

# Improving embryo selection by the development of a laboratory-adapted time-lapse model

Idit Blais, M.Sc.,<sup>a</sup> Mara Koifman, M.Sc.,<sup>a</sup> Ido Feferkorn, M.D.,<sup>a</sup> Martha Dirnfeld, M.D.,<sup>a,b</sup> and Shirly Lahav-Baratz, Ph.D.<sup>a</sup>

<sup>a</sup> Division of Reproductive Endocrinology and In Vitro Fertilization (IVF), Department of Obstetrics and Gynecology, Lady Davis Carmel Medical Center, Haifa, Israel, and <sup>b</sup> Ruth and Bruch Rappaport Faculty of Medicine, Technion – Israel Institute of Technology, Haifa, Israel

**Objective:** To study whether a powerful, in-house, embryo-selection model can be developed for a specific in vitro fertilization (IVF) laboratory where embryos were already selected for transfer using general models.

**Design:** In total, 12,944 fertilized oocytes were incubated in an EmbryoScope (Vitrolife, Göteborg, Sweden) at our laboratory. Embryos were selected for transfer or freezing using general models. There were 1,879 embryos with known implantation data (KID), of which 425 had positive KIDs. For the outcome, we set 3 endpoints for KID's definition: gestational sac, clinical pregnancy, and live birth. Results of a comparison between KID-positive and -negative embryos for cell division timings were analyzed separately for intracytoplasmic sperm injection (ICSI) and IVF embryos in patients aged 18–41 years.

**Setting:** IVF center.

**Patients:** The study included 1,075 women undergoing IVF or ICSI treatment between June 2013 and February 2019.

**Intervention(s):** None.

**Main Outcome Measure(s):** The KID-positive and -negative embryos were analyzed for statistical differences in cell division timing and cell cycle intervals. We used the EmbryoScope Stats software (Unisense FertiTech, Aarhus, Denmark) for model development. The statistically different timing parameters were tested for their contribution to scoring in the model. The algorithms were tested for area under the receiver operating characteristic curve (AUC) in the KID embryos for developing day-2, -3, and -5 embryo-selection models. The validation of these algorithms was performed using calibration/validation procedures.

**Results:** Because significant differences in morphokinetics were found between the KID-positive and KID-negative embryos in our laboratory, it was possible to use our specific KID data to develop an in-house model. The algorithms were developed for embryo selection on days 2, 3, and 5 in the ICSI embryos. In most cases, AUC was >0.65, which indicated that these models were valid in our laboratory. In addition, these AUC values were obtained from all gestational sac, clinical pregnancy, and live birth KID embryo databases tested. An increase in the predictability of the models was observed from days 2–3 to day 5 models. The AUC test results ranged between 0.657 and 0.673 for day 2 and day 3, respectively, and 0.803 for the day 5 model.

**Conclusion:** A model based on laboratory-specific morphokinetics was found to be complementary to general models and an important additive tool for improving single embryo selection. Developing an in-house laboratory-specific model requires many stages of sorting and characterization. Many insights were drawn about the model developing process. These may facilitate and improve the process in other laboratories. (Fertil Steril Sci® 2021;2:176–97. ©2021 by American Society for Reproductive Medicine.)

**Key Words:** In vitro fertilization, time-lapse, embryo-selection algorithm, laboratory-adapted model

**Discuss:** You can discuss this article with its authors and other readers at <https://www.fertstertdialog.com/posts/xfss-d-20-00059>

**E**mbryo incubation and assessment is a vital step in assisted reproductive technology. Traditionally, embryo assessment is performed by removing embryos from a

conventional incubator daily for morphologic evaluation and viewing them under a light microscope (1). In recent years, time-lapse systems (TLS) have been developed. These systems

capture digital images of embryos at frequent time intervals and allow continuous assessment of the embryos' quality without physically removing them from the incubator. One potential advantage of TLS includes the ability to maintain a stable culture environment by limiting the exposure of embryos to changes in gas composition and temperature. Another main potential advantage of TLS includes its ability to improve embryo selection for assisted reproductive technology treatment

Received August 27, 2020; revised February 8, 2021; accepted February 10, 2021.

I.B. has nothing to disclose. M.K. has nothing to disclose. I.F. has nothing to disclose. M.D. has nothing to disclose. S.L.-B. has nothing to disclose.

Reprint requests: Idit Blais, M.Sc., Division of Reproductive Endocrinology and In Vitro Fertilization (IVF), Lady Davis Carmel Medical Center, 7 Michal St., Haifa, Israel, 3436212 (E-mail: [IditBl@clalit.org.il](mailto:IditBl@clalit.org.il) or [IditBlais@gmail.com](mailto:IditBlais@gmail.com)).

Fertil Steril Sci® Vol. 2, No. 2, May 2021 2666-335X/\$36.00

© 2021 American Society for Reproductive Medicine. Published by Elsevier Inc. All rights reserved.

<https://doi.org/10.1016/j.xfss.2021.02.001>

using additional information gained from continuous monitoring of embryo development (2). Multiple clinical trials have shown that the TLS technology, compared with the traditional evaluation, can improve clinical outcomes (3–19). A study conducted in our laboratory showed that applying a TLS model, including morphokinetic parameters, for embryo evaluation in addition to the classical morphology scoring improved IVF outcomes in terms of embryo implantation and clinical pregnancy rate compared with the use of TLS as an incubator without using a model (unpublished data).

The recognition of health risks and financial costs associated with multiple pregnancies has been a driving force for IVF clinics to move toward elective single embryo transfer. The adoption of an elective single embryo transfer policy demands the optimization of not only culture conditions but also embryo-selection or deselection techniques that increase the probability of selecting an embryo with the greatest implantation potential (20). According to fertility societies and the ministry of health guidelines in many countries, only 1 embryo should be transferred in young patients (defined differently in each country). Thus, improving the selection of the best embryo is a major challenge.

Diverse algorithms, using embryo cleavage kinetics available in TLS technology-based imaging, have been proposed in the literature for embryo selection (3, 4, 7, 8, 14, 21–28). Rienzi et al. (29) investigated the morphodynamic characterization of euploid blastocyst development under specific culture conditions at 2 different centers. They found that some morphokinetic timings of euploid embryos cultured at IVF center 2 were consistently slower compared with the timing at IVF center 1 (average 1–2 hours). The study suggested that morphokinetic timing was laboratory-specific.

Various studies have looked into whether reported embryo-selection models are transferable between IVF clinics. Several laboratories have applied the model reported by Meseguer et al. in 2011 (4) to an independent data set (30–32). The model proved ineffective in predicting pregnancy. For example, Best et al. (30) showed the calculated area under the receiver operating characteristic (ROC) curve (AUC) to be 0.61, indicating poor predictive power and ineffectiveness of this model on independent data. When different variables were applied to the model reported by Meseguer et al. (4), AUC of 0.75 was achieved, indicating improved classification power. Fréour et al. (32) also found the model reported by Meseguer et al. (4) to be of low sensitivity in predicting the implantation rate of embryos. However, a simplified version of this model showed acceptable performance. They suggested that morphokinetic time ranges, representing optimum implantation, varies among clinics. Barrie et al. (33) studied the efficacy of 6 reported embryo-selection algorithms when applied to a large, exclusive, set of known implantation embryos. When applied to this data, all 6 examined algorithms achieved AUC of <0.65, indicating reduced predictive capability. The algorithms were also ineffective in embryo classification in terms of implantation rates and aneuploidy risk. This can lead to a potentially misleading result for embryo selection.

The investigators claimed that the algorithms were developed under environmental parameters available in a specific laboratory and, thus, were clinically relevant to that laboratory alone. For external application, the algorithms lost their predictive capabilities (33). The investigators suggested that each IVF laboratory should determine its embryo-selection criteria based on its own data and develop their own selective algorithms (30–33).

To overcome the differences among laboratories, generally applicable morphokinetic algorithms were developed for embryo selection, such as KIDScore D3 for day-3 transferred embryos [as described by Petersen et al. (26, 34)] and KIDScore D5 for day-5 transferred embryos (13, 35). These algorithms are routinely used in IVF laboratories using an EmbryoScope TLS incubator (Vitrolife, Göteborg, Sweden) for the selection and frequent deselection of embryos from a cohort. Broad applicability of these algorithms was achieved using a large and diverse data set for calibration, cell cycle duration instead of actual timings where possible, which compensated for differences relating to incubation conditions and fertilization method, and robust parameterization methods using cross validation.

The KIDScore D3 model enables the deselection of embryos with the lowest probability for implantation, leaving a rather large proportion of embryos available for selection and, thus, the final choice of transfer or freezing to the embryologist. As stated by the investigators, KIDScore enables clinics that have not yet built a large data foundation to support their decision regarding which embryo(s) to transfer immediately after implementing the technology, whereas making the final decision always depends on the embryologist (26). Reigner et al. (13) performed an ROC curve analysis for 2 versions of KIDScore day 5 (versions 1 and 2) and demonstrated that both the models were significant predictors of implantation but with rather low AUCs of 0.59 and 0.60, respectively. They concluded that KIDScore might be considered an objective second opinion on embryo ranking, helping embryologists in making their clinical decision but not replacing them.

The studies presented above showed that there is a need for a laboratory-adapted model for embryo selection in addition to general models used mainly for the deselection of embryos. We report here, for the first time, the process of creating a laboratory-adapted model in a laboratory where embryos were already selected for transfer using general models. We believe that this information will be beneficial for all laboratories that perform time-lapse assessment and embryo selection using general models.

## METHODS

### Stimulation Protocol

All women included were treated with either a gonadotropin-releasing hormone (GnRH) agonist or antagonist protocol. In the agonist protocol, downregulation was achieved using a GnRH agonist (0.2 mg of decapeptyl, Ferring Pharmaceuticals, Saint-Prex, Switzerland) in the midluteal phase, whereas ovarian stimulation was achieved using recombinant follicle-stimulating hormone (FSH) (GONAL-f; Merck Serono,

Geneva, Switzerland, or Puregon; Schering-Plough, Kenilworth, NJ), human Menopausal Gonadotropin (hMG) (Menopur, Ferring, Langley, United Kingdom), or a combination of recombinant FSH and luteinizing hormone (Pergoveris; Merck Serono).

In the antagonist protocol, stimulation by gonadotropins (GT) was performed, as described above. A GnRH antagonist (Orgaluran 0.25 mg, Merck Serono) was administered from the 6<sup>th</sup> day of the cycle. Final follicular maturation was triggered using recombinant human chorionic gonadotropin (Ovitrelle; Merck Serono) when at least 2 leading follicles reached a mean diameter of 18 mm. A GnRH agonist trigger (0.2 mg of decapeptyl) was used to avoid ovarian hyperstimulation syndrome in women with a very high response (>20 follicles on trigger day and/or 17 $\beta$ -estradiol level >10,000 pmole/L), and all embryos were cryopreserved.

Oocytes were collected 36–37 hours later by transvaginal ultrasound-guided needle aspiration of the follicles (day 0).

### Fertilization and Embryo Culture

After retrieval, cumulus-oocyte complexes collected were placed in a human tubal fluid medium (Quinn's advantage fertilization [human tubal fluid] Medium, Sage, Trumbull, U.S.A.) supplemented with 10% serum protein substitute under an oil (Sage) overlay and incubated for 2–3 hours at 37°C in 6% CO<sub>2</sub> and 5% O<sub>2</sub>.

Fertilization was performed using IVF or intracytoplasmic sperm injection (ICSI). During the IVF procedure, 200,000 sperm cells were added to  $\leq$ 6–7 oocytes per well in a 4-well plate (Nunc, Roskilde, Denmark). After 3–4 hours of incubation, which was a part of a short-exposure IVF procedure, or overnight incubation, which was a part of a standard IVF procedure, cumulus and coronal cells were mechanically removed. Before ICSI, cumulus cells surrounding the oocytes were removed by enzymatic digestion with hyaluronidase (Sage). The IVF- and ICSI-fertilized oocytes were placed in individual micro wells within a pre-equilibrated specific culture slide (EmbryoSlide; Vitrolife). Each well was filled with 25  $\mu$ l of a culture medium (Global total, Vitrolife) and overlaid with oil (Sage). The slides were loaded into the EmbryoScope (Vitrolife), a trigas incubator with a built-in microscope that allows the time-lapse monitoring of early embryo development. The embryos were cultured at 37°C in 6% CO<sub>2</sub> and 5% O<sub>2</sub>. Images of each embryo were acquired automatically every 15 minutes on 7 focal planes.

The embryos were selected for transfer or cryopreservation after 2, 3, or 5 days of incubation. Embryo transfer (ET) was performed under transabdominal ultrasound guidance using soft catheters (Wallace; Cooper Surgical, Trumbull, U.S.A.). Embryo cryopreservation was performed by vitrification (Sage).

### Embryo Selection

The selection of the embryos for transfer and/or freezing was performed using general morphokinetic algorithm models, combined with morphologic evaluation of the embryos.

Since we started using the TLS incubator in 2013, we have used the “Carmel modified Alpha/ESHRE” algorithm as a general model for embryo selection (modified in our laboratory as described in [Supplemental Appendix 1](#) and [Supplemental Table 1](#), available online). Once the general applicable KIDScore models were available, we also used KIDScore D3 basic (Vitrolife) to select day-3 embryos (since June 2014) and KIDScore D5 V1.2 (Vitrolife) to select day-5 embryos (version 1 since April 2016 and version 2 since May 2017).

The selection of the embryos in the TLS incubator in our laboratory was preliminary based on the general models “Carmel modified Alpha/ESHRE” and/or the score of the general KIDScore models, and completed by evaluating the morphology of the embryos.

The morphology was assessed using a common embryological practice, as described by the Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology (1). The morphologic evaluation parameters for day 2–3-embryos included cell number, percentage of fragmentation, symmetry, and granularity of the blastomeres. For day-5 embryos, the morphologic evaluation parameters included degree of expansion quality of the inner cell mass and trophoctoderm.

