Development of a dynamic machine learning algorithm to predict clinical pregnancy and live birth rate with embryo morphokinetics

Liubin Yang, M.D., Ph.D.,[a] Mary Peavey, M.D., M.S.C.I.,[a] Khalied Kaskar, Ph.D.,[a] Neil Chappell, M.D., M.S.C.I.,[a] Lynn Zhu, B.S.,[b] Darius Devlin, Ph.D.,[a,c] Cecilia Valdes, M.D.,[a] Amy Schutt, M.D., M.S.C.I.,[a] Terri Woodard, M.D.,[a] Paul Zarutskie, M.D.,[a] Richard Cochran, Ph.D.,[a] and William E. Gibbons, M.D.[a]

[a] Division of Reproductive Endocrinology and Infertility, Department of Obstetrics and Gynecology, Baylor College of Medicine, Huston, Texas; [b] Department of BioSciences, Rice University, Houston, Texas; and [c] Interdepartmental Program in Translational Biology and Molecular Medicine, Baylor College of Medicine, Houston, Texas

Objective: To evaluate the feasibility of generating a center-specific embryo morphokinetic algorithm by time-lapse microscopy to predict clinical pregnancy rates.
Design: A retrospective cohort analysis.
Setting: Academic fertility clinic in a tertiary hospital setting.
Patient(s): Patients who underwent in vitro fertilization with embryos that underwent EmbryoScope time-lapse microscopy and subsequent transfer between 2014 and 2018.
Intervention(s): None.
Main Outcome Measure(s): Clinical pregnancy.
Result(s): A supervised, random forest learning algorithm from 367 embryos successfully predicted clinical pregnancy from a training set with overall 65% sensitivity and 74% positive predictive value, with an area under the curve of 0.7 for the test set. Similar results were achieved for live birth outcomes. For the secondary analysis, embryo growth morphokinetics were grouped into five clusters using unsupervised clustering. The clusters that had the fastest morphokinetics (time to blastocyst = 97 hours) had pregnancy rates of 54%, whereas a cluster that had the slowest morphokinetics (time to blastocyst = 122 hours) had a pregnancy rate of 71%, although the differences were not statistically significant (P=.356). Other clusters had pregnancy rates of 51%–60%.
Conclusion(s): This study shows the feasibility of a clinic-specific, noninvasive embryo morphokinetic simple machine learning model to predict clinical pregnancy rates. (Fertil Steril Rep® 2022;3:116–23. ©2022 by American Society for Reproductive Medicine.)
Key Words: Morphokinetics, artificial intelligence, embryo selection, machine learning, time-lapse microscopy

Embryo quality assessment has historically relied on a variety of snapshot morphological parameters to determine the embryos that have the highest probability of resulting in pregnancy (1), including morphology of the pronuclear stage (2), early embryo cleavage (3, 4), the presence of multinucleation (5, 6), and zona pellucida characteristics (7). However, standard grading only provides a brief view of embryonic morphology at a specific time point, while developmental changes over multiple time points may provide a more robust impression of the implantation potential (8). To obtain objective, real-time embryo morphologic data with potential clinical relevance, we focused on time-lapse microscopy as a noninvasive method to characterize embryo quality. Many studies have identified parameters associated

Received October 28, 2021; revised April 11, 2022; accepted April 12, 2022.
L.Y. has nothing to disclose. M.P. has nothing to disclose. K.K. has nothing to disclose. N.C. has nothing to disclose. L.Z. has nothing to disclose. D.D. has nothing to disclose. C.V. has nothing to disclose. A.S. has nothing to disclose. T.W. has nothing to disclose. P.Z. has nothing to disclose. R.C. has nothing to disclose. W.E.G. has nothing to disclose.
Supported in part by the Robert and Janice McNair Baylor College of Medicine Medical Scientist Training Program Student Scholarship. Supported by National Institute of General Medical Sciences, United States grants T32GM088129 (to the Interdepartmental Program in Translational Biology and Molecular Medicine) and T32GM120011 (to the Training Interdisciplinary Pharmacology Scientists Program) (to D.D.).
Reprint requests: Liubin Yang, M.D., Ph.D., Baylor College of Medicine, 1 Baylor Plaza, Mailstop BCM610, Houston, Texas 77030 (E-mail: liubin.yang@bcm.edu).

Fertil Steril Rep® Vol. 3, No. 2, June 2022 2666-3341

https://doi.org/10.1016/j.xfre.2022.04.004

with improved clinical outcomes, including time of appearance and fading of two pronuclei (9), the timing of cell cycle events (10, 11), the time between 2- to the 8-cell cleavage stage (12–15), cleavage events (16), cytokinesis duration (17), time to initiation of blastulation (18–20), and combined morphokinetic and morphologic or molecular parameters (21–27). Other studies showed that no differences were found in the time-lapse parameters of euploid embryos that led to miscarriages (28). Given the divergent findings, site-specific algorithms have been proposed (29).

