D5 algorithm were the same (mean difference -1.0% , 95% CI $-12.7, 10.7$, $P = 0.95$). KIDScore™ D5 version 1 was not included in the comparisons as it was only used for a few patients ($n = 6$).

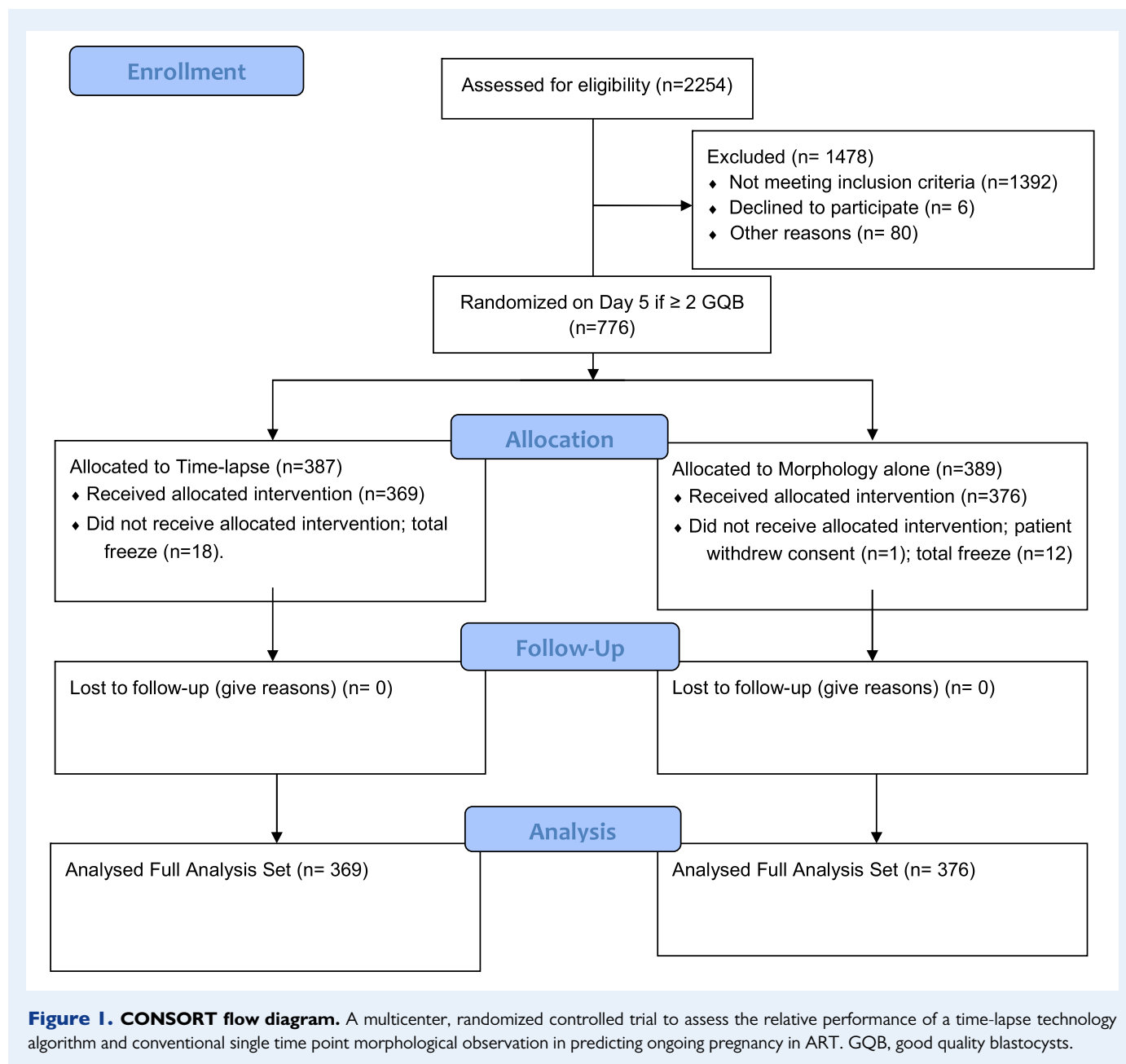
In the TL group, the choice of best blastocyst changed on 42% of occasions (154/369, 95% CI 36.9, 47.2) after the KIDScore™ D5 was applied and this rate was similar for most treatment clinics (Supplementary Table SIV). The ongoing pregnancy rates were the same between concordant and discordant groups for algorithm versions, v1 (66.7% versus 66.7%), v2 (46.7% versus 46.4%) and v3 (50.7% versus 41.0%) (Supplementary Table SV). KIDScore™ D5 version 1 was only used for a few patients ($n = 6$), making the comparison unreliable.

