Original Article

# Optimal Embryo Selection: The Irreplaceable Role of the Embryologist in an Age of Advancing Technology

*Lauren Kendall Rauchfuss, Yulian Zhao, David Walker, Terri Galantis, Jolene Fredrickson, Kathrynne Barud, Chandra Shenoy*

Mayo Clinic, Rochester, MN, USA

**Abstract**

**Background:** Time-lapse incubators allow for ongoing evaluation of embryos without culture condition disruption. The use of time-lapse incubation has been shown to improve outcomes either by improving overall conditions or providing additional information to aid in embryo selection for transfer. Time-lapse incubators can also utilise morphokinetic models to rank embryos based on morphokinetic parameters. We sought to compare a morphokinetic model for embryo comparison to traditional morphologic evaluation. **Aims:** The aim of the study is to compare a morphokinetic model for embryo comparison to traditional morphologic evaluation. **Settings and Design:** This is a retrospective cohort design. **Materials and Methods:** Embryos cultured in a time-lapse culture system that had traditional morphologic evaluation, morphokinetic modelling and known live birth outcomes were included in this study. Embryos with unknown competence were excluded, including when two embryos were transferred with a single live birth resulted. **Statistical Analysis Used:** Receiver operating characteristic (ROC) curves were determined for both the morphologic analysis and the morphokinetic model on culture day 3 and day 5. Using the ROC-determined cutoff that optimised both sensitivity and specificity, a binary outcome for each test was analysed using agreement statistics to determine if one method of embryo evaluation was superior to the other. **Results:** Morphological and morphokinetic grading were both predictive of embryo competence on days 3 and 5. However, on day 3, morphologic grading was superior to morphokinetic grading with area under the curve (AUC) of 0.66 ($P < 0.001$) and 0.58 ($P = 0.009$), respectively. Contrarily, on day 5, the morphokinetic model had a higher AUC of 0.65 ($P = 0.03$) compared to the morphologic grading, AUC 0.56 ($P = 0.02$). **Conclusion:** Traditional morphology was noted to be a better diagnostic tool (higher AUC) on culture day 3 while a morphokinetic model was superior on day 5.

**Keywords:** *Embryoscope, morphokinetics, morphologic grading*

## Introduction

Innovative advancements have sparked many fields to utilise technology to maximise efficiency and optimise patient outcomes. The field of fertility treatment is often thought to be at the forefront of technological advancements in medicine. For years, embryoscopes with time-lapse technology have been used for *in vitro* culture of embryos created through assisted reproductive technology. This storage device provides a stable culture environment with time-lapse capability, enabling the embryologists to observe

***Address for correspondence:*** Dr. Lauren Kendall Rauchfuss, Mayo Clinic, Rochester, MN, USA.
E-mail: kendall.lauren@mayo.edu

**Received:** 01-08-2023 **Revised:** 26-08-2023
**Accepted:** 07-09-2023 **Published:** 29-09-2023

Access this article online

Quick Response Code:

Website: www.jhrsonline.org

DOI: 10.4103/jhrs.jhrs_98_23

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

**For reprints contact:** WKHLRPMedknow_reprints@wolterskluwer.com

**How to cite this article:** Rauchfuss LK, Zhao Y, Walker D, Galantis T, Fredrickson J, Barud K, *et al.* Optimal embryo selection: The irreplaceable role of the embryologist in an age of advancing technology. J Hum Reprod Sci 2023;16:227-32.

embryo development at each moment of development, without disrupting the embryos.[1] Several considerations are important when assessing embryo development to decide which embryo(s) should be transferred to the uterus or cryopreserved for future use.

The factors considered while assessing embryos include both morphologic assessment and morphokinetic parameters. Morphologic grading is the evaluation of the characteristics of the embryo, including degree of fragmentation, overall expansion, tightness of the cells in the inner cell mass and number and cohesiveness of the cells in the trophectoderm. These characteristics each make up a component of grading that is reported as a number or letter grade.[2] Morphologic scoring has been shown to be a relevant way to assess embryo competence.[3,4] In addition, morphokinetic patterns such as cleavage patterns and time to blastulation over time are taken into account.[5-9] The embryoscope is especially helpful in observing growth pattern over time, allowing for a more thorough assessment of embryo development process. Not only can the embryoscope record and display the embryos without removing them but it can also utilise a morphokinetic formula to grade the embryos and produce a morphokinetic score. This score can be utilised in conjunction with morphologic scoring performed by the embryologist.

Both morphologic and morphokinetic grading are quintessential for optimal patient care. In this study, we sought to compare the traditional morphologic grading performed by an embryologist to a morphokinetic model for grading used by the embryoscope by retrospectively reviewing embryo grading for patients with clinical pregnancy and live birth outcomes.