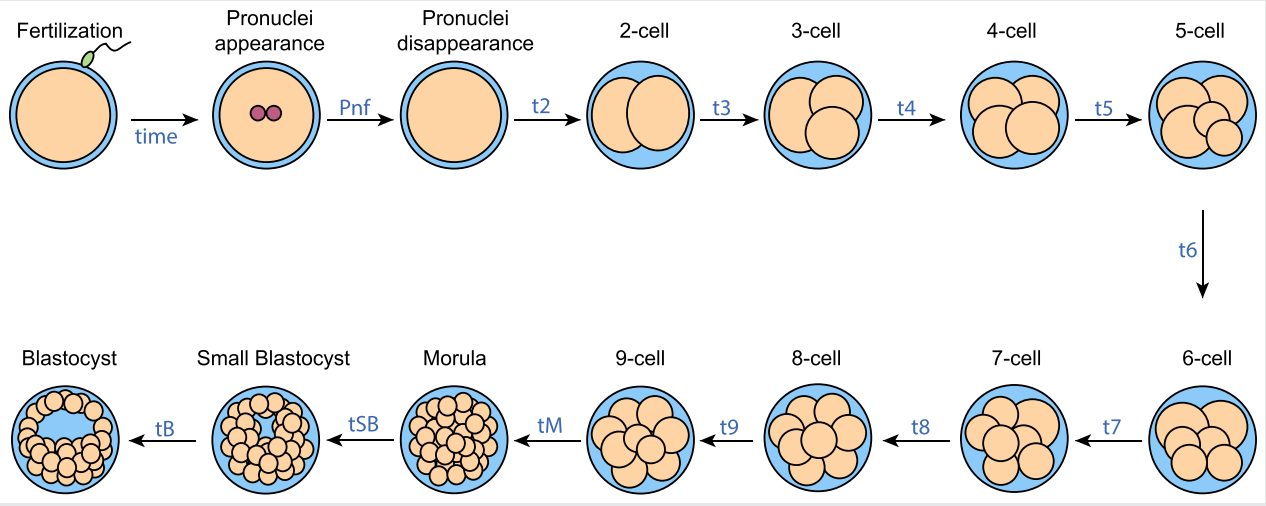
Recently, machine learning has also been employed to predict clinical pregnancy using demographic and clinical data alongside embryo scoring qualities as training data (30–32). Other machine learning techniques aim to predict embryo morphological grading using inputs of still images or videos with high accuracy (33, 34) and also to predict pregnancy rates with high accuracy (35–37). Numerous variables, including demographics, clinical data, patterns in imaging, or video data, have been used to train artificial intelligence (AI) algorithms. Many AI approaches begin with a multistep process of first extracting representations of these complex variables using hand-crafted methods such as feature-engineering or using unsupervised learning (38) before training an algorithm. Factors that may confound or affect the outcome include laboratory parameters (intracytoplasmic sperm insemination), embryo quality, and patient diagnoses (age, prior treatment, ovarian response, endometrial quality) (39, 40), and therefore need to be addressed in AI approaches. For instance, an increase in maternal age was associated with faster early morphokinetics and slower morula and blastocyst formation (40). When computational resources are devoted to high quantity, complex data, the resultant complexity of many parameters can possibly negatively affect the accuracy of algorithms (41, 42). Additionally, this requires expensive, time-consuming multiprocessing cores, a resource to which few laboratories have access. To ameliorate the computational burden and simplify the selection of clinically relevant parameters tied to embryo quality, we hypothesized that a simple machine learning algorithm could be trained on characteristic embryo morphokinetic time points to predict pregnancy outcomes with high accuracy. Our secondary outcome was pregnancy rates in clusters generated by unsupervised machine learning as a proof of concept.

MATERIAL AND METHODS

Study Criteria

A retrospective cohort study was approved by the Institutional Review Board (protocol no. H39094) at Baylor College of Medicine, Houston, Texas. Inclusion criteria were all autologous and donor cycle embryos that were incubated in the EmbryoScope time-lapse microscope (Vitrolife, Sweden) and were subsequently transferred, with in vitro culture of embryos to the blastocyst stage (day 5 to day 7), resulting in a clinical pregnancy (defined by fetal heartbeat after 6 weeks) or no clinical pregnancy (not excluding biochemical pregnancies) from 2014 to 2018 at the Texas Children's Hospital Family Fertility Center. Exclusion criteria included embryo culture in other incubators, embryos without morphokinetic data or incomplete annotations, and dual embryo transfers with discordant outcomes. The sample size was chosen based on other studies that have found significant patterns in <200 embryos and validation cohorts <100 embryos (29). There were no ectopic gestations included in the data sets.

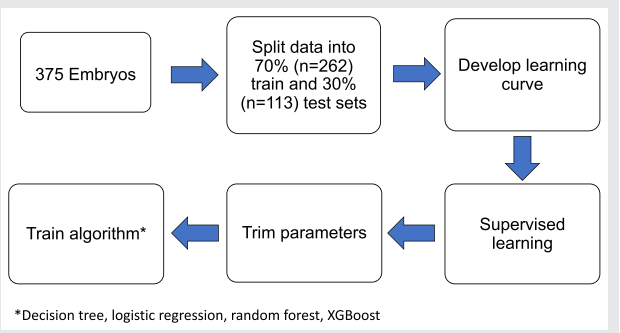
FIGURE 1



Visual abstract of morphokinetic data gathering. Illustration of embryo development following fertilization and the morphokinetic parameters studied, including individual time points.