### Embryo Image Annotations

The EmbryoViewer software (version 7.3.200.16739, Vitrolife) was used for annotation of the timing of embryo blastomeric divisions. The references for the recorded timings were time (t) in hours post insemination (hpi). Start time (t<sub>0</sub>) was defined as the time at which sperm was added to oocytes (IVF) or sperm was injected into an oocyte (ICSI). All the blastomeric division timings of all embryos were annotated.

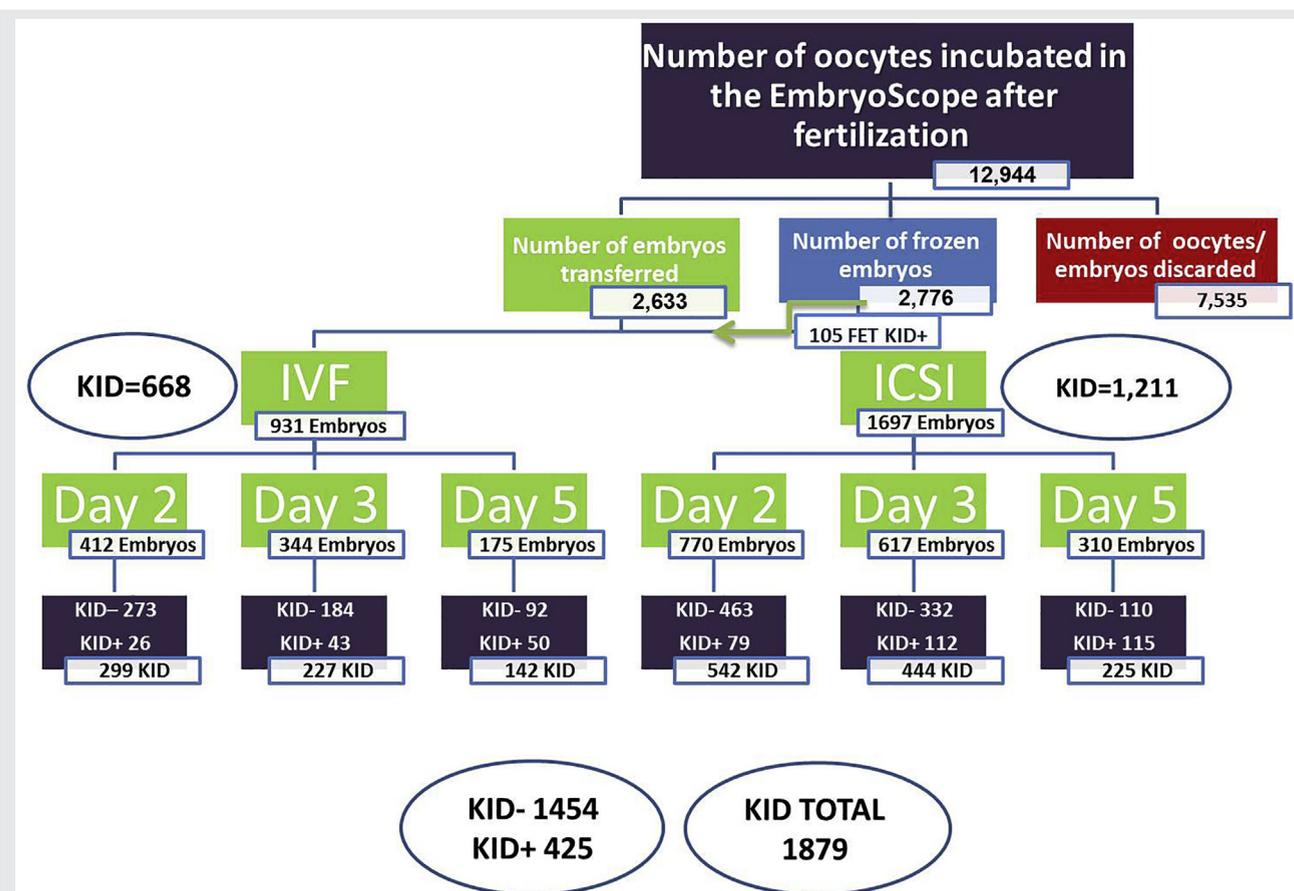
These annotations were performed in accordance with the consensus criteria published in “Proposed guidelines on the nomenclature and annotation” (36) and the Vitrolife Technotes for KIDScore D3 (34) and KIDScore D5 version 3 (35). In general, the fading time of pronuclei (PN) was annotated using the first image after both PNs disappeared (tPNf); blastomeric divisions were annotated using the first image showing complete division (t<sub>2</sub>, t<sub>3</sub>, t<sub>4</sub>, t<sub>5</sub>, t<sub>6</sub>, t<sub>7</sub>, t<sub>8</sub>, and t<sub>9+</sub>); morula was annotated at the end of compaction process (tM); the start of blastulation was set as the first frame in which initiation of cavity formation was observed (tSB); annotation of a full blastocyst was performed using the last image before the embryo started to compress the zona pellucida (tB); expanded blastocyst was annotated when the zona pellucida was 50% thinner (tEB); and herniation blastocyst was annotated using the first frame in which extraction of cells from the zona was observed (tHB).

The annotations were performed by the same 3 senior embryologists in a uniform manner and tested for uniformity by 1 embryologist, who double checked the annotations of all the embryos.

### First Step: Data Collection

A retrospective data analysis included data of 12,944 fertilized oocytes from 1,075 patients aged 18–45 years. All the

FIGURE 1



Dispersion of fertilized oocytes incubated in the EmbryoScope according to selection results. The figure shows the number of fertilized oocytes that developed into embryos and were selected for transfer (green), the number of oocytes that developed into embryos and were selected for freezing (blue), and the number of fertilized oocytes discarded (red). All transferred embryos, including frozen and thawed KID-positive GS embryos (FET KID+) were sorted according to fertilization method (ICSI or IVF). The IVF and ICSI embryos were sorted according to transfer day (2, 3, or 5) and implantation outcome [KID-positive (+) or -negative (-)]. The circuit diagram shows the total number of KID-positive and -negative GS embryos for IVF and ICSI and the total number of GS KID. KID = known implantation data; FET= frozen embryo transfer; GS = gestational sacs; ICSI = intracytoplasmic sperm injection; IVF = in vitro fertilization.

Blais. Laboratory-adapted time-lapse model. Fertil Steril Sci 2021.

fertilized oocytes were incubated in the TLS incubator, EmbryoScope, between June 12, 2013, and February 14, 2019. No preimplantation genetic diagnosis cycles were included. The data were obtained from the Carmel Medical Center IVF Unit after the local institutional review board's approval.

The annotated data were exported from the EmbryoViewer software to Microsoft Office Excel sheet (Microsoft Office Professional Plus 2010). Only embryos that had a full sequence of photographs, from the moment when 2 PN were observed to the moment of collection for transfer or freezing, were included. We found that to get true and reliable data and a list of accurate cleavage timings, some adaptations were required after exporting the data to the Excel sheet, as described in [Supplemental Appendix 1](#).

The data used for developing the algorithm was obtained from embryos with known implantation data (KID). According to the KID's definition, if 1, 2, or more embryos were transferred, only the cases where either all or none of the embryos implanted were used (8, 16, 21, 25, 26, 37).

For the outcome, we set 3 endpoints for the KID embryo definition: gestational sac (GS KID), clinical pregnancy (positive fetal heartbeat in an ultrasound test) (CP KID), and live birth (LB KID). Cycles showing chemical pregnancy ( $\beta$ -human chorionic gonadotropin level  $>5$  U/L 12 days after ET, which later decreased without GS observed using ultrasound) and ectopic pregnancy were excluded.

The frozen embryos were marked for their exact location in the liquid nitrogen container to enable the tracking of implantation results following embryo thawing and transfer. To avoid bias in the dataset, we neutralized the possible effect of the freezing and thawing technique by taking into account only the thawed KID-positive embryos.

The data of the 12,944 fertilized oocytes were sorted according to the embryo-selection results (transferred, frozen, or discarded embryos), fertilization method (ICSI or IVF), ET day (2, 3, or 5), and implantation outcome for the KID embryos (KID-positive or -negative) (Fig. 1). After sorting, 1,879 embryos remained, of which 1,454 were GS

TABLE 1

## Definitions of the equations used to test cell cycle intervals.

Variable name	Equation	Definition	References of sources used these variables
cc2	t3-t2	The time of second cell cycle (duration of the period as 2-cell stage)	(3, 4, 6–9, 14, 16, 17, 20–24, 37–39, 41–47, 50)
s2	t4-t3	The time of synchrony of second cell cycle (duration of the period as 3-cell stage)	(3, 4, 6, 7, 9, 16, 17, 20–24, 29, 36–39, 41–44, 46, 47)
vPN3	t3-tPNf	Duration of the period from pronuclear fading to 3-cell stage	(21, 22, 26)
Equation B	$(t5-t3) / (t5-t2)$	Cell cycle interval that can be used for detecting embryos with irregular divisions	(26)
cc3	t5-t3	Duration of the third cell cycle	(4, 14, 20, 37–39, 42–44, 46, 50)
t5-t2	t5-t2	Duration of the period from 2-cell to 5-cell stage	(20, 37, 39, 42, 44)
vPN5	t5-tPNf	Can be used for detecting embryos, including those with either the 2- or 4-cell stage <5 h	(21)
s3	t8-t5	Time for synchrony of the third cell cycle	(4, 16, 20, 29, 36–39, 41–43, 46)
tM-t9+	tM-t9+	Duration of the period from 9-cell stage to morula	
tB-tSB	tB-tSB	Duration of the period from the start of blastulation to blastocyst	(14, 16, 27, 36, 39)
tB-tM	tB-tM	Duration of the period from morula to blastocyst	
tB-t9	tB-t9	Duration of the period from 9-cell stage to blastocyst	
s1, vPN	t2-tPNf	Duration of the period as 1-cell stage	(22, 37)

Blais. Laboratory-adapted time-lapse model. *Fertil Steril Sci* 2021.

KID-negative and 425 were GS KID-positive. The distribution of the embryos by group is shown in [Figure 1](#).

### Second Step: Analysis of Statistical Differences Between KID-Positive and -Negative Embryos

The KID-positive and -negative embryos were analyzed for statistical differences in cell division timing and cell cycle intervals, which was measured based on the time between 2 consecutive cell divisions, helping in the quantification of cell synchrony (18).

The cell division timing included all 14 timings annotated, as described above: tPNf, t2, t3, t4, t5, t6, t7, t8, t9+, tM, tSB, tB, tEB, and tHB.

The cell cycle intervals included 13 parameters, for which calculation was done to examine the time interval between different cell division timings. These relationships can be represented more generally for any time interval (ti) using the following formula:  $ti = ty - tx$ , where y is a more advanced developmental stage, and x is a defined referent that is always an earlier developmental stage (38). The 13 intervals chosen to be tested for differences between the KID-positive and -negative embryos are described in [Table 1](#).

The GS KID-positive and -negative embryos were compared to determine differences in the characteristics of the patient population. After limiting the age of women to 41 years, no statistical differences were found in the population parameters between the KID-positive and -negative embryo groups. The characteristics of the patients are shown in [Table 2](#).

The main parameters that we expected to affect embryo cleavage timing were fertilization method and women's age. Each GS KID-positive or -negative embryo group was divided into subgroups according to the type of fertilization method (IVF or ICSI) and women's age (<35 years or  $\geq 35$  years). Differences between the subgroups were tested for all cell division timings and cell cycle intervals. As shown in [Supplemental Appendix 2](#) (available online), the fertilization method had a clear effect on the embryo cleavage timing; therefore, the IVF and ICSI embryos were separated for developing the model. No differences were found between short- and conventional-exposure IVF embryos. We found that women's age had no effect on the embryo division timing; hence, the different age subgroups were consolidated into 1 group ([Supplemental Tables 2a-b, 3a-b and 4 in Supplemental Appendix 2](#)).

After setting the abovementioned parameters, the differences between the KID-positive and -negative embryos in terms of cell division timings and cell cycle intervals were tested separately for ICSI and IVF in women aged 18–41 years. The results of a comparison of the timing between the KID-positive and -negative embryos were analyzed using an independent *t*-test or the Mann-Whitney test for continuous variables, as appropriate. Moreover, the correlation between the timing variables and implantation outcome was analyzed using logistic regression (odds ratio with 95% confidence interval [CI]). The abovementioned comparison analysis was conducted using 3 databases: GS, CP, and LB KID embryos.

TABLE 2

## Comparison of the characteristics of the patient population.

Characteristics	ICSI			IVF		
	KID +	KID -	P value	KID +	KID -	P value
GS - KID						
No. of embryos	276	516		116	272	
Woman's age (y) <sup>a</sup>	31.3 ± 4.9	31.7 ± 4.0	.221	34.3 ± 4.7	34.6 ± 4.2	.467
Total amount of FSH dose (U) <sup>a</sup>	1,830 ± 831	1,893 ± 945	.702	1,996 ± 1077	1,997 ± 970	.546
Indication for cause of infertility (no.)			.490			.459
Male factor	184	329		-	-	
Tubal factor	23	61		27	79	
Unexplained infertility	31	63		25	61	
Other factors	34	63		63	132	
CP - KID						
No. of embryos	<b>258</b>	<b>521</b>		<b>102</b>	<b>281</b>	
Woman's age (y) <sup>a</sup>	31.1 ± 4.8	31.6 ± 3.9	.089	33.7 ± 4.7	34.6 ± 4.2	.072
Total amount of FSH dose (U) <sup>a</sup>	1,773 ± 809	1,900 ± 945	.205	1,976 ± 1,037	2,008 ± 962	.418
Indication for cause of infertility (no.)			.066			.533
Male factor	181	331		-	-	
Tubal factor	15	59		24	79	
Unexplained infertility	28	66		23	68	
Other factors	34	65		55	134	
LB - KID						
No. of embryos	<b>178</b>	<b>486</b>		<b>69</b>	<b>206</b>	
Woman's age (y) <sup>a</sup>	30.7 ± 4.5	31.2 ± 3.7	.167	32.5 ± 4.5	33.3 ± 4.1	.094
Total amount of FSH dose (U) <sup>a</sup>	1,727 ± 801	1,865 ± 938	.217	1,702 ± 906	1,827 ± 916	.192
Indication for cause of infertility (no.)			.145			.356
Male factor	126	315		-	-	
Tubal factor	10	48		16	63	
Unexplained infertility	16	62		16	52	

Note: KID = known implantation data; ICSI = intracytoplasmic sperm injection; IVF = in vitro fertilization; GS = gestational sac; CP = clinical pregnancy; LB = live birth; FSH = follicle-stimulating hormone.