Discussion

This large multicenter RCT showed that the addition of TL model KIDScore™ D5 to select single blastocysts for Day 5 transfer did not improve ongoing pregnancy rate compared to selection of blastocysts by morphology alone. Furthermore, the study showed no differences for pregnancy rate and early pregnancy loss between the two groups.

Only two other RCTs have investigated the use of TL selection compared to morphology alone while incubating all embryos in a TL incubator. In agreement with our findings, both studies concluded that embryo selection through TL did not lead to a significant improvement of clinical outcomes (Goodman *et al.*, 2016; Kaser *et al.*, 2017). Unfortunately, both of these studies were underpowered to detect significant improvements after blastocyst transfer. In the study by Goodman *et al.* (2016) only 75% of randomized patients underwent blastocyst transfer, 91 patients in the TL group and 89 in the control group and no significant differences were found for the clinical pregnancy rate (73.6% versus 67.0%, respectively) and implantation rate (55.5% versus 51.2%, respectively). Kaser *et al.* (2017) terminated the study after 60% of the patients were randomized ($n = 163$) and found no significant difference in clinical pregnancy rates between using Early embryo viability assessment™ (Eeva™) test to select embryos on Day 3 (41.1%) or Day 5 (38.9%) when compared with selection by morphology alone on Day 5 (49.1%).

Similarly, our study, although considerably larger, only randomized about half (776/1656, 46.9%) of the patients we calculated were required to detect a superiority of TL greater than 7% for ongoing pregnancy rate ($\alpha = 0.05$, $\beta = 0.20$). With nearly half of the patients randomized, our study had the potential to detect a difference of 10% with 80% power at the 5% significance level between the two randomized groups. Even so, our findings show no indication that we would have found any significant improvement for ongoing pregnancy in the TL group even if we had randomized all intended patients. Thus, this questions the rationale of spending resources on such TL models unless there is a gain in workflow efficiency. However, new developments in TL monitoring, incorporating automatic systems and deep learning models, may provide potential advantages over current TL models (Ueno *et al.*, 2021). To date, Rubio *et al.* (2014) are the only other group that has performed a large RCT ($n = 843$) investigating



the potential benefits of TL. In contrast to our findings, their study reported a significant increase for ongoing pregnancy rate in the TL group compared to the control group (51.4% versus 41.7%, $P=0.005$). However, it is not apparent if the significant increase in the TL group was to the result of potential benefits of the hierarchical TL selection model used and/or the TL incubation system, as all embryos in the control group were cultured in a standard incubator (Rubio et al., 2014).

Even though KIDScore™ D5 was trained on a large heterogeneous multicenter dataset, it has previously been shown that several factors, including patient and treatment characteristics, affect morphokinetic timings and influence model performance when used at new clinical sites with different patient populations and culture conditions (Ciray et al., 2012; Munoz et al., 2012, 2013; Cruz et al., 2013; Freour et al., 2013; Kirkegaard et al., 2013). More recently, a study by Kato et al.

(2021) showed a relationship between KIDScore™ D5, maternal age and pregnancy and live birth outcomes, to suggest that KIDScore™ D5 prediction improved for patients with advanced age (>40 years old) (Kato et al., 2021). When we investigated the possibility of model performance varying within the dataset, our subgroup analyses found no specific subgroup that preferentially benefited from KIDScore™ D5 selection when stratifying for clinical site, maternal age and number of GQB available on Day 5. Although it should be noted that in comparison to Kato et al. (2021) few patients in our study were above 40 years of age. A further analysis investigating additional baseline variables failed to identify any confounders of KIDScore™ D5 performance for prediction of ongoing pregnancy compared to the control group. Possible effects of varying culture conditions were not studied because all clinics used the same media and TL system.

Table 1 Characteristics of the patients in the full analysis set according to study group.