Yang. Embryo selection by machine learning. Fertil Steril Rep 2022.

FIGURE 2



Workflow diagram of supervised machine learning.

Yang. Embryo selection by machine learning. Fertil Steril Rep 2022.

Ovarian Stimulation and Oocyte Retrieval

Patients included were women who underwent oocyte retrieval and subsequent embryo transfer at a single site academic fertility center. For each stimulation, an individualized antagonist ovarian stimulation protocol was employed using follicle-stimulating hormone follitropin alfa (Gonal F; EMD Serono, Rockland, MA) or (Follistim; Merck, and Co.), and follicle-stimulating hormone/luteinizing hormone menotropin (Menopur; Ferring Pharmaceuticals, NJ). Antagonist cetrorelix acetate (Cetrotide; Freedom Fertility Pharmacy, Byfield, MA) was initiated when a lead follicle reached 13 mm. Oocyte trigger was performed with 3–10,000 units of chorionic gonadotropin (Novarel, Ferring Pharmaceuticals) or leuprolide 36 hours before oocyte retrieval when two or more follicles measured greater than 18 mm and in conjunction with appropriate estradiol values. Follicular number and size were obtained via transvaginal ultrasound during ovarian stimulation monitoring, and serum samples were collected via peripheral venipuncture taken within 1 hour of the ultrasound measurement. Transvaginal oocyte aspiration was performed under total intravenous sedation 36 hours post-trigger. The demographic data collected included age and the patient's medical diagnosis.

Fertilization and Embryo Incubation

Oocyte-cumulus complexes were cultured for 3 hours in 60 mm organ culture dishes containing Quinn's Advantage Fertilization Media (Cooper Surgical, Trumbull, CT) with 10% serum protein supplement (Cooper Surgical) and overlaid with OVOIL (Vitrolife, Englewood, CO) in a Planer BT37 benchtop incubator (Cooper Surgical).

After exactly 3 hours, oocytes were denuded using hyaluronidase (Cooper Surgical) and assessed for nuclear maturity as evidenced by the presence of a polar body. Identified metaphase II oocytes underwent intracytoplasmic sperm injection (ICSI) or conventional insemination. After ICSI, the oocytes were immediately placed in the EmbryoScope time-lapse incubator (Vitrolife, Englewood, CO) and were cultured in Sage 1-Step media at 37 oC in 5.5% CO_2 and 6% O_2. Embryos were assessed beginning at 16–18 hours after insemination using the EmbryoScope viewer. For conventional insemination, zygotes were placed in the EmbryoScope after fertilization. Embryo quality was graded on day 3, day 5, and day 6 using established morphological grading criteria (43, 44).

From April 2014 to July 2015, embryos were cultured in sequential media (Quinn's Advantage Plus media and Quinn's Advantage Blastocyst Plus media). From September 2015 onwards, embryos were cultured in Sage 1-Step media. Both culture conditions were compared, and no changes in blastocyst development were observed locally (45), similar to findings at other centers (46).

Embryos selected for preimplantation genetic testing (PGT) underwent assisted hatching by laser on day 3 and were biopsied on day 5 or 6. Embryos were vitrified with Vit-Freeze kits (Irvine Scientific, Santa Ana, CA) using the manufacturer's protocol. Embryos for frozen transfers were warmed using Vit-Thaw Kit (Irvine Scientific) at least 3 hours before the transfer. After warming, assisted hatching was performed by laser (Hamilton-Thorne Beverly, MA) on embryos that had not been previously biopsied for PGT.

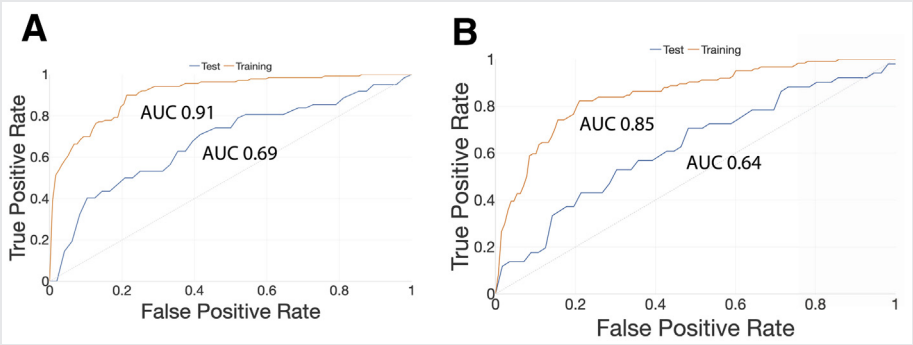
Time-Lapse Microscopy and Morphokinetic Parameters