<sup>a</sup> Mean ± standard deviation. The table shows the results of statistical analysis of the difference between KID-positive (+) and -negative (-) embryos in ICSI and IVF embryos.

Blais. Laboratory-adapted time-lapse model. Fertil Steril Sci 2021.

Additionally, to describe the distribution of the probabilities of implantation per cell division timing and cell cycle interval (GS KID embryos), the timings were converted from continuous variables into categorical variables by dividing them into groups based on their quartiles [as described earlier by Meseguer et al. (4)]. We then calculated the percentage of embryos implanted for each timing quartile to assess the distribution of implantation in the different categories. “*P* for trend” was used to identify a dose response effect. A chi-square test with Bonferroni correction was used for pairwise comparisons between the different quartiles. For each timing variable, an optimal range was defined as the combined range spanned by the quartiles with the significantly highest implantation rates. The categories defined as significant in the quartile test were used to establish optimal ranges of the cell division timings in the model development process.

### Statistical Methods

Statistical analysis was performed using the IBM statistics software (SPSS) version 24. *P* < .05 was considered statistically significant. The continuous variables were presented as mean and standard deviation or as median and interquartile range. The categorical variables were presented as percentages. Difference in the demographic and clinical characteristics between the KID-positive and -negative embryos were compared using an independent *t*-test or the

Mann-Whitney test for the continuous variables and chi-square test for the categorical variables. Correlation between the variables and implantation result was analyzed using logistic regression (odds ratio with 95% CI).

### Third Step: Developing a Model Algorithm

For developing the in-house models, we used the EmbryoScope Stats software (BETA version, A/S – version 0.8.0.0, Unisense FertilTech, Aarhus, Denmark). The EmbryoScope Stats software was developed as a beta version by Unisense FertilTech for the building, evaluation, and validation of models on the basis of retrospective data exported from the EmbryoViewer software database [also used by Basile et al. (8)]. The data are exported to a Microsoft Office Excel sheet and saved in a fixed file format, which the software recognizes. The sheet created consists of retrospective morphokinetic data obtained from the time-lapse monitoring of embryo development in an EmbryoScope TLS incubator. The adjustments we were required to make for the software to correctly identify the KID data and for developing the model are described in Supplemental Appendix 1.

Developing a model using the EmbryoScope Stats software is based on the ROC curve test. Receiver operating curve is used to evaluate the classification power of a model by plotting true positive rate as a function of false positive rate and calculating ROC AUC. As also described by Petersen et al. (26),

AUC is a commonly used quantifier of the overall predictive capability of an algorithm. Algorithms with zero predictive capability have an AUC value of 0.5 on an average, whereas those with perfect prediction have an AUC value of 1 (26). We set a minimum criterion of an AUC value of 0.65 for a valid model for clinical use.

As stated before, the embryos in this study were selected for transfer using the general models. The AUC values of the general KIDScore models for the KID embryos in our laboratory were calculated to examine the likelihood of distinguishing between embryos that actually resulted in birth and those that did not and verify the need to develop a laboratory-adapted model. The AUC values of KIDScoreD3 V1.2 and KIDScoreD5 V3.1 were tested on our KID data for embryos transferred on days 3 and 5, respectively. The test was first performed on all the KID embryos according to the day of transfer and was then performed again only on KID embryos with cycles in which it was necessary to choose between several embryos for transfer. The aim was to test the ability of the general models to distinguish between good- and poor-quality embryos.

### Calibration/Validation of the Developed Algorithms of the In-House Model

The algorithms were developed separately for each fertilization method (IVF or ICSI) and ET day (2, 3, and 5 days). For each transfer day, the algorithm type was tested for 3 different endpoints in the algorithms calibration/validation tests; GS KID embryos, CP KID embryos, and LB KIDs databases.

For day 2 and day 3 algorithms, calibration/validation tests were performed using 2 types of databases: “All KIDs” (including all KID embryos database available) and “Age  $\leq 41$  KIDs” (including only the database of KID embryos of women aged  $\leq 41$  years). For day 5 algorithm, the calibration/validation tests were performed using 3 types of databases: “All KIDs” and “Age  $\leq 41$  KIDs” (including data from embryos transferred on days 2, 3, or 5) and “ET Day 5 KIDs” (using only the data from embryos transferred on day 5).

Each KID embryo database uploaded to the EmbryoScope Stats software was automatically randomly divided into 2 datasets: training and holdout data. The training data were used to build and evaluate the model algorithm, and the holdout data were used to validate the model.

The algorithm was verified by performing 4 calibration procedures on each database. Each database was calibrated after randomly excluding a 25%, 33%, 40%, or 50% holdout dataset at a time (calibration subset). The part of the data that was excluded (holdout data) was then used for validation (validation subset). The AUC values were calculated both for the calibration (AUC calibration) and validation subsets (AUC validation). When the curve form was very different with the holdout data compared with that with the training data or when AUC was  $< 0.65$ , the model algorithm was adjusted by testing different variables, time ranges, and weight of the variables (as described in [Supplemental Appendix 1](#)).

## RESULTS

The cell division timing was found to be significantly different between the KID-positive and -negative embryos. The results of the comparison of cell division timings among the GS, CP, and LB KID ICSI and IVF embryos are presented in [Figures 2–4](#) and [Supplemental Appendix 3](#) (available online) ([Supplemental Table 5](#), available online). The mean and standard deviation obtained for the timings tested are shown in [Supplemental Appendix 3](#) ([Supplemental Table 5](#)). The 4 quartiles for the timing of each investigated parameter are presented in [Table 3](#), together with the percentage of implanted embryos in each quartile (GS KID embryos).

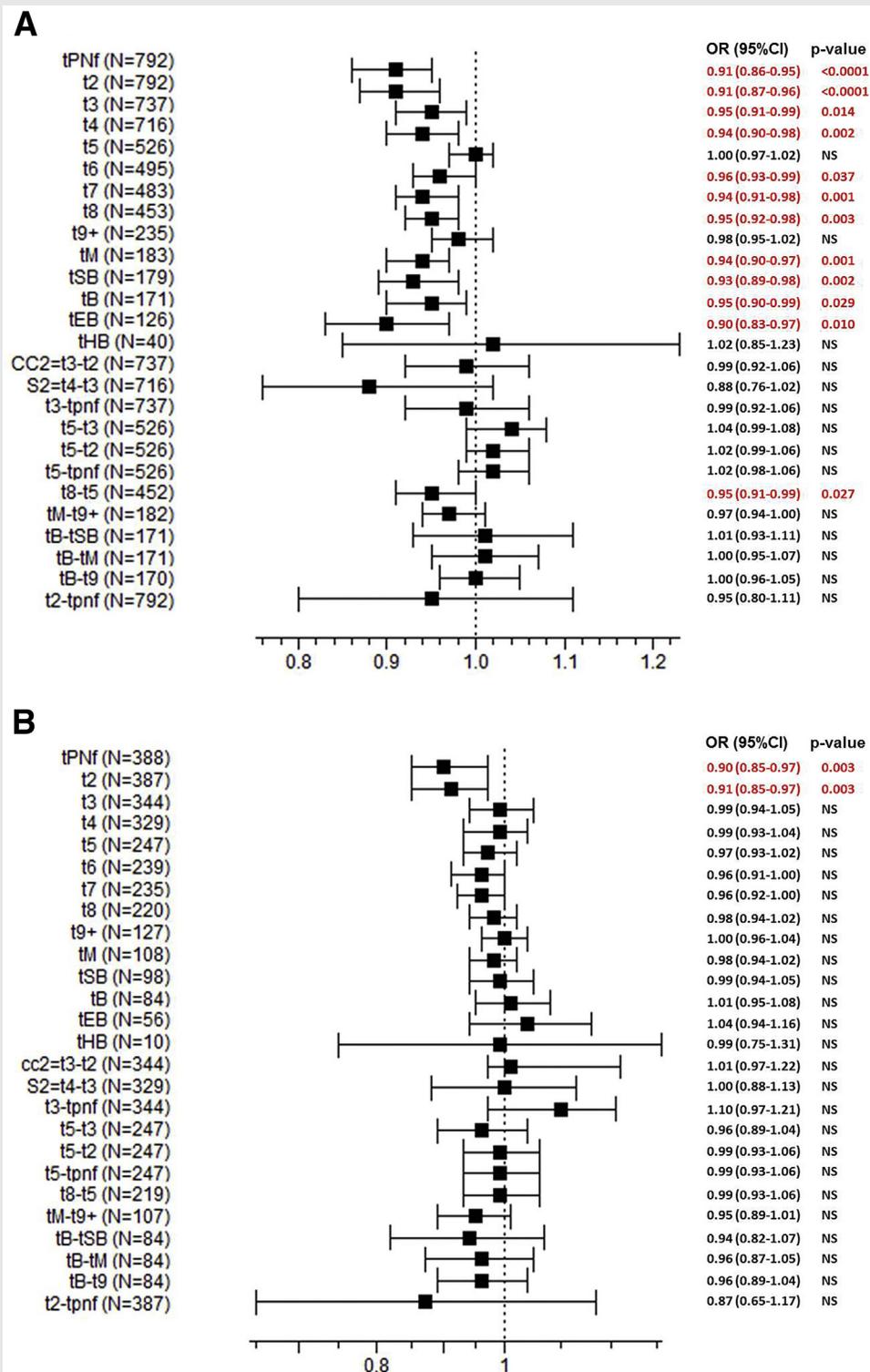
For the ICSI embryos, significant differences in the cell division timing between the KID-positive and -negative embryos were found throughout the GS, CP, and LB KID embryos at all the timings tested, except for t5, t9+, and tHB ([Figs. 2–4A](#) and [Supplemental Appendix 3](#) [[Supplemental Table 5](#)]). Significant differences were found between the quartiles for embryo implantation likelihood at all cell division timings, except for tB and tHB. The likelihood of embryo implantation was significantly higher in the first quartiles (Q1 and Q2) and, in some cases, in the third quartile (Q3) ([Table 3](#)). Differences in tHB timing were not statistically significant possibly because of the small number of tHB KID embryos for comparison.

For the IVF embryos, significant differences in the cell division timing between the KID-positive and -negative embryos were found only in the timings tPNF and t2 in the GS and CP KID embryos. There were no significant differences in the division timings of the IVF LB KID embryos ([Figs. 2–4B](#) and [Supplemental Appendix 3](#) [[Supplemental Table 5](#)]). Differences between the quartiles were found only in the tPNF and t2 cell division timings, where the likelihood of embryo implantation was significantly higher in the first 3 quartiles of timing for these parameters (Q1, Q2, and Q3) ([Table 3](#)).

The cell cycle interval timing analysis showed no differences between the KID-positive and -negative embryos at most intervals examined, except for t5-t3/t5-t2 and t8-t5 intervals in the GS KID ICSI embryos ([Figs. 2–4](#) and [Table 4](#)). Interval t5-t3/t5-t2 is shown in a separate table ([Table 4](#)) because the line range exceeded the graph size in [Figures 2–4](#). The independent *t*-test or Mann-Whitney test showed differences in cc2 (second cell cycle) interval in the ICSI LB KID embryos and t5-t3/t5-t2 interval in the IVF GS and CP KID embryos (but not in LB KID IVF embryos). The mean and standard deviation obtained for the timings tested are shown in [Supplemental Appendix 3](#) ([Supplemental Table 6](#), available online).

Differences between the quartiles for cell cycle interval timing were found in t3-tPNF, t5-t3, t5-t2, t5-tPNF, t8-t5, and t2-tPNF cell cycle intervals in the ICSI GS KID embryos. In the IVF GS KID embryos, differences between the quartiles were found at t3-tPNF and t5-t3/t5-t2 intervals. The likelihood of embryo implantation was significantly higher in the middle quartiles (Q2 and Q3) and in 1 case in Q1 (t5-t3/t5-t2 in the IVF KID embryos) ([Table 5](#)).