Variable	Time-lapse group (n = 369)	Control group (n = 376)
Female age (years)	33.4 (4.2)	33.6 (4.1)
	33 (30; 37)	33 (30; 37)
Female height (cm)	168.2 (6.5)	167.7 (6.2)
	169 (164; 172)	168 (163; 172)
Female weight (kg)	70.4 (13.1)	69.6 (13.6)
	68 (61; 77)	67 (60; 78)
Female BMI (kg/m ²)	24.9 (4.4)	24.7 (4.4)
	24 (21.5; 27.7)	23.6 (21.5; 27.3)
Smoking female		
Non-smoker	357 (98.6%)	361 (97.6%)
<5 cigarettes/day	1 (0.3%)	7 (1.9%)
5–10 cigarettes/day	4 (1.1%)	1 (0.3%)
11–20 cigarettes/day	0 (0.0%)	1 (0.3%)
Taking snuff (snus) female		
Non-snuffer	348 (96.1%)	358 (96.8%)
Occasionally	5 (1.4%)	5 (1.4%)
Daily	9 (2.5%)	7 (1.9%)
Partner age (years)	35.8 (5.7)	35.6 (5.6)
	35 (31; 39)	35 (31; 39)
Smoking partner		
Non-smoker	311 (97.2%)	299 (94.9%)
<5 cigarettes/day	2 (0.6%)	8 (2.5%)
5–10 cigarettes/day	3 (0.9%)	2 (0.6%)
11–20 cigarettes/day	2 (0.6%)	5 (1.6%)
>20 cigarettes/day	2 (0.6%)	1 (0.3%)
Taking snuff (snus) partner		
Non-snuffer	251 (78.4%)	245 (77.8%)
Occasionally	11 (3.4%)	6 (1.9%)
Daily	58 (18.1%)	64 (20.3%)
Previous pregnancies in current relationship		
0	218 (59.1%)	220 (58.5%)
1	85 (23.0%)	85 (22.6%)
2	40 (10.8%)	44 (11.7%)
3	12 (3.3%)	12 (3.2%)
4	7 (1.9%)	9 (2.4%)
5 or more	7 (1.9%)	6 (1.6%)
Previous births in current relationship		
0	287 (77.8%)	280 (74.5%)
1	76 (20.6%)	90 (23.9%)
2	6 (1.6%)	5 (1.3%)
3	0 (0.0%)	1 (0.3%)
Number of started IVF cycles leading to oocyte retrieval	0.775 (1.288)	0.870 (1.622)
	0 (0; 1)	0 (0; 1)

(continued)

Table 1 Continued

Variable	Time-lapse group (n = 369)	Control group (n = 376)
Type of menstruation		
Regular	312 (85.0%)	323 (86.1%)
Irregular	55 (15.0%)	52 (13.9%)
Indication		
Hormonal (hypogonadotropic hypogonadism World Health Organization class I)	20 (5.4%)	11 (2.9%)
Male	83 (22.5%)	81 (21.6%)
Endometriosis	19 (5.1%)	21 (5.6%)
Cervical factor	1 (0.3%)	0 (0.0%)
Polycystic ovary syndrome (Rotterdam criteria)	19 (5.1%)	21 (5.6%)
Tubal factor	34 (9.2%)	22 (5.9%)
Same sex	18 (4.9%)	23 (6.1%)
Unexplained	139 (37.7%)	152 (40.5%)
Other reasons	29 (7.9%)	35 (9.3%)
Single	7 (1.9%)	9 (2.4%)
Origin of sperm		
Ejaculated	310 (84.0%)	308 (81.9%)
Donated	47 (12.7%)	58 (15.4%)
Percutaneous epididymal sperm aspiration (PESA)	2 (0.5%)	2 (0.5%)
Testicular sperm aspiration (TESA)	5 (1.4%)	5 (1.3%)
Testicular sperm aspiration extraction (TESE)	1 (0.3%)	0 (0.0%)
Thawed	4 (1.1%)	3 (0.8%)

For categorical variables, n (%) is presented.

For continuous variables, mean (SD), median (Q1; Q3) is presented.

Several retrospective studies have shown KIDScore™ D5 to have significant positive associations to implantation and live birth rates after single blastocyst transfer (Lee *et al.*, 2019; Reignier *et al.*, 2019; Kato *et al.*, 2021). In smaller studies, blastocysts with higher KIDScores were associated with increased glucose consumption on Day 5, a positive indicator of viability, and increasing euploidy rate (Ferrick *et al.*, 2020; Yap *et al.*, 2019). Certainly, evidence generated via retrospective analyses remains important, but as this study demonstrates, findings were not translated into true clinical value when validated in an RCT. As KIDScore™ D5 is a commercially available model protected by copyright, we are not privileged to know how annotations made for t2, t3, t4 t5 tSB and morphological grades for TE and ICM are used or weighted to calculate a score (i.e. timings to specific cell stages, cleavage cycle durations). This limits our ability to compare specific markers and timing intervals to other published studies. Certainly, there are dilemmas in developing TL models that are applicable to heterogeneous patient populations in different clinical settings. Many of the

Table II Treatment characteristics of the IVF cycles in the full analysis set according to study group.