The interval image acquisition for the EmbryoScope was every 10 minutes at seven focal planes; images were obtained starting from the time after fertilization until the embryo was either transferred or cryopreserved. All morphokinetic parameters starting from the fading of pronuclei to blastocyst formation were annotated manually using EmbryoViewer software. The morphokinetic parameters were determined by a team of trained personnel that included two physicians and one embryologist. Morphokinetic annotation, measurement consistency, and internal quality control were verified by the senior embryologist to decrease annotator bias. For ICSI, t0 was the time of insemination, and for conventional in vitro fertilization (IVF), t0 was defined as the time of addition of sperm to oocytes, similar to other reported studies (23, 47, 48). Parameter information collected included: pronuclei fade (PNf), time to 2-cell (t2), t3, t4, t5, t8, t9, time to morula (defined as 50% cells with indistinct membranes), time to start of early blastulation, time to blastocyst (Fig. 1).

Embryo Transfer and Clinical Outcomes

All embryo transfers were performed under ultrasound guidance as fresh (n = 30) or frozen (n = 350) transfers. All frozen embryos were transferred on day 5 after progesterone administration of approximately 5 days. Luteal support was provided by vaginal progesterone gel (Crinone, Juniper Pharmaceuticals) and continued if pregnancy occurred. Embryo selection was based on a standard grading system of morphologic characteristics (43, 49) for both PGT for aneuploidy (PGT-A) and non-PGT-A tested embryos. Serum β human chorionic gonadotropin (hCG) levels were obtained 9–10 days after embryo transfer. Once a positive β-hCG was established, a repeat hCG test was performed 48–72 hours later, and

FIGURE 3



Receiver operating characteristic (ROC) curve based on clinical pregnancy and live birth rate outcomes. Two ROC curves developed from the training and test sets based on morphokinetics in *orange* and *blue*, respectively, based on clinical pregnancy outcome (**A**). Two ROC curves developed from the training and test sets were based on morphokinetic parameters of time to pronuclei fade, t4, time to morula, t9, and time to blastocyst (**B**). The area under the curve (AUC) is noted next to the graph. The x-axis is the true positive rate. y-axis is the false positive rate (1 − specificity).

Yang. Embryo selection by machine learning. Fertil Steril Rep 2022.

then a transvaginal ultrasound was performed between 6 and 7 weeks of gestational age to determine the presence of clinical pregnancy as defined by a detectable fetal heartbeat.

Supervised Classification Training

Embryo morphokinetic data were randomly divided into 70% training sets and 30% test sets (Fig. 2). The ratio of data assignment to training and test was set by the operator (70%/30%), and this parameter was entered into a ranger. Embryo assignment was achieved with the use of a *random* module in python based on the Mersenne Twister function that assigns a pseudo-random number starting with an original number or seed, such as 1, that changes with each iteration for each embryo (50). Learning curves were developed using scikit-learn (51) to assess the optimal number of training samples, which was at least 150 samples. (Supplemental Fig. 1, available online). Total variables used included all morphokinetic time points. Data were trimmed manually by removing time points that had high collinearity with variable inflation factor >10 which resulted in the use of only six-time points (Supplemental Fig. 2A). Binary outcomes included either positive or negative pregnancy (Supplemental Fig. 2B). Models trained included logistic regression, XGBoost, decision tree, and random forest using Exploratory and ranger (52) (Fig. 2). For each model, the training set was used to develop an algorithm using the training variables and applied to the data on the test set.

Model Comparison

For each model, the predicted probability of pregnancy was calculated for each embryo and compared with the actual outcome. A receiver operating characteristic curve was developed by graphing the true positive rate (sensitivity) with the false positive rate (1 − specificity). The area under the curve (AUC) was calculated for each receiver operating characteristic curve.

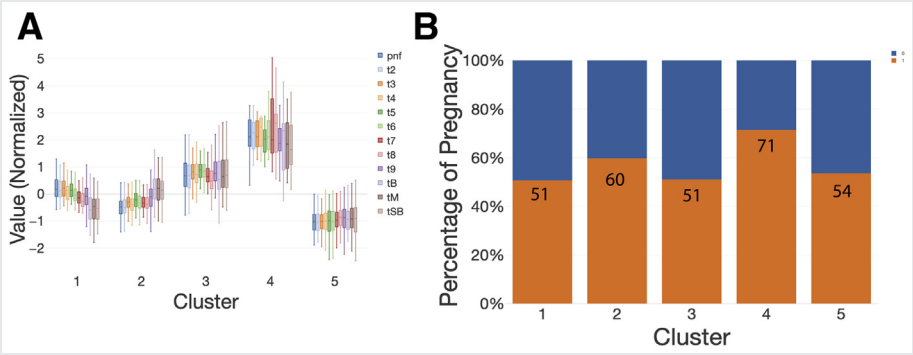
Unsupervised Clustering

All morphokinetic time points were normalized and used in K-means clustering (Supplemental Fig. 3A). To determine the number of clusters, the average sum of squares (distances between all points per cluster) was graphed against the number of clusters using the Elbow method (53) (Supplemental Fig. 3B). Thus, 5 clusters were selected. The Hartigan-Wong algorithm was used to group the data into 5 clusters with a random seed of 1 and a maximum iteration of 10. The normalized number of hours per cluster and the pregnancy rate per cluster were calculated.