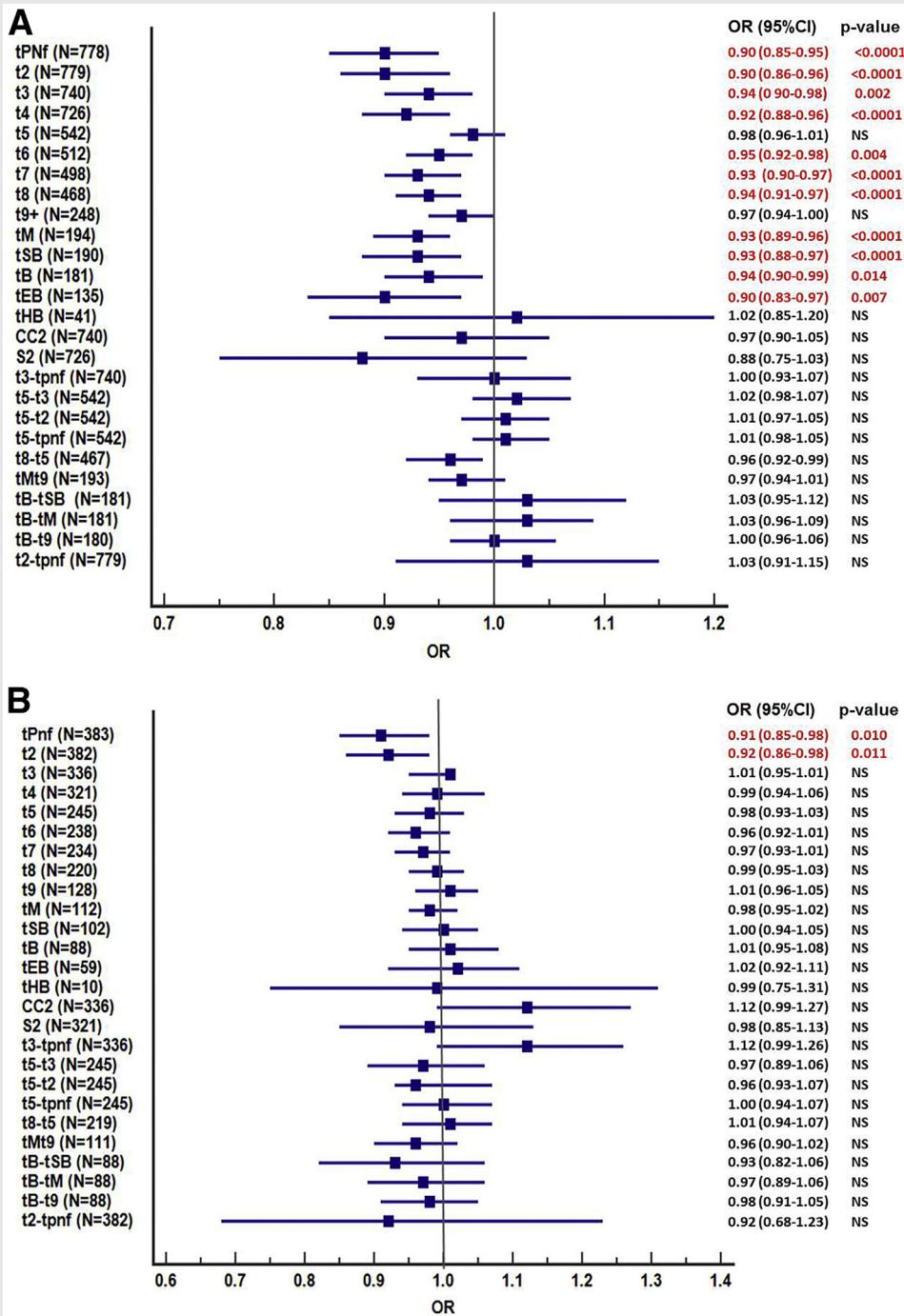
FIGURE 2



Univariate logistic regression analysis for the prediction of implantation in GS KID ICSI (A) and IVF (B) embryos for cell division timing and cell cycle intervals. Significant timings are marked in red. NS = nonsignificant; KID = known implantation data; GS = gestational sacs; ICSI = intracytoplasmic sperm injection; IVF = in vitro fertilization.

Blais. Laboratory-adapted time-lapse model. Fertil Steril Sci 2021.

**FIGURE 3**



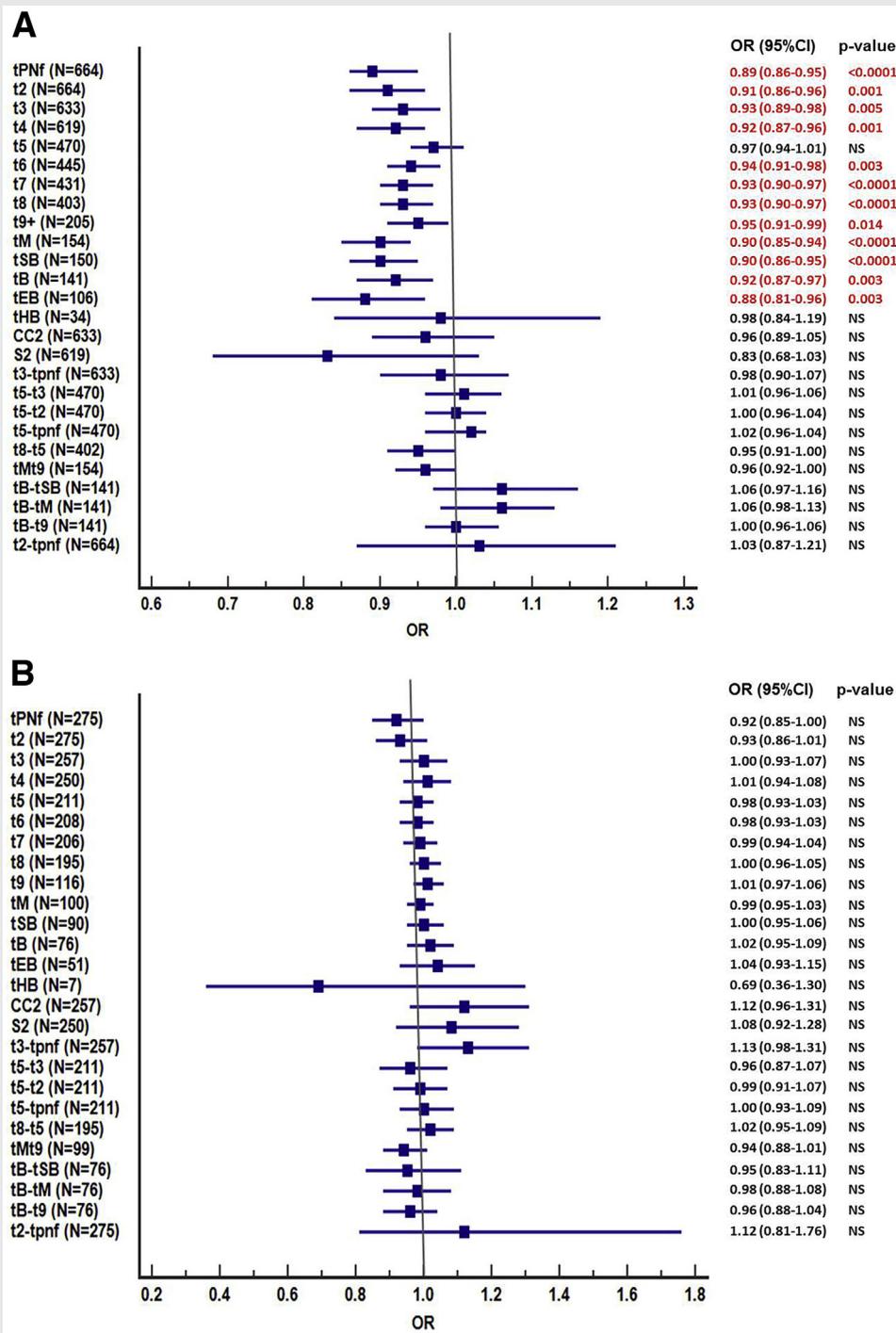
Univariate logistic regression analysis for the prediction of pregnancy in CP KID ICSI (A) and IVF (B) embryos for cell division timing and cell cycle intervals. Significant timings are marked in red. NS = nonsignificant; CP = clinical pregnancy; KID = known implantation data; ICSI = intracytoplasmic sperm injection; IVF = in vitro fertilization.

Blais. Laboratory-adapted time-lapse model. *Fertil Steril Sci* 2021.

Interesting results were observed for t5 cell cleavage timing. No significant differences were found in cleavage time t5 between the KID-positive and -negative embryos. Even for ICSI embryos that led to live birth, almost all the cell division timings were found to be significantly different

between the KID-positive and -negative embryos; the only difference in timing that was not found significant was at t5 (tBH was also not found to be statistically significant, most probably because of the small number of embryos in the group) (Figs. 2-4, Supplemental Appendix 3, and

FIGURE 4



Univariate logistic regression analysis for the prediction of live birth in LB KID ICSI (A) and IVF (B) embryos for cell division timing and cell cycle intervals. Significant timings are marked in red. NS = nonsignificant; LB = live birth; KID = known implantation data; ICSI = intracytoplasmic sperm injection; IVF = in vitro fertilization.

Blais. Laboratory-adapted time-lapse model. Fertil Steril Sci 2021.

TABLE 3

Cell division timing grouped into quartiles (Q1, Q2, Q3, and Q4) from ICSI and IVF GS KID embryos.

		Q1		Q2		Q3		Q4		P for trend
	Variable (no.)	Limit (hpi)	Implantation (%)	Limit (hpi)	Implantation (%)	Limit (hpi)	Implantation (%)	Limit (hpi)	Implantation (%)	
ICSI	tpnf (792)	≤22.24	<b>86 (43%)*</b>	<b>22.25–24.08</b>	<b>76 (38%)*</b>	24.09–26.36	64 (32%)	>26.36	50 (25%)	.001
	t2 (792)	≤24.68	<b>84 (42%)*</b>	<b>24.69–26.6</b>	<b>80 (40%)*</b>	26.61–28.97	63 (32%)	>28.97	49 (25%)	.001
	t3 (737)	≤35.94	<b>76 (41%)*</b>	<b>35.95–38.15</b>	<b>77 (42%)*</b>	38.16–40.75	62 (34%)	>40.75	52 (28%)	.019
	t4 (716)	≤36.56	<b>77 (43%)*</b>	<b>36.57–38.84</b>	<b>73 (41%)*</b>	38.85–41.52	64 (36%)	>41.52	48 (27%)	.007
	t5 (526)	≤47.75	46 (35%)	<b>47.76–51.66</b>	<b>66 (50%)*</b>	<b>51.67–55.44</b>	<b>62 (47%)*</b>	>55.44	36 (28%)	<.0001
	t6 (495)	≤49.8	<b>60 (49%)*</b>	<b>49.81–53.03</b>	<b>58 (47%)*</b>	53.04–56.69	50 (40%)	>56.69	37 (30%)	.015
	t7 (483)	≤51.51	<b>68 (56%)*</b>	<b>51.52–54.6</b>	<b>53 (44%)*</b>	54.61–58.27	47 (39%)	>58.27	37 (31%)	.001
	t8 (453)	≤52.48	<b>61 (54%)*</b>	<b>52.49–56.0</b>	<b>55 (49%)*</b>	56.01–60.28	45 (40%)	>60.28	39 (35%)	.018
	t9 (235)	≤65.38	29 (49%)	<b>65.39–70.03</b>	<b>36 (61%)*</b>	<b>70.04–74.86</b>	<b>35 (59%)*</b>	>74.86	22 (38%)	.046
	tM (183)	≤79.76	<b>32 (70%)*</b>	<b>79.77–84.96</b>	<b>31 (67%)*</b>	84.97–91.7	19 (41%)	>91.70	19 (42%)	.004
	tSB (179)	≤90.7	<b>33 (73%)*</b>	90.71–96.6	27 (60%)	96.61–100.8	22 (49%)	>100.8	19 (43%)	.022
	tB (171)	≤100.65	29 (67%)	100.66–104.97	28 (65%)	104.98–110.2	23 (51%)	>110.2	21 (53%)	.285
	tEB (126)	≤106.13	<b>23 (72%)*</b>	<b>106.14–110.1</b>	<b>22 (71%)*</b>	110.2–113.4	21 (66%)	>113.4	13 (42%)	.048
	tHB (40)	≤107.85	6 (60%)	107.86–110.66	9 (90%)	110.67–113.3	9 (82%)	>113.3	6 (67%)	.413
IVF	tpnf (388)	≤24.29	<b>33 (34%)*</b>	<b>24.30–26.30</b>	<b>34 (35%)*</b>	<b>26.31–28.75</b>	<b>33 (34%)*</b>	>28.76	16 (17%)	.011
	t2 (387)	≤26.72	<b>34 (35%)*</b>	<b>26.73–28.96</b>	<b>32 (33%)*</b>	<b>28.97–31.46</b>	<b>35 (36%)*</b>	>31.47	15 (16%)	.005
	t3 (344)	≤37.67	25 (29%)	37.68–39.81	32 (37%)	39.82–42.24	34 (40%)	>42.25	21 (24%)	.120
	t4 (329)	≤38.45	28 (34%)	38.46–40.39	26 (32%)	40.40–42.72	33 (40%)	40.40–42.72	25 (31%)	.558
	t5 (247)	≤51.03	24 (39%)	51.04–54.57	27 (44%)	54.58–58.66	24 (39%)	>58.67	17 (28%)	.326
	t6 (239)	≤52.02	26 (43%)	52.03–55.89	24 (40%)	55.90–59.64	24 (40%)	>59.65	17 (29%)	.386
	t7 (235)	≤54.08	25 (42%)	54.09–57.61	26 (44%)	57.62–61.31	23 (39%)	>61.32	16 (28%)	.254
	t8 (220)	≤54.76	23 (42%)	54.77–58.72	23 (42%)	58.73–62.73	19 (34.5)	>62.74	20 (36%)	.630
	t9+ (127)	≤68.68	13 (41%)	68.69–74.30	17 (53%)	74.31–79.94	15 (47%)	>79.95	12 (39%)	.650
	tM (108)	≤84.25	11 (41%)	84.26–89.65	14 (52%)	89.66–95.40	14 (52%)	>95.41	8 (30%)	.292
	tSB (98)	≤92.90	12 (48%)	92.91–98.50	14 (58%)	98.51–103.30	9 (36%)	>103.31	12 (50%)	.474
	tB (84)	≤100.31	10 (48%)	100.32–106.21	12 (57%)	106.22–111.30	13 (59%)	>111.31	9 (45%)	.749
	tEB (56)	≤108.22	8 (57%)	108.23–111.15	8 (57%)	111.16–113.30	7 (47%)	>113.31	9 (69%)	.694
	tHB (10)	≤106.59	2 (67%)	106.60–109.65	1 (50%)	109.66–113.27	1 (33%)	>113.28	1 (50%)	>.99

Note: % = percentage of implanting embryos in each quartile, hpi = hours post insemination; GS = gestational sacs; KID = known implantation data; ICSI = intracytoplasmic sperm injection; IVF = in vitro fertilization.