Variable	Time lapse group (n = 369)	Control group (n = 376)	P-value
FSH start dose (IU)	192.1 (73.7)	195.0 (75.0)	0.60
Total FSH dosage (IU)	150 (150; 225)	150 (150; 225)	0.98
	1946 (845)	1947 (889)	
	1725 (1323; 2350)	1738 (1325; 2325)	
Number of stimulation days	10.2 (2.2)	9.98 (1.59)	0.16
	10 (9; 11)	10 (9; 11)	
Type of drug used for ovarian stimulation			
Bemfola	69 (18.7%)	72 (19.1%)	
Gonal-F	197 (53.4%)	201 (53.5%)	
Menopur	90 (24.4%)	87 (23.1%)	
Puregon	7 (1.9%)	8 (2.1%)	
Recombinant FSH	6 (1.6%)	8 (2.1%)	0.98
GnRH downregulation			
Agonist	39 (10.6%)	41 (10.9%)	
Antagonist	330 (89.4%)	335 (89.1%)	0.98
Number of oocytes at retrieval	12.1 (3.9)	12.1 (4.0)	0.90
	12 (9; 14)	11 (9; 14)	
Method of fertilization			
Standard IVF	227 (61.5%)	226 (60.1%)	
ICSI	108 (29.3%)	112 (29.8%)	
50/50	34 (9.2%)	38 (10.1%)	0.89
Number of two pronuclear zygotes	8.04 (2.81)	8.04 (2.75)	1.00
	7 (6; 10)	8 (6; 9.5)	
Number of blastocysts on Day 5	4.51 (2.20)	4.38 (2.03)	0.43
	4 (3; 5)	4 (3; 5)	
Number of good quality blastocysts on Day 5	3.59 (1.73)	3.40 (1.49)	0.12
	3 (2; 4)	3 (2; 4)	
Number of blastocysts on Day 6	2.08 (1.67)	2.27 (1.80)	0.17
	2 (1; 3)	2 (1; 3)	
	n = 368	n = 366	
Number of good quality blastocysts on Day 6	0.659 (0.947)	0.648 (0.900)	0.91
	0 (0; 1)	0 (0; 1)	
	n = 367	n = 366	
Number of cryopreserved embryos (Day 5 + 6)	3.28 (1.88)	3.19 (3.19)	0.70
	3 (2; 4)	3 (2; 4)	
	n = 368	n = 366	

For categorical variables, n (%) is presented.

For continuous variables, mean (SD)/median/(Q1; Q3)/n = is presented for variables when missing values for full analysis set population. For comparison between groups, Fisher's exact test (lowest one-sided P-value multiplied by 2) was used for dichotomous variables and the Mantel-Haenszel's chi-square exact test was used for ordered categorical variables and the chi-square exact test was used for non-ordered categorical variables and the Fisher's non-parametric permutation test was used for continuous variables.

models developed have not been transferable to other clinical settings without modifications (Kirkegaard et al., 2014; Freour et al., 2015; Ahlstrom et al., 2016; Ferrick, et al., 2020). Particular risks include defining too narrow or too broad time intervals for optimal divisions, which can change the balance between model specificity and sensitivity. It may also be of importance that several atypical cleavage or morphological features (chaotic cleavage, reverse cleavage, fragment internalization, multinucleation, blastocoel collapse)

indicated for deselection models were not evaluated in this study and our results cannot be generalized to these types of markers (Athayde Wirka et al., 2014; Desai et al., 2014; Goodman et al., 2016). Furthermore, aberrant division patterns associated with developmental arrest may be more valuable when selecting embryos for transfer at early stages of development. In comparison, the hierarchical model by Meseguer et al. (2011) used several exclusion criteria before assessing the timings of morphokinetic markers.

Table III Efficacy variables for the full analysis set population.

Variable	Time lapse group (n = 369)	Control group (n = 376)	P-value	Difference between groups Mean (95% CI)	Relative risk (95% CI)
Primary efficacy variable					
No. of ongoing pregnancies (%)	175 (47.4%) (42.2–52.7%)	181 (48.1%) (43.0–53.3%)	0.90	−0.7 (−8.2, 6.7)	0.985 (0.848, 1.145)
Secondary efficacy variables					
No. of positive β -hCG pregnancies (%)	222 (60.2%) (55.0–65.2%)	222 (59.0%) (53.9–64.1%)	0.81	1.1 (−6.2, 8.4)	1.019 (0.905, 1.147)
No. of early pregnancy losses (%)	47 (21.2%) (16.0–27.1%)	41 (18.5%) (13.6–24.2%)	0.55	2.7 (−5.2, 10.6)	1.146 (0.788, 1.668)

For categorical variables, n (%) and exact 95% CI is presented. For comparison between groups, Fisher's exact test (lowest one-sided P-value multiplied by 2) was used for dichotomous variables. The CI for dichotomous variables is the unconditional exact confidence limits. If no exact limits can be computed the asymptotic Wald confidence limits with continuity correction are calculated instead.