RESULTS

Out of 479 embryos surveyed, a total of 367 embryos met the inclusion criteria for the current study (Supplemental Fig. 4A). Embryos were transferred based on morphology. Embryo development from fertilization to blastocyst formation (Fig. 1) was manually annotated hours postinsemination. No differences were observed between the frequency of donor eggs, patient age, and diagnoses (Supplemental Fig. 4B). Multiple types of algorithms were tested and compared, including logistic regression, XGBoost, decision tree, and random forest (Supplemental Table 1), using six morphokinetic parameters as described in the Methods section. The best performing algorithm was random forest, which was named Yang-Peavey Embryo Enhancement Algorithm, with an AUC of 0.91 and 0.69 in the training and test sets, respectively (Fig. 3A). The calculated sensitivity was 65%, specificity 60%, positive predictive value 74%, and negative predictive value 50% from the validation data set only, not including the training set. To address whether PGT-A or single embryo transfer status affected the quality of the algorithm, we incorporated both statuses as covariates (Supplemental Fig. 5B and C, respectively). Similar AUC was achieved (0.92). However, lower predictive power was noted in the test sets as the sample size reached n = 4 in one of the arms (Supplemental Figs. 6 and 7). When the algorithm was applied to live birth rate data,

FIGURE 4



Unsupervised clustering characteristics. Distribution of values of each normalized morphokinetic parameter grouped by cluster (**A**). Box plot includes the first and third quartile with minimum and maximum. The *center* of the box plot presents the median. The x-axis is a cluster. The y-axis is the normalized value. The key for each morphokinetic time point is listed in the figure legend—frequency of embryos with positive outcomes in each cluster (**B**). *Blue* indicates no pregnancy. *Orange* indicates pregnancy.

Yang. Embryo selection by machine learning. Fertil Steril Rep 2022.

we found similar predictive power with the AUC of 0.85 and 0.64 in the training and test sets, respectively (Fig. 3B).

For our secondary analysis, we sought to understand the pattern of embryo growth in an unbiased manner by unsupervised machine learning. We asked whether the embryos could be grouped based on shared morphokinetic characteristics. Given the large range of the number of hours from 19 hours to 160 hours, the data were normalized or rescaled from −3 to 3. Data were grouped into 5 clusters by the distance between each time point (range, n = 21 to n = 97) (Fig. 4B). The normalized morphokinetic values were compared for each cluster showing high and low normalized values (Fig. 4A). Notably, cluster 4 had higher normalized values or slower time points, and cluster 5 had lower values or faster time points (Fig. 4A). The differences in pregnancy rates among the 5 clusters were not statistically significant (χ^2 test P=.356). When we compared the pregnancy rates among the clusters, cluster 4 (71%) and cluster 2 (60%) had a trend of higher rates of pregnancy, whereas clusters 1 (51%) and cluster 3 (51%) had lower rates (Fig. 4B). In this patient population, a small group of embryos (cluster 4, n = 25) with high rates of pregnancy (71%) developed overall more slowly, whereas another cohort of embryos (cluster 2, n = 78) with faster compaction and blastulation had a 60% pregnancy rate (Fig. 4B and Supplemental Table 2) and may represent intrinsic embryo qualities for future hypothesis testing. Overall, cluster 5 represented a group of embryos with faster overall morphokinetics, and cluster 4 represented those with slower morphokinetics (Supplemental Fig. 8).

DISCUSSION

We show that simple machine learning can be used to develop an algorithm to predict clinical pregnancy rates with as few as 200 embryo outcomes with AUCs of 0.91 and 0.69, respectively, in the training and test cohorts. The algorithm also predicted live birth rate outcomes. When PGT-A and single embryo transfer statuses of each embryo were incorporated, the predictive power of the training set declined, likely due to the overfitting of the data because of the small samples size ranging from 4 to 50 patients per arm (Supplemental Figs. 6 and 7). Embryos that had faster blastulation and compaction and a select few with slower development were also associated with favorable pregnancy outcomes. A small, distinct group of embryos with slower development was noted to have higher pregnancy rates, although it was not statistically significant. This is in contrast to other studies that have reported faster development associated with more favorable outcomes (19). Most likely, there was sample size bias, or there was likely an embryo-related biological bias associated with this small group of embryos, or other factors such as uterine or environmental factors unrelated to the embryo.

This algorithm is unique in that it distills innumerable clinical, laboratory, and genetic factors into a few quintessential features to predict pregnancy, assuming that these characteristics are represented by individualized embryo morphokinetics. Many studies use over 32 demographic, laboratory, or grading criteria in the algorithms (1, 30, 47) or complex imaging processing (35, 36), which increases model learning difficulty requiring complex calculations (41). This algorithm can be run on any platform on-site at the laboratory and can be quickly adapted or retrained to the changes in laboratory conditions such as media with input from only 200 embryos. This technology is also noninvasive and may aid in embryo selection after PGT-A. Further tests are needed to determine whether the Yang-Peavey Embryo Enhancement Algorithm can predict pregnancy rates in embryos that did or did not undergo PGT-A.