## Materials and Methods

After obtaining IRB approval (IRB number 15-007799), a retrospective chart review was performed in accordance with the Principles of Helsinki Declaration. Patients who did not consent to the use of medical record information for research purposes were excluded from the review. As this was a retrospective study, sample size was determined by available data rather than by power analysis. The inclusion criteria were patients who underwent fresh embryo transfer with known live birth outcomes from August 2014 to August 2016. Patients with unknown live birth outcomes, multiple embryos transferred with live birth outcome that did not correlate with number of embryos transferred and frozen embryo transfers were excluded. An embryo that was transferred and resulted in live birth was considered competent. Incompetent embryos did not result in live birth. Single, double and triple embryo transfers were included if competency could be determined. Only double or triple embryo transfers with a negative pregnancy test or live births with an equivalent number of live-born neonates to embryos transferred were included. Multiple embryo transfers where that had some embryos achieve implantation and others did not, were excluded. For example, if a double embryo transfer was performed and a single live birth was achieved, both embryos were excluded from the analysis due to incomplete competency determination.

Embryos included were cultured in a time-lapse culture system, had traditional morphologic evaluation and had mathematical modelling performed to determine a morphokinetic grade on day 3 and day 5. Risk of bias was reduced by utilising embryos graded by multiple embryologists with differing opinions reviewed and final decision made by the laboratory director.

An internally derived grading system was developed for day 3 embryos based on factors predictive of live birth as shown in the literature.[10,11] Gardner grading system[12] was utilized for day 5. On day 3, a morphologic score is 0, 0.25–1.0, 1.25–2.0 or <2.0 based on the number of cells, percent fragmentation and symmetry[10,11] [Supplementary Table 1]. Day 5 grading considered stage of embryo expansion, size, shape and compaction of the inner cell mass and distribution, scalloping and cell numbers of the trophectoderm. On day 5, a score of 1, 2, 3, 4, 5 or 6 was assigned to the inner cell mass [Supplementary Table 2] and 1, 2, 3, 4 or 5 assigned to the trophectoderm [Supplementary Table 3].

Throughout embryo culture, the embryoscope assigned a score based on the morphokinetic grade based on the embryoscope's compare and select model in parallel to the embryologist morphologic grading. The morphokinetic formula for day 3 and day 5 was based on the parameters that are known to be characteristics of good-quality embryos based on the literature and a review of internal data. This included time from insemination to division of 2 cells,[6,13] time from 2 cells to 3 cells,[13] and whether or not the embryo had a certain number of cells at different time points.[2] The day 5 morphokinetic formula took into account all the factors and scores assigned on day 3 as well as time from insemination to initiation and completion of blastulation,[14] time from 3 to 4 cells,[6] time from 3 to 5 cells and a ratio of the time it took for the embryo to develop from 3 to 5 cells compared to 2–3 cells.[15] A weighted score was assigned if the embryo met these predetermined characteristics at time points detailed in Table 1. These values were incorporated into a formula described in Figure 1. Scores were then added together to create a total final score. An embryo could score a maximum score of 1.5 on day 3 and 2.1 on day 5.

**Table 1: Score assignments based on embryo characteristics at time points utilised for the morphopkinetic formula**

| Variable | Embryo characteristics at time points | Score (day 3) | Score (day 5) |
|---|---|---|---|
| t2 | Time from insemination to complete division of 2 cells was 20–32 h | 0.1 | 0.5 |
| cells4 | An embryo had 4 cells at 42–44 h | 0.5 | 0.5 |
| cells8 | An embryo had 7–9 cells at 64–66 h | 0.3 | 0.1 |
| cc2 | Time from a 2-cell to 3-cell embryo was 10–14 h | 0.6 | 0.1 |
| s2 | Time from 3 cell to 4 cell embryo was 0–2 h | 0 | 0.1 |
| cc3 | Time from 3 cell to 5 cell embryo was 10–14 h | 0 | 0.1 |
| tSB | Time from insemination to start of blastulation of 70–105 h | 0 | 0.3 |
| tB | Time from insemination to complete blastulation of 80–115 h | 0 | 0.3 |
| ccRatio | Calculated ratio of 0.08–2.0 when comparing time it took to develop from 3–5 cells to the time it took to develop from 2–3 cells | 0 | 0.1 |
| Total possible score | | 1.5 | 2.1 |

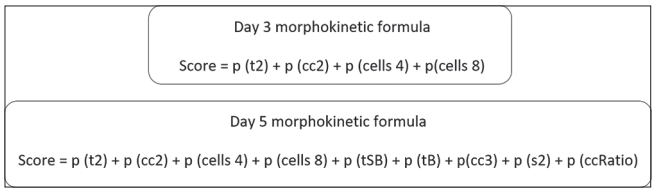



**Figure 1:** The morphokinetic formulas are illustrated here. The scores for day 3 and day 5 are calculated by evaluating the embryo characteristics up until the specified time point. If criteria are met, the weighted points are assigned to the embryo. The points are then added for a summative score. Definitions of each individual parameter and the weight of points assigned are outlined in Table 1

### Statistical analysis

In this study, we compared morphologic grading by the human embryologist to the time-based mathematical machine model by looking at how their grading predicted embryo competence. Age, body mass index (BMI) and anti-Müllerian hormone (AMH) were compared between the two groups using a two-sample *t*-test. To compare the two different grading methods, we evaluated the area under the curve (AUC) for receiver operating characteristic (ROC) curves created for each grading method on both days 3 and 5 of embryo development. Using the ROC-determined cutoff that optimised both sensitivity and specificity, a binary outcome for each test was analysed. Agreement statistics were utilised to determine if one method of embryo evaluation was superior to the other.