Timings that were found to be statistically significant were compared between the different quartiles (chi-square test with Bonferroni correction was used for pairwise comparisons). The quartiles showing a significantly higher implantation rate are marked with an asterisk and *in bold*.

The quartiles showing a significantly higher implantation rate are marked with an asterisk.

Blais. Laboratory-adapted time-lapse model. *Fertil Steril Sci* 2021.

TABLE 4

Univariate logistic regression analysis for cell cycle interval t5-t3/t5-t2 for the prediction of GS, CP, and LB KID in ICSI and IVF embryos.

Variable	No.	ICSI		No.	IVF		
		OR	95% CI		P value	OR	95% CI
GS							
t5-t3 / t5-t2	526	4.65	(1.17–18.5)	.029	247	0.35 (0.02–5.19)	.442
CP							
t5-t3 / t5-t2	542	3.5	(0.90–14.1)	.069	245	0.51 (0.03–9.2)	.649
LB							
t5-t3 / t5-t2	470	3.1	(0.66–14.4)	.151	211	0.39 (0.01–15.2)	.613

Note: GS = gestational sacs; CP = clinical pregnancy; LB = live birth; KID = known implantation data; ICSI = intracytoplasmic sperm injection; IVF = in vitro fertilization; OR = odds ratio; CI = confidence interval.

Significant timings are marked in bold.

Blais. Laboratory-adapted time-lapse model. Fertil Steril Sci 2021.

Table 5). However, t5 timing was found to have an optimal range in the quartile test and was part of most cell cycle intervals, which were found to be significantly different (Figs. 2–4, Table 4, and Supplemental Appendix 3 [Supplemental Table 5]).

### AUC Values of the General KIDScoreD3 and KIDScoreD5 Models in our Laboratory

The AUC values (95% CI) accepted for the KIDScoreD3 general model tested on all day-3 transferred embryos were 0.643 (0.589–0.696) and 0.641 (0.583–0.698) for the CP KID (n = 438) and LB KID (n = 410) embryos, respectively. For the KIDScoreD5 general model tested on all day-5 transferred embryos, the AUC values (95% CI) were 0.644 (0.580–0.707) and 0.651 (0.585–0.718) for the CP KID (n = 286) and LB KID (n = 258) embryos, respectively.

Lower AUC values (95% CI) were accepted for testing the KID embryos of cycles in which it was necessary to choose between several embryos for transfer, so that the AUC result could be considered as showing random prediction capability. For the KIDScoreD3 general model tested on select day-3 transferred embryos, the AUC values (95% CI) were 0.583 (0.509–0.658) and 0.572 (0.493–0.652) for the CP KID (n = 224) and LB KID (n = 204) embryos, respectively. For the KIDScoreD5 general model tested on select day-5 transferred embryos, the AUC values (95% CI) were 0.590 (0.513–0.666) and 0.607 (0.527–0.687) for the CP KID (n = 214) and LB KID (n = 189) embryos, respectively.

### In-House Model Algorithm Development and Validation

Because significant differences were found in cell division timing between KID-positive and -negative embryos, it was possible to develop an algorithm based on these differences. The algorithm was developed using only the cleavage timings and cell cycle intervals of the ICSI KID embryos, which showed clear and strong statistical differences between the KID-positive and -negative embryos.

Three algorithms were developed separately to fit the ET day: day 2 ICSI Carmel model, day 3 ICSI Carmel model, and day 5 ICSI Carmel model.

Great effort was invested in determining the combination of variables that gave the best AUC value and shape of the ROC curve for the model. The score weight and timing range values of the selected variables were also adjusted accordingly. The calibration/validation results of the algorithms for the models on different database files (“All KIDs,” “Age ≤41 KIDs,” and “ET Day 5 KIDs”) are shown in Tables 6–8. Once the right combination of variables was identified, only minor adjustments (weight and timing range of the variables) in the algorithm were needed for its implementation to different database files. Based on this result, we assumed that the algorithms developed were suitable for a wide range of embryos in our laboratory.

An increase in the predictability of the models was observed from the day 2–3 to day 5. The AUC values obtained ranged between 0.657 and 0.673 for the day 2 and day 3 models, respectively, and 0.803 for the day 5 model. For the day 2 model, the AUC calibration value obtained was 0.662 ± 0.011 (mean ± SD), ranging from 0.634 to 0.677. The AUC validation value for the day 2 model was 0.650 ± 0.033 (mean ± SD), ranging from 0.586 to 0.704. For the day 3 model, the AUC calibration value obtained was 0.679 ± 0.014 (mean ± SD), ranging from 0.656 to 0.714. The AUC validation value for the day 3 model was 0.665 ± 0.025 (mean ± SD), ranging from 0.632 to 0.722. For the day 5 model, the AUC calibration value obtained was 0.803 ± 0.023 (mean ± SD), ranging from 0.751 to 0.848. The AUC validation value for the day 5 model was 0.802 ± 0.040 (mean ± SD), ranging from 0.673 to 0.872 (Tables 6–8).

Area under the characteristic curve was tested for all the databases: GS, CP, and LB KID embryo databases. Although the number of KID embryos in the LB KID database was relatively smaller compared with that in the GS KID database, we obtained high values of AUC. Area under the characteristic curve for all calibration and validation cases was approximately ≥0.65, which indicated that these models were valid for use in our laboratory (Tables 6–8).

An example of the results obtained in one of the runs is presented in Figure 5 (for day 5 ICSI Carmel model, LB KID

TABLE 5

Cell cycle intervals grouped into quartiles (Q1, Q2, Q3, and Q4) from ICSI and IVF GS KID embryos.

	Variable (no.)	Q1		Q2		Q3		Q4		P for trend
		Limit (hpi)	Implantation (%)	Limit (hpi)	Implantation (%)	Limit (hpi)	Implantation (%)	Limit (hpi)	Implantation (%)	
ICSI	CC2 (737)	≤11	69 (37%)	11.01–11.76	74 (40%)	11.77–12.75	70 (38%)	>12.75	54 (29%)	.140
	S2 (716)	0	98 (36%)	0.01–0.5	33 (40%)	0.51–1.0	77 (43%)	>1	54 (30%)	.077
	t3-tpnf (737)	≤13.5	67 (36%)	<b>13.5–14.47</b>	<b>84 (46%)*</b>	14.48–15.26	58 (32%)	>15.26	58 (32%)	<b>.014</b>
	t5-t3 (526)	≤12.02	44 (33%)	<b>12.03–13.55</b>	<b>66 (50%)*</b>	<b>13.56–15.25</b>	<b>59 (45%)*</b>	>15.25	41 (31%)	<b>.003</b>
	t5-t2 (526)	≤23.03	46 (35%)	<b>23.04–25.63</b>	<b>70 (53%)*</b>	25.64–27.51	53 (40%)	>27.51	41 (31%)	<b>.001</b>
	t5-t3/t5-t2 (526)	≤0.51	48 (37%)	0.52–0.54	50 (38%)	0.55–0.56	66 (50%)	>0.57	46 (35%)	.053
	t5-tpnf (526)	≤25.53	44 (33%)	<b>25.54–28.09</b>	<b>72 (55%)*</b>	28.1–30.01	53 (40%)	>30.01	41 (31%)	<b>&lt;.0001</b>
	t8-t5 (452)	≤1.76	52 (46%)	1.77–3.25	46 (41%)	<b>3.26–5.75</b>	<b>61 (54%)*</b>	>5.75	40 (35%)	<b>.034</b>
	tM-t9 (94)	≤8.51	28 (62%)	8.52–13.28	25 (54%)	13.29–18.3	25 (56%)	>18.3	22 (48%)	.628
	tB-tSB (171)	≤7.49	24 (56%)	7.5–9.3	30 (70%)	9.31–11.73	27 (63%)	>11.73	20 (48%)	.191
	tB-tM (171)	≤17.3	26 (61%)	17.31–19.76	27 (63%)	19.77–23.2	22 (51%)	>23.2	26 (61%)	.675
	tB-t9 (170)	≤29.67	26 (61%)	29.68–33.05	23 (55%)	33.06–37.52	28 (65%)	>37.52	23 (55%)	.724
	t2-tpnf (792)	≤2.25	61 (31%)	<b>2.26–2.5</b>	<b>81 (41%)*</b>	<b>2.51–2.8</b>	<b>77 (39%)*</b>	>2.81	57 (29%)	<b>.026</b>
	IVF	CC2 (344)	≤10.75	23 (27%)	10.76–11.51	30 (35%)	11.52–12.26	35 (41%)	>12.27	24 (28%)
S2 (329)		0.0	39 (31%)	0.1–0.5	16 (41%)	0.51–0.75	29 (35%)	>0.76	28 (34%)	.723
t3-tpnf (344)		≤13.26	21 (24%)	<b>13.27–14.01</b>	<b>38 (44%)*</b>	14.02–15.01	26 (30%)	>15.02	27 (31%)	<b>.043</b>
t5-t3 (247)		≤12.62	25 (40%)	12.63–14.00	24 (39%)	14.01–15.51	21 (34%)	>15.52	22 (36%)	.886
t5-t2 (247)		≤23.76	21 (34%)	23.77–25.51	27 (44%)	25.52–27.74	25 (40%)	>27.75	19 (31%)	.461
t5-t3/t5-t2 (247)		≤ <b>0.52</b>	<b>27 (44%)*</b>	<b>0.53–0.55</b>	<b>26 (42%)*</b>	<b>0.56–0.57</b>	<b>26 (42%)*</b>	>0.58	13 (21%)	<b>.031</b>
t5-tpnf (247)		≤26.26	20 (32%)	26.27–28.14	30 (48%)	28.15–30.26	22 (36%)	>30.27	20 (33%)	.208
t8-t5 (219)		≤1.75	19 (35%)	1.76–3.25	25 (46%)	3.26–5.75	19 (34%)	>5.76	23 (43%)	.536
tMt9 (107)		≤9.25	12 (46%)	9.26–12.25	13 (46%)	12.26–17.26	15 (56%)	>17.27	7 (27%)	.199
tB-tSB (84)		≤6.90	12 (57%)	6.91–8.63	11 (52%)	8.64–11.25	9 (43%)	>11.26	9 (43%)	.766
tB-tM (84)		≤15.13	13 (62%)	15.14–18.75	9 (43%)	18.76–21.55	12 (57%)	>21.56	10 (48%)	.591
tB-t9 (84)		≤26.39	12 (57%)	26.40–31.01	10 (48%)	31.02–34.04	12 (57%)	>34.05	10 (48%)	.858
t2-tpnf (387)		≤2.25	26 (27%)	2.26–2.50	31 (32%)	2.51–2.89	36 (37%)	>2.90	23 (24%)	.192

Note: % = percentage of implanting embryos in each quartile, hpi =hours post insemination; GS = gestational sacs; KID = known implantation data; ICSI = intracytoplasmic sperm injection; IVF = in vitro fertilization.

Intervals that were found to be statistically significant were compared between the different quartiles (chi-square test with Bonferroni correction was used for pairwise comparisons). The quartiles showing a significantly higher implantation rate are marked with an asterisk and in bold.

The quartiles showing a significantly higher implantation rate are marked with an asterisk.

Blais. Laboratory-adapted time-lapse model. Fertil Steril Sci 2021.

TABLE 6

## Algorithm calibration/validation results for day 2 ICSI Carmel model.

Database for developing the algorithm	Data subset used for validation (%)	Calibration subset (no.)	Validation subset (no.)	AUC calibration	AUC validation
ICSI day 2 GS					
All KIDs (KID neg. 912)	25	906	302	0.654	0.678
KID pos. 296)	33	809	399	0.655	0.673
	40	725	483	0.644	0.692
	50	604	604	0.659	0.667
Age ≤41 y KIDs (KID neg. 652)	25	709	236	0.668	0.610
	33	633	312	0.675	0.610
KID pos. 293	40	567	378	0.663	0.643
	50	472	473	0.669	0.644
ICSI day 2 CP					
All KIDs (KID neg. 923)	25	885	295	0.667	0.668
KID pos. 257)	33	791	386	0.666	0.672
	40	708	472	0.659	0.685
	50	590	590	0.680	0.655
Age ≤41 y KIDs (KID neg. 663)	25	688	229	0.677	0.615
	33	614	303	0.680	0.624
KID pos. 254)	40	550	367	0.666	0.661
	50	458	459	0.674	0.656
ICSI day 2 LB					
All KIDs (KID neg. 940)	25	838	279	0.650	0.685
KID pos. 177)	33	748	369	0.649	0.683
	40	670	447	0.634	0.704
	50	558	559	0.655	0.664
Age ≤41 y KIDs (KID neg. 680)	25	643	214	0.668	0.586
	33	574	283	0.673	0.593
KID pos. 177)	40	514	343	0.664	0.619
	50	428	429	0.669	0.624

Note: ICSI = intracytoplasmic sperm injection; AUC = area under the characteristic curve; KID = known implantation data; GS = gestational sac; CP = clinical pregnancy; LB = live birth; neg = negative; pos = positive.

Tests were performed using the GS, CP, and LB KID embryos databases: "All KID embryos" and "KID embryos from women aged ≤41 years." The AUC values are shown for both the scored calibration subset (AUC calibration) and validation subset (AUC validation).

Blais. Laboratory-adapted time-lapse model. *Fertil Steril Sci* 2021.

embryos). As can be seen in the figure, the shape of the ROC calibration and validation curves were similar, which was also an indicator of the reliability of the models developed.