A proposed assumption is that an algorithm's potential to increase pregnancy rates is greater when the agreement with standard selection methods is low (Storr *et al.*, 2018). In a retrospective analysis, Reigner *et al.* (2019) found a moderate level of agreement, 61%, between selections by KIDScore™ D5 (v2) and morphology alone, and suggested that there was potential for improvement when applying KIDScore™ D5 in an RCT because of a positive association between KIDScores and live birth outcome. Similarly, we observed that the preliminary decision made by embryologists was in agreement with KIDScore™ D5 in 55% of cases for algorithm version 2, and in 65% of cases for version 3. The fact that clinical outcomes were the same between our study groups (control versus intervention) shows that the choice made by the embryologist was not inferior to the KIDScore™ D5 model and vice versa when selecting from a number of GQB.

As part of our study design, blastocysts in the control group were selected for transfer after a single static observation on Day 5. In practice, this may not reflect how many clinics perform routine morphological selection, instead it is more common to perform a number of static observations on multiple days and weight morphological events to improve ranking of viable embryos. We chose not to include a more extensive morphological ranking system in order to simplify and standardize selection methods between clinics. Positively, the ongoing pregnancy rate in the control group was comparable to published performance indicators and Nordic register data (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011), and did not affect our findings. However, in retrospect, since the TL intervention group included assessments also from earlier cleavage stages into the KIDScore™ grading, we may have given this group a small advantage enabling better assessment of blastocyst expansion as well as TE and ICM grade.

A strength of this study was that all patients underwent single blastocyst transfer (SET). This allowed us to know the direct outcome of each blastocyst and avoid the influence of confounding factors associated with the practice of double embryo transfer (DET) (Liu *et al.*, 2019). In comparison, previous studies performing multiple embryo transfers have failed to describe whether all embryos were optimal (according to the selection criteria) or the decision behind choosing DET, often practiced in cycles for poor prognosis patients, with poor

embryo quality and for patients with advanced maternal age (Rubio *et al.*, 2014; Goodman *et al.*, 2016; Liu *et al.*, 2019). Furthermore, several studies performing mixed transfer days were unable to indicate on which day the model performs best. Another strength of this study was the enrollment of patients from multiple clinical sites that strengthens the generalizability of our findings.

A possible weakness of this study is that we performed a manual pre-selection and only allowed blastocysts of good quality to be considered for selection by the models. It could be argued that using such a strict morphological criterion restricts the possibility for a theoretical selection power of TL to take place, as well as restricting the results mainly to a good prognosis patient group. We decided to use a higher threshold in order to ensure that patients would at least get a GQB transferred and not risk a 'substandard' embryo transfer. Indeed, the ideal study would have been to select from all available two pronuclear zygotes and really challenge morphological grading and what can be considered acceptable embryo quality before transfer, but we considered that unethical.

Another weakness of our study was caused by the COVID-19 pandemic, as many participating clinics temporarily closed down, greatly reducing the number of IVF cycles started or increasing the number of freeze-all cycles to postpone fresh transfers. This reduced the randomization of eligible patients and could influence the generalizability of our study.

The findings of our study, alongside many others, demonstrate the limitations of morphology and morphokinetic markers to fully reflect the reproductive potential of developing embryos, including the chromosomal status (Ziebe *et al.*, 2003; Alfarawati *et al.*, 2011). The implicated relaxation of cell cycle checkpoints, allowing embryo development to progress even in the presence of chromosomal errors, explains why cleavage patterns will be of little use to fully separate normal from abnormal embryos (Ambartsumyan and Clark, 2008). Future hopes now lie invested in the use of artificial intelligence technologies to interpret the full scope of TL images and identify key features that have been missed or not quantified in current models and without manual annotations (Tran *et al.*, 2019; VerMilyea *et al.*, 2020).

In conclusion, this large RCT showed that TL selection did not improve ongoing pregnancy rate after Day 5 single blastocyst transfer when compared with selection by morphology alone.

Supplementary data

Supplementary data are available at *Human Reproduction* online.

Data availability

The data underlying this article will be shared on reasonable request to the corresponding author.