Whereas previous studies emphasized single parameters or ratios of time points as the best predictors of blastulation (8, 54) or implantation (10, 17) or that parameters cannot be used to predict blastulation (55), our data demonstrate that a prediction model is best developed from multiple morphokinetic time points (56) and can successfully predict pregnancy.

Reassuringly, our model is also consistent with other reports of time to morula and time to blastocyst time points being predictive of implantation (17, 19, 20, 57). Further work is needed to explore whether a few key morphokinetic parameters can universally predict implantation using the built-in EmbryoViewer software (58, 59).

A limitation of this study is that all transferred embryos underwent morphologic grading, which influenced embryo selection. Our findings were made after standard morphologic grading. Future randomized studies are needed to compare predictions between the algorithm and morphological grading. Other limitations include multiple transfers per patient and donor embryos that could be affected by maternal factors. Another limitation is that laboratory parameters included many varying treatments, including ICSI and conventional insemination, PGT-A or untested, assisted hatching, fresh and frozen embryo transfer cycles, and fresh and frozen oocytes.

Time of insemination (t = 0) in conventional IVF may affect or skew the embryo morphokinetics compared with that of ICSI. Some studies reported up to a 1.5-hour delay in early morphokinetics and up to 4.1 hours of advancement in development to the blastocyst stage in conventional IVF compared with ICSI (60). Because of conventional insemination, only time points after time to PNf (48) were included in our study. In our data, a 1.5-hour difference is within the SD (1.7 to 3.5 hours) at the time to PNf, and a 4.1-hour advancement is also within the SD at the blastocyst stage (3.6 to 6.5 hours) observed among the morphokinetic groups by unsupervised clustering (Supplemental Table 2). However, more validation studies are needed to determine whether t0 in conventional IVF significantly impacts morphokinetics from that of ICSI. For instance, to control for the exact time of t0, one would need to study embryos that underwent ICSI only for future larger cohorts as part of a continuously updated, dynamic algorithm.

Assisted hatching was performed for all samples undergoing PGT testing and could have been accounted for by studying PGT status. In the analysis, including PGT-A status, we did not observe an improvement in predictive power because of low sample numbers. The addition of confounding factors likely depletes the accuracy of the model given the low representative sample numbers. Therefore, it would be useful to analyze the effect of assisted hatching with a larger data set. In addition, sex selection based on PGT-A results could confound results. However, in our cohort, the few patients who requested sex selection had embryos that were transferred based on morphological grading. These factors could be included in the algorithm for future predictions with larger cohorts or applied to a cohort with similar characteristics as the current cohort. If conditions change, the algorithm would need to be updated to be applied prospectively. One of the challenges of applying this algorithm is the need for manual annotation for each embryo prospectively which requires additional time by a trained embryologist or technical staff. Unfortunately, at this facility, there was no access to software that performed automated annotations. In the future, the algorithm could be combined with automated annotation software to improve efficiency.

Another limitation is a lack of radiological examination to confirm zygosity to distinguish the difference between two embryo transfers that resulted in monozygotic duplication and loss of the other embryo compared with true implantation of both embryos.

CONCLUSION

We demonstrate that embryo morphokinetics are associated with pregnancy outcomes in a dynamic machine learning algorithm that is specific to a clinic. We also showed that embryos that exhibit slow overall morphokinetics (time to blastocyst = 122 hours vs. 97 hours) had higher rates of pregnancy (71% vs. 54%). As a proof of concept, machine learning can be used at a local IVF center to design a selection algorithm with as few as 200 embryos. As the sample size and the parameters and characteristics change with time, a new model will need to be continually optimized. Rather than using a single universal model, we demonstrate that a model tailored to the local embryo and laboratory characteristics with continuous updates could be used.

Acknowledgments: The authors thank all the members of the Texas Children's Hospital Pavilion for Women's Family Fertility Center staff, including Kim Akina, Katie Walworth, Daneeka Hamilton, Blesson Chellakkan Selvanesan, and Amanda David. They also thank the Division of Reproductive Endocrinology and Infertility and the Department of Obstetrics and Gynecology Residency Program for their support.