The algorithm scoring system developed for day 2, day 3, and day 5 ICSI Carmel models are presented in Tables 9, 10, and 11, respectively. The variables chosen as well as the weight and timing range values for these variables are shown. The model formulas presented were obtained as follows:

**Day 2 ICSI Carmel model.** (Table 9) includes variables of cell division timing (t2, t3, and t4) and cell cycle intervals (cc2, s2, vPN2, and vPN3). An additional variable introduced in this model was "cells 44." This variable refers to embryos that reached a 4-cell division stage within ≤44 hours after fertilization, as described by Basile et al. (16). Its purpose was to dispose late embryos. The model algorithm is described using the following formula:

$$\text{Score} = P(\text{cc2}) + P(\text{vPN2}) + P(\text{s2}) + P(\text{t2}) + P(\text{cc2}) + P(\text{t4}) + P(\text{cell44}) + P(\text{t3}) + P(\text{vPN3})$$

**Day 3 ICSI Carmel model.** (Table 10) includes variables of cell division timing (t2, t4, and t5) and cell cycle intervals (vPN2, cc2, s2, cc3, and s3). The model algorithm is described using the following formula:

$$\text{Score} = P(\text{cc2}) + P(\text{vPN2}) + P(\text{s2}) + P(\text{t2}) + P(\text{cc2}) + P(\text{t4}) + P(\text{t5}) + P(\text{s3}) + P(\text{t5-t3})$$

**Day 5 ICSI Carmel model.** (Table 11) includes variables of cell division timing (t2, t5 and t9+, tM, and tSB) and cell cycle intervals (vPN2, cc2, s2, and s3). The model algorithm is described using the following formula:

$$\text{Score} = P(\text{cc2}) + P(\text{vPN2}) + P(\text{s2}) + P(\text{t2}) + P(\text{cc2}) + P(\text{t5}) + P(\text{s3}) + P(\text{t9}) + P(\text{tM}) + P(\text{tSB})$$

Moreover, when we added later timing variables, such as t5 for the day 2 algorithm or t9 for the day 3 algorithm, the AUC value obtained was much higher. However, for the model to allow a true comparison of embryos on day 2 or day 3, the variables we chose for developing the algorithms were t4 for the day 2 and t8 for day 3 models.

### Results of Embedding the Models in the EmbryoScope EmbryoViewer Incubator Software

Applying the new models to the EmbryoViewer software of the EmbryoScope incubator was relatively simple. We have been using the new in-house models for embryo selection since May 2019.

When the model was applied, we noted that the models did not score some of the embryos. We found that when an embryo was selected for transfer or freezing before the models defined hpi or before reaching the maximum number of cells defined by the model, no score was accepted. For example,

TABLE 7

## Algorithm calibration/validation results for day 3 ICSI Carmel model.

Data for developing the algorithm	Data subset used for validation (%)	Calibration subset (no.)	Validation subset (no.)	AUC calibration	AUC validation
ICSI day 3 GS					
All KIDs (KID neg. 912)	25	906	302	0.680	0.632
KID pos. 296)	33	809	399	0.685	0.635
	40	725	483	0.684	0.647
	50	604	604	0.687	0.651
Age $\leq$ 41 y KIDs (KID neg. 652)	25	709	236	0.656	0.704
KID pos. 293)	33	633	312	0.661	0.678
	40	567	378	0.672	0.660
	50	472	473	0.690	0.648
ICSI day 3 CP					
All KIDs (KID neg. 923)	25	885	295	0.683	0.668
KID pos. 257)	33	791	389	0.686	0.666
	40	708	472	0.688	0.669
	50	590	590	0.691	0.669
Age $\leq$ 41 y KIDs (KID neg. 663)	25	688	229	0.663	0.722
KID pos. 254)	33	614	303	0.667	0.696
	40	550	367	0.681	0.674
	50	458	459	0.699	0.663
ICSI day 3 LB					
All KIDs (KID neg. 940)	25	838	279	0.662	0.699
KID pos. 177)	33	748	369	0.673	0.666
	40	670	447	0.684	0.649
	50	558	559	0.714	0.619
Age $\leq$ 41 y KIDs (KID neg. 680)	25	643	214	0.659	0.699
KID pos. 177)	33	574	283	0.676	0.663
	40	514	343	0.691	0.640
	50	428	429	0.709	0.642

Note: ICSI = intracytoplasmic sperm injection; AUC = area under the characteristic curve; KID = known implantation data; GS = gestational sac; CP = clinical pregnancy; LB = live birth; neg = negative; pos = positive.

Tests were performed using the GS, CP, and LB KID embryos databases: "All KID embryos" and "KID embryos from women aged  $\leq$ 41 years." The AUC values are shown for both the scored calibration subset (AUC calibration) and validation subset (AUC validation).

Blais. Laboratory-adapted time-lapse model. *Fertil Steril Sci* 2021.

embryos selected for transfer or freezing before 44 hpi (for day-2 embryos) or those selected before their division into 8 cells (for day-3 embryos) were not scored. To overcome this, we defined another version of the model, in which the following parameters appeared as "informational" only and not as a part of the score. For example, for the day 2 model, we defined the variable "Cells 44" as information only in cycles in which embryo selection occurred in  $<$ 44 hpi so that the scores for 4-celled embryos could be obtained even if they did not reach 44 hpi.

## DISCUSSION

Our primary objective was to assess whether it is possible to develop an in-house laboratory-adapted model for embryo selection in a laboratory where embryos have already been selected for transfer using the general models ("modified Alpha/ESHRE," KIDScore day 3 and day 5 models). Because significant differences in morphokinetics were found between the KID-positive and -negative embryos in our laboratory, it was possible to use our specific KID data to develop a model.

Our secondary objective was to develop and validate the in-house laboratory-adapted models. Using the EmbryoScope Stats software, we were able to develop laboratory-specific models for embryo selection on days 2, 3, and 5 in ICSI em-

bryos. The day 3 and day 5 ICSI Carmel models were complementary to the general models (KIDScore day 3 and day 5), and they refined our ability to select the best embryo from a cohort. There is no general model available for embryo selection on day 2; therefore, our day 2 ICSI Carmel model might assist in selecting day-2 embryos.

Different studies have previously shown that multiple cultural, environmental, and clinical factors influence morphokinetic values (36, 39). Mumusoglu et al. (39) analyzed the effects of patient and ovarian stimulation-related factors on morphokinetics. They showed that no single factor, such as body mass, total FSH dose, duration of infertility, number of previous cycles, antral follicle count, ovarian stimulation protocol, and estradiol trigger, elicits a systematic effect on any morphokinetic parameter. As was also shown in our data, women's age was not found to be a significant parameter (39). The culture conditions and patient characteristics in our KID data were examined and found to be similar between the KID-positive and -negative embryos. Consequently, the differences observed between the KID-positive and -negative embryos can be attributed mainly to the differences in embryo morphokinetics.

In studies that examined specific conditions, fertilization method (ICSI vs. IVF) (26, 40, 41) was found to be an important factor influencing differences in embryo morphokinetics.

TABLE 8

## Algorithm calibration/validation results for day 5 ICSI Carmel model.

Data for developing the algorithm	Data subset used for validation (%)	Calibration subset (no.)	Validation subset (no.)	AUC calibration	AUC validation
ICSI day 5 GS					
All KIDs (KID neg. 912)	25	906	302	0.788	0.830
KID pos. 296)	33	809	399	0.778	0.846
	40	725	483	0.751	0.872
	50	604	604	0.757	0.844
Age ≤41 y KIDs (KID neg. 652)	25	709	236	0.811	0.745
KID pos. 293)	33	633	312	0.818	0.749
	40	567	378	0.805	0.774
	50	472	473	0.817	0.772
ET day5 KIDs (KID neg. 113)	25	169	56	0.834	0.673
KID pos. 112)	33	151	74	0.830	0.735
	40	135	90	0.807	0.794
	50	112	113	0.779	0.832
ICSI day 5 CP					
All KIDs (KID neg. 923)	25	885	295	0.841	0.763
KID pos. 257)	33	791	389	0.838	0.784
	40	708	472	0.833	0.798
	50	590	590	0.839	0.803
Age ≤41 y KIDs (KID neg. 663)	25	688	229	0.802	0.830
KID pos. 254)	33	614	303	0.802	0.826
	40	550	367	0.793	0.833
	50	458	459	0.786	0.840
ET day5 KIDs (KID neg. 118)	25	166	55	0.823	0.787
KID pos. 103)	33	148	73	0.816	0.822
	40	133	88	0.807	0.834
	50	110	111	0.848	0.807
ICSI day 5 LB					
All KIDs (KID neg. 940)	25	838	279	0.832	0.727
KID pos. 177)	33	748	369	0.816	0.788
	40	670	447	0.797	0.815
	50	558	559	0.792	0.805
Age ≤41 y KIDs (KID neg. 680)	25	643	214	0.787	0.826
KID pos. 177)	33	574	283	0.801	0.792
	40	514	343	0.796	0.809
	50	428	429	0.791	0.817
ET day5 KIDs (KID neg. 125)	25	142	47	0.799	0.792
KID pos. 64)	33	127	62	0.783	0.820
	40	113	76	0.761	0.844
	50	94	95	0.753	0.833

Note: ICSI = intracytoplasmic sperm injection; AUC = area under the characteristic curve; KID = known implantation data; GS = gestational sac; CP = clinical pregnancy; LB = live birth; neg = negative; pos = positive.

Tests were performed using the GS, CP, and LB KID embryos databases: "All KID embryos," "KID embryos from women aged ≤41 years," and "Day 5 KID embryos." The AUC values are shown for both the scored calibration subset (AUC calibration) and validation subset (AUC validation).

Blais. Laboratory-adapted time-lapse model. *Fertil Steril Sci* 2021.

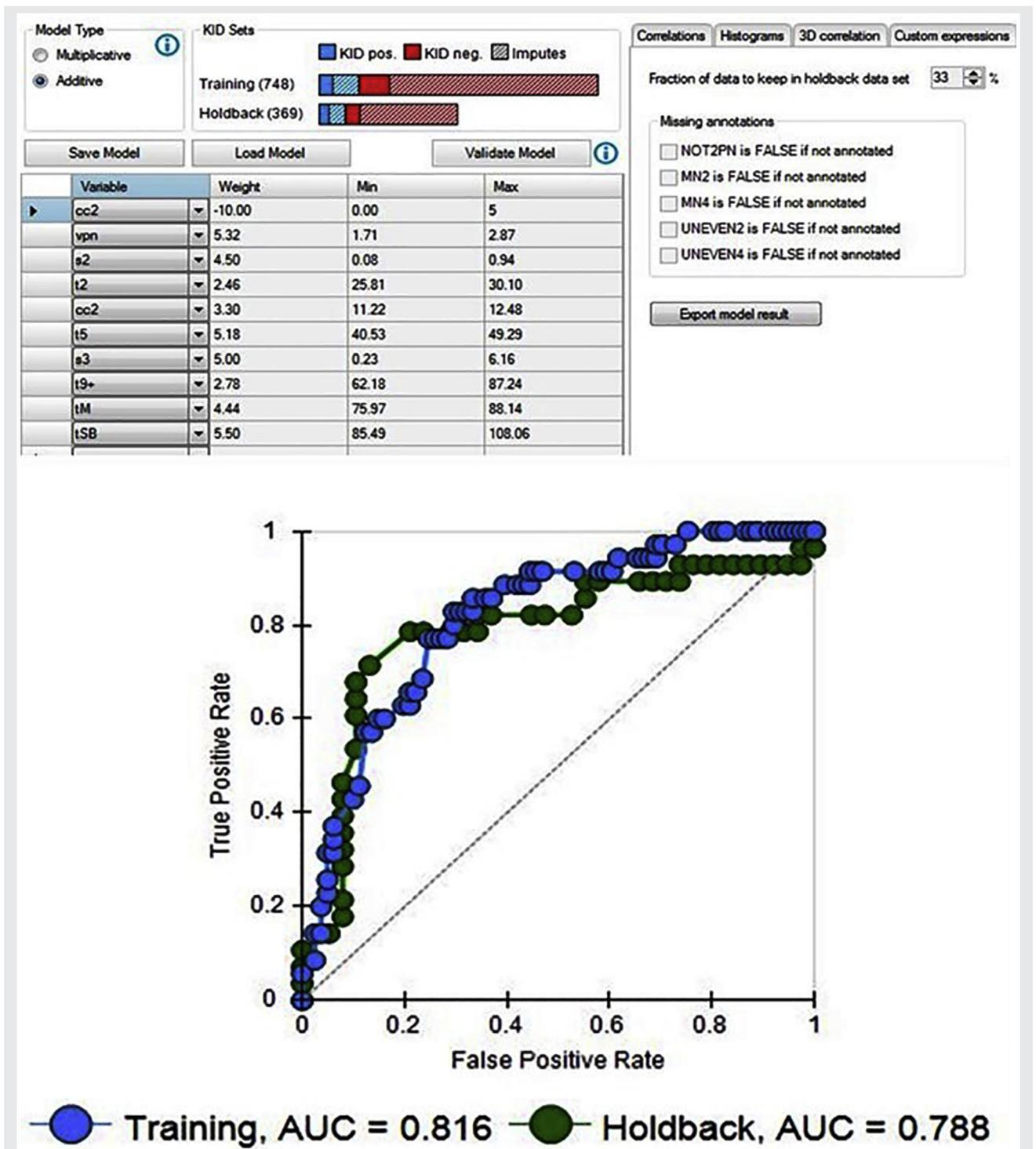
Our study also demonstrated clear differences in morphokinetics between the ICSI and IVF embryos. Major biases occurred at the time of fertilization in the IVF embryos because the time at which we added the sperm to the oocytes ( $t_0$ ) in culture was not necessarily the time of fertilization. This is probably the reason that no clear differences were found in the cell cleavage timing between the KID-positive and -negative IVF embryos. Consequently, information on timing variables was not accurate enough for designing an algorithm based on IVF timing data. Minor biases can also be seen in ICSI cases that require the injection of many oocytes, where a relatively long time period passes from the start of the injection of the first oocyte to the injection of the last one. However, in most cases, the injection time is short and does not exceed more than a few minutes (10–15 minutes).