Acknowledgements

The authors would like to thank statisticians Nils-Gunnar Pehrsson and Christopher Backström for their invaluable support and Per Lindquist who provided eCRF data services. We would also like to extend our gratitude to all coworkers at participating clinics for their ongoing support during the study period.

Authors' roles

K.L. T.H. and A.A. participated in the design of the study, enrollment of patients, analysis and interpretation of data, writing of the article and approval of the final version. A.K.L. and A.T.H. participated in the design of the study, enrollment of patients, analysis and interpretation of data and approval of the final version. All other authors have participated in the enrolment of patients, interpretation of data, revision of the article and approval of the final version.

Funding

The study was financed by a grant from the Swedish State under the ALF-agreement between the Swedish Government and the County Councils (ALFGBG-723141).

Conflict of interest

Vitrolife supported the study with embryo culture dishes and culture media. Vitrolife had no influence in the design of the study, analyses or writing of the manuscript. The manuscript was sent to Vitrolife before submission, but Vitrolife did not influence the content of the manuscript. During the study period, T.H. changed his employment from Livio AB to Vitrolife AB. All other authors have no conflicts of interests to disclose.

References

Ahlstrom A, Park H, Bergh C, Selleskog U, Lundin K. Conventional morphology performs better than morphokinetics for prediction of live birth after day 2 transfer. *Reprod Biomed Online* 2016;**33**:61–70.

Alfarawati S, Fragouli E, Colls P, Stevens J, Gutierrez-Mateo C, Schoolcraft WB, Katz-Jaffe MG, Wells D. The relationship between blastocyst morphology, chromosomal abnormality, and embryo gender. *Fertil Steril* 2011;**95**:520–524.

Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. *Hum Reprod* 2011;**26**:1270–1283.

Ambartsumyan G, Clark AT. Aneuploidy and early human embryo development. *Hum Mol Genet* 2008;**17**:R10–R15.

Arce JC, Ziebe S, Lundin K, Janssens R, Helmggaard L, Sorensen P. Interobserver agreement and intraobserver reproducibility of embryo quality assessments. *Hum Reprod* 2006;**21**:2141–2148.

Armstrong S, Bhide P, Jordan V, Pacey A, Marjoribanks J, Farquhar C. Time-lapse systems for embryo incubation and assessment in assisted reproduction. *Cochrane Database Syst Rev* 2019;**5**:CD011320.

Athayde Wirka K, Chen AA, Conaghan J, Ivani K, Gvakharia M, Behr B, Suraj V, Tan L, Shen S. Atypical embryo phenotypes identified by time-lapse microscopy: high prevalence and association with embryo development. *Fertil Steril* 2014;**101**:1637–1648. e1631–1635.

Beral V, Doyle P, Tan SL, Mason BA, Campbell S. Outcome of pregnancies resulting from assisted conception. *Br Med Bull* 1990;**46**:753–768.

Ciray HN, Aksoy T, Goktas C, Ozturk B, Bahceci M. Time-lapse evaluation of human embryo development in single versus sequential culture media—a sibling oocyte study. *J Assist Reprod Genet* 2012;**29**:891–900.

Ciray HN, Campbell A, Agerholm IE, Aguilar J, Chamayou S, Esbert M, Sayed S. Proposed guidelines on the nomenclature and annotation of dynamic human embryo monitoring by a time-lapse user group. *Hum Reprod* 2014;**29**:2650–2660.

Conaghan J, Chen AA, Willman SP, Ivani K, Chenette PE, Boostanfar R, Baker VL, Adamson GD, Abusief ME, Gvakharia M et al. Improving embryo selection using a computer-automated time-lapse image analysis test plus day 3 morphology: results from a prospective multicenter trial. *Fertil Steril* 2013;**100**:412–419. e415.

Craft I. Factors affecting the outcome of assisted conception. *Br Med Bull* 1990;**46**:769–782.

Cruz M, Garrido N, Gadea B, Munoz M, Perez-Cano I, Meseguer M. Oocyte insemination techniques are related to alterations of embryo developmental timing in an oocyte donation model. *Reprod Biomed Online* 2013;**27**:367–375.

De Croo I, Colman R, De Sutter P, Tilleman K. Blastocyst transfer for all? Higher cumulative live birth chance in a blastocyst-stage transfer policy compared to a cleavage-stage transfer policy. *Facts Views Vis Obgyn* 2019;**11**:169–176.

De Vos A, Van Landuyt L, Santos-Ribeiro S, Camus M, Van de Velde H, Tournaye H, Verheyen G. Cumulative live birth rates after fresh and vitrified cleavage-stage versus blastocyst-stage embryo transfer in the first treatment cycle. *Hum Reprod* 2016;**31**:2442–2449.