REFERENCES

1. Herrero J, Meseguer M. Selection of high potential embryos using time-lapse imaging: the era of morphokinetics. Fertil Steril 2013;99:1030–4.
2. Tesarik J, Greco E. The probability of abnormal preimplantation development can be predicted by a single static observation on pronuclear stage morphology. Hum Reprod 1999;14:1318–23.
3. Lundin K, Bergh C, Hardarson T. Early embryo cleavage is a strong indicator of embryo quality in human IVF. Hum Reprod 2001;16:2652–7.
4. Isiklar A, Mercan R, Balaban B, Alatas C, Aksoy S, Urman B. Early cleavage of human embryos to the two-cell stage. A simple, effective indicator of implantation and pregnancy in intracytoplasmic sperm injection. J Reprod Med 2002;47:540–4.
5. Jackson KV, Ginsburg ES, Hornstein MD, Rein MS, Clarke RN. Multinucleation in normally fertilized embryos is associated with an accelerated ovulation induction response and lower implantation and pregnancy rates in vitro fertilization-embryo transfer cycles. Fertil Steril 1998;70:60–6.
6. Yakin K, Balaban B, Urman B. Impact of the presence of one or more multinucleated blastomeres on the developmental potential of the embryo to the blastocyst stage. Fertil Steril 2005;83:243–5.
7. Palmstierna M, Murkes D, Csemiczky G, Andersson O, Wramsby H. Zona pellucida thickness variation and occurrence of visible mononucleated blastomers in preembryos are associated with a high pregnancy rate in IVF treatment. J Assist Reprod Genet 1998;15:70–5.
8. 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.
9. Azzarello A, Hoest T, Mikkelsen AL. The impact of pronuclei morphology and dynamicity on live birth outcome after time-lapse culture. Hum Reprod 2012;27:2649–57.
10. 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.

11. Sayed S, Reigstad MM, Petersen BM, Schwennicke A, Wegner Hausken J, Storeng R. Time-lapse imaging derived morphokinetic variables reveal association with implantation and live birth following in vitro fertilization: a retrospective study using data from transferred human embryos. PLoS One 2020; 15:e0242377.
12. Dal Canto M, Coticchio G, Mignini Renzini M, 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.
13. 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.
14. Chappell NR, Barsky M, Shah J, Peavey M, Yang L, Sangi-Haghpeykar H, et al. Embryos from polycystic ovary syndrome patients with hyperandrogenemia reach morula stage faster than controls. F&S Reports 2020;1:125–32.
15. Jacobs C, Nicolielo M, Erberelli R, Mendez F, Fanelli M, Cremonesi L, et al. Correlation between morphokinetic parameters and standard morphological assessment: what can we predict from early embryo development? A time-lapse-based experiment with 2085 blastocysts. JBRA Assist Reprod 2020;24:273–7.
16. Rubio I, Kuhlmann R, Agerholm I, Kirk J, Herrero J, Escriba MJ, et al. Limited implantation success of direct-cleaved human zygotes: a time-lapse study. Fertil Steril 2012;98:1458–63.
17. Basile N, Vime P, Florensa M, Aparicio Ruiz B, Garcia Velasco JA, Remohi 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.
18. 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.
19. Campbell A, Fishel S, Bowman N, Duffy S, Sedler M, Hickman CF. Modelling a risk classification of aneuploidy in human embryos using non-invasive morphokinetics. Reprod Biomed Online 2013;26:477–85.
20. Gazzo E, Pena F, Valdez F, Chung A, Bonomini C, Ascenzo M, et al. The Kidscore™ D5 algorithm as an additional tool to morphological assessment and PGT-A in embryo selection: a time-lapse study. JBRA Assist Reprod 2020;24:55–60.
21. Kaser DJ, Racowsky C. Clinical outcomes following selection of human preimplantation embryos with time-lapse monitoring: a systematic review. Hum Reprod Update 2014;20:617–31.
22. Racowsky C, Stern JE, Gibbons WE, Behr B, Pomeroy KO, Biggers JD. National collection of embryo morphology data into Society for Assisted Reproductive Technology Clinic Outcomes Reporting System: associations among day 3 cell number, fragmentation and blastomere asymmetry, and live birth rate. Fertil Steril 2011;95:1985–9.
23. Lee CI, Chen CH, Huang CC, Cheng EH, Chen HH, Ho ST, et al. Embryo morphokinetics is potentially associated with clinical outcomes of single-embryo transfers in preimplantation genetic testing for aneuploidy cycles. Reprod Biomed Online 2019;39:569–79.
24. Hernández-Vargas P, Muñoz M, Domínguez F. Identifying biomarkers for predicting successful embryo implantation: applying single to multi-OMICs to improve reproductive outcomes. Hum Reprod Update 2020;26:264–301.
25. Alegre L, Del Gallego R, Arrones S, Hernández P, Muñoz M, Meseguer M. Novel noninvasive embryo selection algorithm combining time-lapse morphokinetics and oxidative status of the spent embryo culture medium. Fertil Steril 2019;111:918–27.e3.
26. Bartolacci A, Dal Canto M, Guglielmo MC, Mura L, Brigante C, Mignini Renzini M, et al. Early embryo morphokinetics is a better predictor of post-ICSI live birth than embryo morphology: speed is more important than beauty at the cleavage stage. Zygote 2021;29:495–502.
27. Coticchio G, Pennetta F, Rizzo R, Tarozzi N, Nadalini M, Orlando G, et al. Embryo morphokinetic score is associated with biomarkers of developmental competence and implantation. J Assist Reprod Genet 2021;38:1737–43.
28. McQueen DB, Mazur J, Kimelman D, Confino R, Robins JC, Bernardi LA, et al. Can embryo morphokinetic parameters predict euploid pregnancy loss? Fertil Steril 2021;115:382–8.
29. 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.
30. Blank C, Wildeboer RR, DeCroo I, Tilleman K, Weyers B, de Sutter P, et al. Prediction of implantation after blastocyst transfer in in vitro fertilization: a machine-learning perspective. Fertil Steril 2019;111:318–26.
31. Hernández-González J, Inza I, Crisol-Ortíz L, Guembe MA, Iñarra MJ, Lozano JA. Fitting the data from embryo implantation prediction: learning from label proportions. Stat Methods Med Res 2018;27:1056–66.
32. Giscard d'Estaing S, Labrune E, Forcellini M, Edel C, Salle B, Lornage J, et al. A machine learning system with reinforcement capacity for predicting the fate of an ART embryo. Syst Biol Reprod Med 2021;67:64–78.
33. Feyeux M, Reignier A, Mocaer M, Lammers J, Meistermann D, Barrière P, et al. Development of automated annotation software for human embryo morphokinetics. Hum Reprod 2020;35:557–64.
34. VerMilyea M, Hall JMM, Diakiw SM, Johnston A, Nguyen T, Perugini D, et al. 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;35:770–84.
35. Khosravi P, Kazemi E, Zhan Q, Malmsten JE, Toschi M, Zisimopoulos P, et al. Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization. NPJ Digit Med 2019;2:21.
36. 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;34:1011–8.
37. Bori L, Paya E, Alegre L, Viloria TA, Remohi JA, Naranjo V, et al. Novel and conventional embryo parameters as input data for artificial neural networks: an artificial intelligence model applied for prediction of the implantation potential. Fertil Steril 2020;114:1232–41.
38. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. Nat Rev Clin Oncol 2019;16:703–15.
39. Chéles DS, Molin EAD, Rocha JC, Nogueira MFG. Mining of variables from embryo morphokinetics, blastocyst's morphology and patient parameters: an approach to predict the live birth in the assisted reproduction service. JBRA Assist Reprod 2020;24:470–9.
40. Barrie A, McDowell G, Troup S. An investigation into the effect of potential confounding patient and treatment parameters on human embryo morphokinetics. Fertil Steril 2021;115:1014–22.
41. Cai J, Luo J, Wang S, Yang S. Feature selection in machine learning: a new perspective. Neurocomputing 2018;300:70–9.
42. Ali M, Aittokallio T. Machine learning and feature selection for drug response prediction in precision oncology applications. Biophys Rev 2019; 11:31–9.
43. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. Hum Reprod 2011;26:1270–83.
44. Racowsky C, Vernon M, Mayer J, Ball GD, Behr B, Pomeroy KO, et al. Standardization of grading embryo morphology. J Assist Reprod Genet 2010;27: 437–9.
45. Kaskar K, Hamilton DP, Miller K, Zarutskie PW, Gibbons WE. Blastocyst development using sequential media versus one-step media in Embryoscope and Planer incubators. Fertil Steril 2016;103.
46. Sfontouris IA, Martins WP, Nastri CO, Viana IG, Navarro PA, Raine-Fenning N, et al. Blastocyst culture using single versus sequential media in clinical IVF: a systematic review and meta-analysis of randomized controlled trials. J Assist Reprod Genet 2016;33:1261–72.
47. 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.
48. 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.e1.
49. Gardner DK, Schoolcraft WB. Culture and transfer of human blastocysts. Curr Opin Obstet Gynecol 1999;11:307–11.