Significant differences in morphokinetics between the KID-positive and -negative embryos were found in most cell

division timing variables and some of the cell cycle intervals, which allowed us to develop the model. Various studies have shown the importance of these parameters for embryo development and implantation. Long or short cell cycle intervals can serve as an indicator of the probability of an embryo for implantation (3–9, 20–24, 26, 37, 41–47). The most striking example is a particularly short interval in the first cell cycle (<5 hours) that can indicate direct uneven cleavage (DUC). Direct uneven cleavage significantly reduces the chances of an embryo for implantation (4, 8, 16, 20, 22, 36, 45).

Although this study found clear differences between the KID-positive and -negative embryos for most of the cell division timing variables, differences in cell cycle intervals were found only in few of the cases. These results can probably be explained by the fact that the embryos were already selected for transfer using the general models. The algorithm

**FIGURE 5**



Example of algorithm development result for day 5 ICSI Carmel model in embryos with known live birth outcomes (LB KID embryos) using the EmbryoScope Stats software. Above the image is the user-defined algorithm, tested on 67% of training data and validated on 33% holdout data. The resulting graphical form and the AUC (ROC test) results of the training (AUC = 0.816) and holdout data (AUC = 0.788) can be seen below the illustration. ICSI = intracytoplasmic sperm injection; LB = live birth; KID = known implantation data; AUC = area under the characteristic curve; ROC = receiver operating curve.

Blais. Laboratory-adapted time-lapse model. *Fertil Steril Sci* 2021.

TABLE 9

The algorithm scoring system developed for day 2 ICSI Carmel model.

Model type: day 2, additive

Variable	Weight	Min	Max	Description	P (Variable)
cc2 = t3-t2	-10.00	0.00	5	Avoid	-10, if $0.0 \leq cc2 \leq 5.0$ 0, if $0.0 > cc2$ or $cc2 > 5.0$
vPN = t2-tpnf	3.98	1.37	3.85	Prefer	3.98, if $1.4 \leq vPN2 \leq 3.9$ 0, if $1.4 > vPN2$ or $vPN2 > 3.9$
s2 = t4-t3	2.92	0.27	1.00	Prefer	2.92, if $0.3 \leq s2 \leq 1.0$ 0, if $0.3 > s2$ or $s2 > 1.0$
t2	2.60	25.81	27.05	Prefer	2.6, if $25.8 \leq t2 \leq 27.1$ 0, if $25.8 > t2$ or $t2 > 27.1$
cc2 = t3-t2	1.70	11.19	12.17	Prefer	1.7, if $11.2 \leq cc2 \leq 12.2$ 0, if $11.2 > cc2$ or $cc2 > 12.2$
t4	1.64	33.18	37.77	Prefer	1.64, if $33.2 \leq t4 \leq 37.8$ 0, if $33.2 > t4$ or $t4 > 37.8$
cells44 = cells (44)	3.48	4.0	4.0	Prefer	3.48, if $4.0 \leq cells44 \leq 4.0$ 0, if $4.0 > cells44$ or $cells44 > 4.0$
t3	2.08	33.45	41.05	Prefer	2.08, if $33.5 \leq t3 \leq 41.0$ 0, if $33.5 > t3$ or $t3 > 41.0$
vPN3 = t3-tpnf	1.46	12.05	14.83	Prefer	1.46, if $12.1 \leq vPN3 \leq 14.8$ 0, if $12.1 > vPN3$ or $vPN3 > 14.8$

Note: ICSI = intracytoplasmic sperm injection; min = minimum; max = maximum.

Blais. Laboratory-adapted time-lapse model. Fertil Steril Sci 2021.

TABLE 10

The algorithm scoring system developed for day 3 ICSI Carmel model.

Model type: day 3, additive

Variable	Weight	Min	Max	Description	P (variable)
cc2 = t3-t2	-10.00	0.00	5	Avoid	-10, if $0.0 \leq cc2 \leq 5.0$ 0, if $0.0 > cc2$ or $cc2 > 5.0$
vPN = t2-tpnf	5.64	1.73	5.42	Prefer	5.64, if $1.7 \leq vPN2 \leq 5.4$ 0, if $1.7 > vPN2$ or $vPN2 > 5.4$
s2 = t4-t3	1.70	0.04	0.78	Prefer	1.7, if $0.0 \leq s2 \leq 0.8$ 0, if $0.0 > s2$ or $s2 > 0.8$
t2	2.46	24.98	27.44	Prefer	2.46, if $25.0 \leq t2 \leq 27.4$ 0, if $25.0 > t2$ or $t2 > 27.4$
cc2 = t3-t2	2.08	11.25	12.51	Prefer	2.08, if $11.2 \leq cc2 \leq 12.5$ 0, if $11.2 > cc2$ or $cc2 > 12.5$
t4	3.22	33.46	37.23	Prefer	3.22, if $33.5 \leq t4 \leq 37.2$ 0, if $33.5 > t4$ or $t4 > 37.2$
t5	2.02	44.34	52.36	Prefer	2.02, if $44.3 \leq t5 \leq 52.4$ 0, if $44.3 > t5$ or $t5 > 52.4$
s3 = t8-t5	3.74	0	6.12	Prefer	3.74, if $0 \leq s3 \leq 6.1$ 0, if $0 > s3$ or $s3 > 6.1$
t5t3 = t5-t3	1.70	11.32	14.83	Prefer	1.7, if $11.3 \leq t5t3 \leq 14.8$ 0, if $11.3 > t5t3$ or $t5t3 > 14.8$

Note: ICSI = intracytoplasmic sperm injection; min = minimum; max = maximum.

Blais. Laboratory-adapted time-lapse model. Fertil Steril Sci 2021.

of the general models used for embryo selection relies primarily on time intervals (26) (Supplemental Appendix 1). However, we chose to include some of the cell cycle intervals in the algorithms we developed. Interval cc2 appeared twice in the

model algorithms because it had 2 important aims. The first was to exclude DUC embryos. To do this, an embryo undergoing cleavage between 2 and 3 cells within 0–5 hours received a negative weight of -10 points [a similar criteria was used in

TABLE 11

The algorithm scoring system as developed for day 5 ICSI Carmel model.

Model type: day 5, additive

Variable	Weight	Min	Max	Description	P (Variable)
cc2 = t3-t2	-10.00	0.00	5	Avoid	-10, if $0.0 \leq cc2 \leq 5.0$ 0, if $0.0 > cc2$ or $cc2 > 5.0$
vPN = t2-tpnf	5.32	1.71	2.87	Prefer	5.32, if $1.7 \leq vPN2 \leq 2.9$ 0, if $1.7 > vPN2$ or $vPN2 > 2.9$
s2 = t4-t3	4.50	0.08	0.94	Prefer	4.5, if $0.1 \leq s2 \leq 0.9$ 0, if $0.1 > s2$ or $s2 > 0.9$
t2	2.46	25.81	30.10	Prefer	2.46, if $25.8 \leq t2 \leq 30.1$ 0, if $25.8 > t2$ or $t2 > 30.1$
cc2 = t3-t2	3.30	11.22	12.48	Prefer	3.3, if $11.2 \leq cc2 \leq 12.5$ 0, if $11.2 > cc2$ or $cc2 > 12.5$
t5	5.18	40.53	49.29	Prefer	5.18, if $40.5 \leq t5 \leq 49.3$ 0, if $40.5 > t5$ or $t5 > 49.3$
s3 = t8-t5	5.00	0.23	6.16	Prefer	5, if $0.2 \leq s3 \leq 6.2$ 0, if $0.2 > s3$ or $s3 > 6.2$
t9+	2.78	62.18	87.24	Prefer	2.78, if $62.2 \leq t9 \leq 87.2$ 0, if $62.2 > t9$ or $t9 > 87.2$
tM	4.44	75.97	88.14	Prefer	4.44, if $76.0 \leq tM \leq 88.1$ 0, if $76.0 > tM$ or $tM > 88.1$
tSB	5.50	85.49	108.06	Prefer	5.5, if $85.5 \leq tSB \leq 108.1$ 0, if $85.5 > tSB$ or $tSB > 108.1$

Note: ICSI = intracytoplasmic sperm injection; min = minimum; max = maximum.

Blais. Laboratory-adapted time-lapse model. Fertil Steril Sci 2021.

the model by Basile et al. (8)]. Because there is currently no general model for embryo selection on day 2, we found this to be an important criterion for embryo deselection. The second reason to include cc2 timing interval was for its added value in embryo selection. We found that when the intervals cc2, s2 vPN2, vPN3, and t5-t3 were used as inclusion criteria, higher AUC values were accepted. By defining the range of laboratory-specific timing for these intervals, they contribute to the selection of embryos that have the highest chance of implantation, thus improving our laboratory-specific models.

Interesting results were obtained for t5 cell cleavage timing. A similar average timing was found for the KID-positive and -negative embryos. We assumed, as reported by Herrero et al. (42), that a short t5 cleavage timing was recorded for some of the embryos because these embryos had partial DUC at the 2-cell stage (cleavage from 2 to 5 cells), and had not cleaved normally from 4- to 5-cells. In this case, a short t5 time interval was recorded, and a similar average was obtained for the KID-positive and -negative embryos. This masked the ability to distinguish between normal and partially abnormal embryos. For this reason, although no differences were found between the KID-positive and -negative embryos for the average t5 cleavage timing, t5 was found to have an optimal range in the quartile test and was found to be a dominant parameter in most cell cycle intervals. The same phenomenon was not observed for t3 in our database probably because we excluded embryos showing full DUC from 1 to 3 cells.

Other models also highlight the importance of t5. The timing of cell division t5 was defined by Meseguer et al. (4)

as a key event in early embryo development and was chosen to be the first timing for the classification of embryos using a hierarchical model. They found that the implantation rate of ICSI embryos with t5 cleavage within the best 2 quartile ranges was 2.6 times the implantation rate for embryos outside this range. Thus, it was the investigators' opinion that t5 provides the best early cleavage single criteria to select embryos with an improved implantation potential. Chamayou et al. (43) found that the morphokinetic parameter most significantly associated with implantation was cc3 (t5-t3), whereas no significant differences were found for t5 in ICSI embryos. Embryos with a poor prognosis for in vitro development had a wider range of cc3 values in comparison to a group of implanted embryos. Wu et al. (46) showed that IVF embryos that reached the best range of time interval for t5-t4 had significantly higher implantation rates than those outside of this range for t5-t4, again highlighting the importance of this parameter in implantation prediction. Basile et al. (44), who studied differences in cleavage timing between chromosomally normal and abnormal embryos, found that embryos falling within optimal ranges for t5, cc3 (t5-t3), and t5-t2 exhibited a significantly greater proportion of normal embryos than those falling outside of these ranges (44).

An embryo DUC of 2-5 cells is an interesting issue and is also presented in the general model KIDScore algorithm using the equation  $B = (t5 - t3) / (t5 - t2)$ . According to the general model, equation B is used for embryos that show an irregular cleavage pattern (26). In our study, no differences in equation B were found between the KID-positive and -negative

embryos for the ICSI KID embryos, but differences were found for the IVF KID embryos. In our laboratory, we do not deselect embryos with scores 3–4 in the KIDScore D3 model (score range 1–5) unless there are other factors that influence the decision to deselect the embryos. Even without the deselection of embryos with scores of 3–4 in the KIDScore model, equation B was not found to be statistically significant in our study and, therefore, was not included in our laboratory-specific model. The equation of interval cc3 (t5-t3) was introduced into our day 3 ICSI in-house model because it had significantly contributed to the scoring of the model. Further studies with TLS are needed to understand the influence of 2–5-cell embryo DUC on implantation and LB.

An improved predictability of the laboratory-specific models was observed from day 2–3 to day 5. The AUC values were 0.657 and 0.673 for the day 2 and day 3 models, respectively, and 0.803 for the day 5 model.

The AUC values achieved in our laboratory-specific model were similar and even higher than the general applicable KIDScore model's AUC values, as shown in our study and several other studies (13, 26, 48). Petersen et al. (26) and Liu et al. (48) showed that in the KIDScore D3 model, the implantation potential of embryos was predicted with AUC values of 0.650 and 0.614, respectively. Reignier et al. (13) demonstrated AUC values of 0.59 and 0.60, respectively, for ROC curves for both KIDScore day 5 versions 1 and 2.

Meseguer et al. (4) and Basile et al. (8) developed algorithms in which all kinetic variables included were up to day 2 of development. They agreed that the use of a day-2 algorithm did not necessarily replace blastocyst culture with better embryo selection. Aparicio-Ruiz et al. (17) also used laboratory-specific day-2 morphokinetic parameters (t2-t4) for day-3 and day-5 embryo selection and obtained an AUC value of 0.65. Based on these studies, it can be concluded that a combination of blastocyst culture and selection based on morphology and early kinetic markers seems to be the best approach for improving clinical outcomes (4, 8, 17).