Desai N, Ploskonka S, Goodman LR, Austin C, Goldberg J, Falcone T. Analysis of embryo morphokinetics, multinucleation and cleavage anomalies using continuous time-lapse monitoring in blastocyst transfer cycles. *Reprod Biol Endocrinol* 2014;**12**:54.

Fernandez-Shaw S, Cercas R, Brana C, Villas C, Pons I. Ongoing and cumulative pregnancy rate after cleavage-stage versus blastocyst-stage embryo transfer using vitrification for cryopreservation: impact of age on the results. *J Assist Reprod Genet* 2015;**32**:177–184.

Ferrick L, Lee YSL, Gardner DK. Metabolic activity of human blastocysts correlates with their morphokinetics, morphological grade, KIDScore and artificial intelligence ranking. *Hum Reprod* 2020;**35**:2004–2016.

Freour T, Dessolle L, Lammers J, Lattes S, Barriere P. Comparison of embryo morphokinetics after *in vitro* fertilization-intracytoplasmic

- sperm injection in smoking and nonsmoking women. *Fertil Steril* 2013; **99**:1944–1950.
- Freour T, Le Fleuter N, Lammers J, Splingart C, Reignier A, Barriere P. External validation of a time-lapse prediction model. *Fertil Steril* 2015; **103**:917–922.
- Gardner DK, Schoolcraft WB. In vitro culture of human blastocyst. In Janson R, Mortimer D (eds). *Towards Reproductive Certainty: Infertility and Genetics Beyond 1999*. Carnforth: Parthenon Press, 1999, 378–388.
- Gissler M, Malin Silverio M, Hemminki E. In-vitro fertilization pregnancies and perinatal health in Finland 1991–1993. *Hum Reprod* 1995; **10**:1856–1861.
- Goodman LR, Goldberg J, Falcone T, Austin C, Desai N. Does the addition of time-lapse morphokinetics in the selection of embryos for transfer improve pregnancy rates? A randomized controlled trial. *Fertil Steril* 2016; **105**:275–285.e210.
- Kaser DJ, Bormann CL, Missmer SA, Farland LV, Ginsburg ES, Racowsky C. A pilot randomized controlled trial of day 3 single embryo transfer with adjunctive time-lapse selection versus day 5 single embryo transfer with or without adjunctive time-lapse selection. *Hum Reprod* 2017; **32**:1598–1603.
- Kato K, Ueno S, Berntsen J, Ito M, Shimazaki K, Uchiyama K, Okimura T. Comparing prediction of ongoing pregnancy and live birth outcomes in patients with advanced and younger maternal age patients using KIDScore day 5: a large-cohort retrospective study with single vitrified-warmed blastocyst transfer. *Reprod Biol Endocrinol* 2021; **19**:98.
- Kirkegaard K, Ahlstrom A, Ingerslev HJ, Hardarson T. Choosing the best embryo by time lapse versus standard morphology. *Fertil Steril* 2015; **103**:323–332.
- Kirkegaard K, Campbell A, Agerholm I, Bentin-Ley U, Gabrielsen A, Kirk J, Sayed S, Ingerslev HJ. Limitations of a time-lapse blastocyst prediction model: a large multicentre outcome analysis. *Reprod Biomed Online* 2014; **29**:156–158.
- Kirkegaard K, Hindkjaer JJ, Ingerslev HJ. Effect of oxygen concentration on human embryo development evaluated by time-lapse monitoring. *Fertil Steril* 2013; **99**:738–744. e734.
- Kovacs P, Matyas S, Forgacs V, Sajgo A, Molnar L, Pribenszky C. Non-invasive embryo evaluation and selection using time-lapse monitoring: results of a randomized controlled study. *Eur J Obstet Gynecol Reprod Biol* 2019; **233**:58–63.
- Lee CI, Chen CH, Huang CC, Cheng EH, Chen HH, Ho ST, Lin PY, Lee MS, Lee TH. Embryo morphokinetics is potentially associated with clinical outcomes of single-embryo transfers in preimplantation genetic testing for aneuploidy cycles. *Reprod Biomed Online* 2019; **39**:569–579.
- Liu Y, Feenan K, Chapple V, Matson P. Assessing efficacy of day 3 embryo time-lapse algorithms retrospectively: impacts of dataset type and confounding factors. *Hum Fertil (Camb)* 2019; **22**:182–190.
- Meseguer M, Herrero J, Tejera A, Hilligsoe KM, Ramsing NB, Remohi J. The use of morphokinetics as a predictor of embryo implantation. *Hum Reprod* 2011; **26**:2658–2671.