50. Matsumoto M, Nishimura T, editors. Mersenne Twister - a pseudo random number generator and its variants; 2009.
51. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. J Mach Learn Res 2011;12:2825–30.
52. Wright MN, Ziegler A. ranger: a fast implementation of random forests for high dimensional data in C++ and R. J Stat Softw 2017;77.
53. Ketchen DJ, Shook CL. The application of cluster analysis in strategic management research: an analysis and critique. Strateg Manag J 1996;17:441–58.
54. Cruz M, Garrido N, Herrero J, Perez-Cano I, Munoz M, Meseguer M. Timing of cell division in human cleavage-stage embryos is linked with blastocyst formation and quality. Reprod Biomed Online 2012;25:371–81.
55. Rienzi L, Capalbo A, Stoppa M, Romano S, Maggiulli R, Albricci L, et al. No evidence of association between blastocyst aneuploidy and morphokinetic assessment in a selected population of poor-prognosis patients: a longitudinal cohort study. Reprod Biomed Online 2015;30:57–66.
56. Milewski R, Kuć P, Kuczyńska A, Stankiewicz B, Łukaszuk 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.
57. Meseguer M, Herrero J, Tejera A, Hilligsøe KM, Ramsing NB, Remohi J. The use of morphokinetics as a predictor of embryo implantation. Hum Reprod 2011;26:2658–71.
58. 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.
59. 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.
60. 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.e1–2.