Although previous studies have emphasized the importance of early morphokinetic markers, more recent studies have demonstrated the importance of later morphokinetic markers. Late morphokinetic timing parameters can suggest the ploidy of embryos. Mumusoglu et al. (39) demonstrated a significant delay in 10 morphokinetic timings (PNa, t2, t7-tEB, and t9-t2) when aneuploid blastocysts were compared with euploid blastocysts. However, when patient- and ovarian stimulation-related factors were taken into account, only 5 late parameters (t9, tM, tSB, tB, and tEB) remained significant for ploidy status. They noted that aneuploid embryos appeared to have a significantly delayed time for blastocyst development at the postcleavage stage. Campbell et al. (25) and Desai et al. (20) showed in their studies that late kinetic parameters appeared to be associated with the likelihood of euploidy. Campbell et al. (25) developed an aneuploidy risk model algorithm based on late morphokinetic parameters (tSB and tB), with AUC values of 0.75 for CP and 0.74 for LB. The investigators concluded that time-lapse imaging using defined morphokinetic data could be used to classify embryos according to their risk of aneuploidy, without performing biopsy and preimplantation genetic screening

(PGS), and that this correlated well with clinical outcomes. Desai et al. (20) concluded that early kinetic markers did not predict normal chromosomal content but only the ability to form blastocysts. They concluded that although DUC and other dysmorphisms, like irregular chaotic division, had a high negative predictive value regardless of other blastocyst kinetic parameters included in the selection model, late kinetic parameters, like tSB, tEB, and tEB-tSB, are predictive of euploidy. Similar results were shown by Rienzi et al. (29), who demonstrated the importance of the time of morulation (tM) in predicting a euploid blastocyst's reproductive competence based on preimplantation genetic testing for aneuploidy. Basil et al. (18) suggested that morphokinetics should not be considered a competitive technology or replacement to preimplantation genetic testing for aneuploidies or other technologies but rather a complimentary technology. The investigators suggested that time-lapse be used as an opportunity to improve selection even within a pool of euploid embryos (18). We adopted this approach and argued that a laboratory-optimized in-house model serves as an important tool, in addition to other embryo-selection tools, in accurately predicting our ability to select embryos.

As previously noted, our model also showed an increase in the predictability of the models from days 2–3 to day 5 possibly because it was reasonable to assume that it was associated with ploidy. Moreover, a fully automated deep learning model for day-5 embryos has recently been described and appears to be able to predict a pregnancy based on a positive fetal heart beat, with an extremely high AUC of 0.93 (49). Although artificial intelligence (AI) algorithms for time-lapse embryo selection are being developed, these methods are not yet readily available in most IVF laboratories. A laboratory-specific model seems to be, at present, more feasible for widespread use. It will be interesting to follow the AI model development and its implementation in IVF laboratories.

## CONCLUSIONS

From a clinical perspective, a model based on in-house laboratory-specific morphokinetics was found to be complementary to the general models for embryo selection and serves as an important tool for improving single embryo selection. The present model developed is not only based on implantation results but has also been tested and validated using live birth data. The in-house model refines the ability to objectively choose an embryo that has the highest chance for implantation and live birth under the specific environmental and clinical condition in laboratories. The ability to identify an embryo with the highest implantation potential would likely increase the widespread practice of single ET.

From a laboratory practical aspect, developing an in-house laboratory-specific model requires many stages of sorting and characterization of the data. It demands accurate annotation and statistical analysis for every parameter alone and in combination to evaluate which parameter should be included in the model. The model developing process, as described, could facilitate and improve the process of in-house model development in other laboratories.

## REFERENCES

- 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;26:1270–83.
- Lundin K, Ahlström A. Quality control and standardization of embryo morphology scoring and viability markers. *Reprod Biomed Online* 2015; 31:459–71.
- 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.
- Meseguer M, Herrero J, Tejera A, Hilligsøe KM, Ramsing NB, Remohí J. The use of morphokinetics as a predictor of embryo implantation. *Hum Reprod* 2011;26:2658–71.
- Meseguer M, Rubio I, Cruz M, Basile N, Marcos J, Requena A. Embryo incubation and selection in a time-lapse monitoring system improves pregnancy outcome compared with a standard incubator: a retrospective cohort study. *Fertil Steril* 2012;98:1481–9.
- Chen AA, Tan L, Suraj V, Pera RR, Shen S. Biomarkers identified with time-lapse imaging: discovery, validation, and practical application. *Fertil Steril* 2013;99:1035–43.
- Conaghan J, Chen AA, Willman SP, Ivani K, Chenette PE, Boostanfar R, et al. Improving embryo selection using a computer-automated time-lapse image analysis test plus day 3 morphology: results from a prospective multicenter trial. *Fertil Steril* 2013;100:412–9.
- Basile N, Vime P, Florensa M, Aparicio Ruiz B, García Velasco JA, Remohí J, et al. The use of morphokinetics as a predictor of implantation: a multicentric study to define and validate an algorithm for embryo selection. *Hum Reprod* 2015;30:276–83.
- Rubio I, Galán A, Larreategui Z, Ayerdi F, Bellver J, Herrero J, et al. Clinical validation of embryo culture and selection by morphokinetic analysis: a randomized, controlled trial of the EmbryoScope. *Fertil Steril* 2014;102:1287–94.
- Matyas SZ, Kovacs P, Forgacs V, Sajgó A, Pribenszky CS. Selection of single blastocyst for transfer using time-lapse monitoring during in vitro fertilization in good prognosis patients: a randomized controlled trial. *Hum Reprod* 2015;30:119.
- Pribenszky C, Nilselid AM, Montag M. Time-lapse culture with morphokinetic embryo selection improves pregnancy and live birth chances and reduces early pregnancy loss: a meta-analysis. *Reprod Biomed Online* 2017; 35:511–20.
- Pribenszky C, Nilselid AM, Montag M. Response: time-lapse systems for ART. *Reprod Biomed Online* 2018;36:290–2.
- Reignier A, Girard JM, Lammers J, Chtourou S, Lefebvre T, Barriere P, et al. Performance of Day 5 KIDScore™ morphokinetic prediction models of implantation and live birth after single blastocyst transfer. *J Assist Reprod Genet* 2019;36:2279–85.
- Fishel S, Campbell A, Montgomery S, Smith R, Nice L, Duffy S, et al. Live births after embryo selection using morphokinetics versus conventional morphology: a retrospective analysis. *Reprod Biomed Online* 2017;35: 407–16.
- Adolfsson E, Porath S, Andershed AN. External validation of a time-lapse model; a retrospective study comparing embryo evaluation using a morphokinetic model to standard morphology with live birth as endpoint. *JBRA Assist Reprod* 2018;22:205–14.
- Basile N, Caiazzo M, Meseguer M. What does morphokinetics add to embryo selection and in-vitro fertilization outcomes? *Curr Opin Obstet Gynecol* 2015;27:193–200.
- Aparicio-Ruiz B, Basile N, Albalá SP, Bronet F, Remohí J, Meseguer M. Automatic time-lapse instrument is superior to single-point morphology observation for selecting viable embryos: retrospective study in oocyte donation. *Fertil Steril* 2016;106:1379–85.
- Basile N, Elkhatib I, Meseguer M. A Strength, weaknesses, opportunities and threats analysis on time lapse. *Curr Opin Obstet Gynecol* 2019;31:148–55.
- Kalleas D, McEvoy K, Horne G, Roberts SA, Brison DR. Live birth rate following undisturbed embryo culture at low oxygen in a time-lapse incubator compared to a high-quality benchtop incubator. *Hum Fertil (Camb)* 2020:1–7.
- Desai N, Goldberg JM, Austin C, Falcone T. Are cleavage anomalies, multinucleation, or specific cell cycle kinetics observed with time-lapse imaging predictive of embryo developmental capacity or ploidy? *Fertil Steril* 2018; 109:665–74.
- Liu Y, Chapple V, Feenan K, Roberts P, Matson P. Time-lapse deselection model for human day 3 in vitro fertilization embryos: the combination of qualitative and quantitative measures of embryo growth. *Fertil Steril* 2016; 105:656–62.
- Yang SH, Wu CH, Chen YC, Yang CK, Wu TH, Chen PC, et al. Effect of morphokinetics and morphological dynamics of cleavage stage on embryo developmental potential: a time-lapse study. *Taiwan J Obstet Gynecol* 2018;57:76–82.
- Milewski R, Kuć P, Kuczyńska A, Stankiewicz B, Łukaszk K, Kuczyński W. A predictive model for blastocyst formation based on morphokinetic parameters in time-lapse monitoring of embryo development. *J Assist Reprod Genet* 2015;32:571–9.
- Milewski R, Milewska AJ, Kuczyńska A, Stankiewicz B, Kuczyński W. Do morphokinetic data sets inform pregnancy potential? *J Assist Reprod Genet* 2016;33:357–65.
- Campbell A, Fishel S, Bowman N, Duffy S, Sedler M, Thornton S. Retrospective analysis of outcomes after IVF using an aneuploidy risk model derived from time-lapse imaging without PGS. *Reprod Biomed Online* 2013;27: 140–6.
- Petersen BM, Boel M, Montag M, Gardner DK. Development of a generally applicable morphokinetic algorithm capable of predicting the implantation potential of embryos transferred on Day 3. *Hum Reprod* 2016;31:2231–44.
- Fishel S, Campbell A, Montgomery S, Smith R, Nice L, Duffy S, et al. Time-lapse imaging algorithms rank human preimplantation embryos according to the probability of live birth. *Reprod Biomed Online* 2018;37:304–13.
- Bodri D, Milewski R, Serna JY, Sugimoto T, Kato R, Matsumoto T, et al. Predicting live birth by combining cleavage and blastocyst-stage time-lapse variables using a hierarchical and a data mining-based statistical model. *Reprod Biol* 2018;18:355–60.
- Renzi L, Cimadomo D, Delgado A, Minasi MG, Fabozzi G, Gallego RD, et al. Time of morulation and trophoctoderm quality are predictors of a live birth after euploid blastocyst transfer: a multicenter study. *Fertil Steril* 2019;112: 1080–93.
- Best L, Campbell A, Duffy S, Montgomery S, Fishel S. Does one model fit all? Testing a published embryo selection algorithm on independent time-lapse data. *Hum Reprod* 2013;28(Suppl 1):i87–90.
- Yalçınkaya E, Ergin EG, Çalışkan E, Öztel Z, Ozay A, Özörnek H. Reproducibility of a time-lapse embryo selection model based on morphokinetic data in a sequential culture media setting. *J Turk Ger Gynecol Assoc* 2014;15:156–60.
- Fréour T, Le Fleuter N, Lammers J, Splingart C, Reignier A, Barrière P. External validation of a time-lapse prediction model. *Fertil Steril* 2015;103: 917–22.
- Barrie A, Homburg R, McDowell G, Brown J, Kingsland C, Troup S. Examining the efficacy of six published time-lapse imaging embryo selection algorithms to predict implantation to demonstrate the need for the development of specific, in-house morphokinetic selection algorithms. *Fertil Steril* 2017; 107:613–21.
- Technote KIDScore D3 Basic decision support tool v2\_INT.pdf. Vitrolife.
- Vitrolife. Technote – KIDScore D5 v.3 decision support tool. Available at: [Technote\\_kidscore-d5.3\\_v3\\_v2.pdf](#).
- Ciray HN, Campbell A, Agerholm IE, Aguilar J, Chamayou S, Esbert M, et al. Proposed guidelines on the nomenclature and annotation of dynamic human embryo monitoring by a time-lapse user group. *Hum Reprod* 2014; 29:2650–60.
- Desai N, Ploskonka S, Goodman LR, Austin C, Goldberg J, Falcone T. Analysis of embryo morphokinetics, multinucleation and cleavage anomalies us-

- ing continuous time-lapse monitoring in blastocyst transfer cycles. *Reprod Biol Endocrinol* 2014;12:54.
38. Kaser DJ, Racowsky C. Clinical outcomes following selection of human pre-implantation embryos with time-lapse monitoring: a systematic review. *Hum Reprod Update* 2014;20:617–31.
  39. Mumusoglu S, Yarali I, Bozdag G, Ozdemir P, Polat M, Sokmensuer LK, et al. Time-lapse morphokinetic assessment has low to moderate ability to predict euploidy when patient- and ovarian stimulation-related factors are taken into account with the use of clustered data analysis. *Fertil Steril* 2017;107:413–21.
  40. Bodri D, Sugimoto T, Serna JY, Kondo M, Kato R, Kawachiya S, et al. Influence of different oocyte insemination techniques on early and late morphokinetic parameters: retrospective analysis of 500 time-lapse monitored blastocysts. *Fertil Steril* 2015;104:1175–81.
  41. Dal Canto M, Cotichio G, Renzini MM, De Ponti E, Novara PV, Brambillasca F, et al. Cleavage kinetics analysis of human embryos predicts development to blastocyst and implantation. *Reprod Biomed Online* 2012;25:474–80.
  42. Herrero J, Tejera A, Albert C, Vidal C, de los Santos MJ, Meseguer M. A time to look back: analysis of morphokinetic characteristics of human embryo development. *Fertil Steril* 2013;100:1602–9.
  43. Chamayou S, Patrizio P, Storaci G, Tomaselli V, Alecci C, Ragolia C, et al. The use of morphokinetic parameters to select all embryos with full capacity to implant. *J Assist Reprod Genet* 2013;30:703–10.
  44. Basile N, Nogales M del C, Bronet F, Florensa M, Riqueiros M, Rodrigo L, et al. Increasing the probability of selecting chromosomally normal embryos by time-lapse morphokinetics analysis. *Fertil Steril* 2014;101:699–704.
  45. Rubio I, Kuhlmann R, Agerholm I, Kirk J, Herrero J, Escribá MJ, et al. Limited implantation success of direct-cleaved human zygotes: a time-lapse study. *Fertil Steril* 2012;98:1458–63.
  46. Wu L, Han W, Zhang X, Wang J, Liu W, Xiong S, et al. A retrospective analysis of morphokinetic parameters according to the implantation outcome of IVF treatment. *Eur J Obstet Gynecol Reprod Biol* 2016;197:186–90.
  47. Wong CC, Loewke KE, Bossert NL, Behr B, De Jonge CJ, Baer TM, et al. Non-invasive imaging of human embryos before embryonic genome activation predicts development to the blastocyst stage. *Nat Biotechnol* 2010;28:1115–21.
  48. 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;22:182–90.
  49. 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* 2018;34:1011–8.
  50. Hlinka D, Kařatová B, Uhrinová I, Dolinská S, Rutarová J, Rezáčová J, et al. Time-lapse cleavage rating predicts human embryo viability. *Physiol Res* 2012;61:513–25.