- Meseguer M, Rubio I, Cruz M, Basile N, Marcos J, Requena A. Embryo incubation and selection in a time-lapse monitoring system improves pregnancy outcome compared with a standard incubator: a retrospective cohort study. *Fertil Steril* 2012; **98**:1481–1489. e1410.
- Munoz M, Cruz M, Humaidan P, Garrido N, Perez-Cano I, Meseguer M. Dose of recombinant FSH and oestradiol concentration on day of HCG affect embryo development kinetics. *Reprod Biomed Online* 2012; **25**:382–389.
- Munoz M, Cruz M, Humaidan P, Garrido N, Perez-Cano I, Meseguer M. The type of GnRH analogue used during controlled ovarian stimulation influences early embryo developmental kinetics: a time-lapse study. *Eur J Obstet Gynecol Reprod Biol* 2013; **168**:167–172.
- Paternot G, Wetzels AM, Thonon F, Vansteenbrugge A, Willemen D, Devroe J, Debrock S, D'Hooghe TM, Spiessens C. Intra- and interobserver analysis in the morphological assessment of early stage embryos during an IVF procedure: a multicentre study. *Reprod Biol Endocrinol* 2011; **9**:127.
- Reignier A, Girard JM, Lammers J, Chtourou S, Lefebvre T, Barriere P, Freour T. Performance of day 5 KIDScore morphokinetic prediction models of implantation and live birth after single blastocyst transfer. *J Assist Reprod Genet* 2019; **36**:2279–2285.
- Rubio I, Galan A, Larreategui Z, Ayerdi F, Bellver J, Herrero J, Meseguer M. Clinical validation of embryo culture and selection by morphokinetic analysis: a randomized, controlled trial of the EmbryoScope. *Fertil Steril* 2014; **102**:1287–1294. e1285.
- Rubio I, Kuhlmann R, Agerholm I, Kirk J, Herrero J, Escriba MJ, Bellver J, Meseguer M. Limited implantation success of direct-cleaved human zygotes: a time-lapse study. *Fertil Steril* 2012; **98**:1458–1463.
- Storr A, Venetis C, Cooke S, Kilani S, Ledger W. Time-lapse algorithms and morphological selection of day-5 embryos for transfer: a preclinical validation study. *Fertil Steril* 2018; **109**:276–283.e3.
- Tran D, Cooke S, Illingworth PJ, Gardner DK. Deep learning as a predictive tool for fetal heart pregnancy following time-lapse incubation and blastocyst transfer. *Hum Reprod* 2019; **34**:1011–1018.
- Ueno S, Berntsen J, Ito M, Uchiyama K, Okimura T, Yabuuchi A, Kato K. Pregnancy prediction performance of an annotation-free embryo scoring system on the basis of deep learning after single vitrified-warmed blastocyst transfer: a single-center large cohort retrospective study. *Fertil Steril* 2021; **116**:1172–1180.
- VerMilyea M, Hall JMM, Diakiw SM, Johnston A, Nguyen T, Perugini D, Miller A, Picou A, Murphy AP, Perugini M. Development of an artificial intelligence-based assessment model for prediction of embryo viability using static images captured by optical light microscopy during IVF. *Hum Reprod* 2020; **35**:770–784.
- Westergaard HB, Johansen AM, Erb K, Andersen AN. Danish National In-Vitro Fertilization Registry 1994 and 1995: a controlled study of births, malformations and cytogenetic findings. *Hum Reprod* 1999; **14**:1896–1902.
- Yang L, Cai S, Zhang S, Kong X, Gu Y, Lu C, Dai J, Gong F, Lu G, Lin G. Single embryo transfer by day 3 time-lapse selection versus day 5 conventional morphological selection: a randomized, open-label, non-inferiority trial. *Hum Reprod* 2018; **33**:869–876.
- Yap WY, Lee CSS, Lim YX, Lim MW. Relationship between euploidy rates and D5 KIDScoreTM of blastocysts from embryoscope. *Reprod Biomed Online* 2019; **39**:e39.
- Ziebe S, Lundin K, Loft A, Bergh C, Nyboe Andersen A, Selleskog U, Nielsen D, Grondahl C, Kim H, Arce JC; CEMAS II and Study Group. FISH analysis for chromosomes 13, 16, 18, 21, 22, X and Y in all blastomeres of IVF pre-embryos from 144 randomly selected donated human oocytes and impact on pre-embryo morphology. *Hum Reprod* 2003; **18**:2575–2581.