## Results

In total, 195 embryos were included from 129 patients: 79 single embryo transfers, 49 double embryo transfers and 6 triple embryo transfers.

Not surprisingly, patients were younger in the competent embryo group, with an average age of 32.9 compared to 34.3 in the incompetent group ($P < 0.0001$). Patients had similar BMI in both groups (28.3 vs. 27.1; $P = 0.80$). Patients with competent embryos had higher AMH values than those with incompetent embryos (5.1 vs. 3.8; $P = 0.034$). The demographics are presented in Table 2.

Day 3 morphologic grading was predictive of embryo competence ($P < 0.001$), ROC AUC was 0.66 with a 95% confidence interval of (0.57–0.74). The ROC curve for day 3 morphology suggests a cutoff grade of 0.5 to predict live birth which gives this test a sensitivity of 53.5% and specificity of 72.7%.

Day 3 morphokinetic grading was also predictive of embryo competence ($P = 0.009$), however, the ROC AUC was letter at 0.58 with a 95% confidence interval of (0.53–0.63). The ROC curve for day 3 morphokinetic grading suggests a cutoff grade of 1.2 to predict live birth which gives this test a sensitivity of 90.7% and specificity of 24.5%.

Day 5 inner cell mass morphologic grading was predictive of embryo competence ($P = 0.02$), ROC AUC was 0.56 with a 95% confidence interval of (0.52–0.60). The ROC curve for day 5 ICM suggests a cutoff grade of 1.0 which gives this test a sensitivity of 97% and a specificity of 16.1%.

Day 5 morphologic trophectoderm grading was not predictive of embryo competence ($P = 0.12$), ROC AUC was 0.52 with a 95% confidence interval of (0.49–0.62). The ROC curve for day 5 trophectoderm grading suggests a cutoff grade of 1 which gives this test a sensitivity of 90.5% and a specificity of 25.8%.

Day 5 morphokinetic modelling was predictive of embryo competence ($P = 0.03$). The ROC AUC was 0.65 with a 95% confidence interval of (0.56–0.73). The ROC curve for day 5 morphokinetic grading suggests a cutoff grade of 2.0 which gives this test a sensitivity of 75% and a specificity of 52.5%. All ROC curves are graphically depicted in Figure 2. Average grades of competent and incompetent embryos are presented in Table 3. Distribution of grades is depicted in histograms in Supplementary Figure 1.

## Discussion

This study highlights the importance of the human experience and judgement in a world that is quickly integrating the use of artificial intelligence in reproductive medicine. This is certainly an exciting time in our field. Faced with advances in embryo selection such as pre-implantation genetic testing and development of artificial intelligence, we must decide how to best utilise technological advancements to enhance care provided by embryologists and clinicians without causing harm.

**Table 2: Patient demographics**

| | Competent | Incompetent | *P* |
|---|---|---|---|
| Age | 32.9 | 34.3 | <0.001 |
| BMI | 28.3 | 27.1 | 0.80 |
| Anti-Müllerian hormone | 5.1 | 3.8 | 0.034 |

*P* value are two sided and were considered significant at <0.05.
BMI: Body mass index

**Table 3: Morphologic and morphokinetic grades assigned for competent and incompetent groups**

| | Competent, mean (SD) | Incompetent, mean (SD) |
|---|---|---|
| Day 3 | | |
| Morphologic grade | 0.37 (0.37) | 0.67 (0.54) |
| Morphokinetic grade | 1.45 (0.16) | 1.3 (036) |
| Day 5 | | |
| Inner cell mass grade | 1.02 (0.15) | 1.16 (0.40) |
| Trophectoderm grade | 1.10 (0.20) | 1.50 (0.11) |
| Morphokinetic grade | 2.03 (0.20) | 1.90 (0.23) |

SD: Standard deviation

For example, pre-implantation genetic testing (PGT) is a tool that allows for the selection of euploid embryos for transfer. In the era of PGT, the practice of grading and cleavage stage transfers have been called into question. However, in clinical practice, many patients do not choose to pursue PGT given the cost and testing limitations such as interpretation of mosaicism. Furthermore, one study showed that performing PGT did not improve live birth rates for patients with poor prognoses compared to embryo scoring alone.[3] Hence, grading of embryos is especially important for this group of patients.

Overall, our study agrees with published data that have shown that both morphologic and morphokinetic parameters have a role in selecting blastocyst embryos for transfer[3,6,16,17] and even on day 3 can predict embryo competence.[7,8] However, when comparing morphologic to morphokinetic parameters on day 3, the expertise of an embryologist was superior to the morphokinetic model

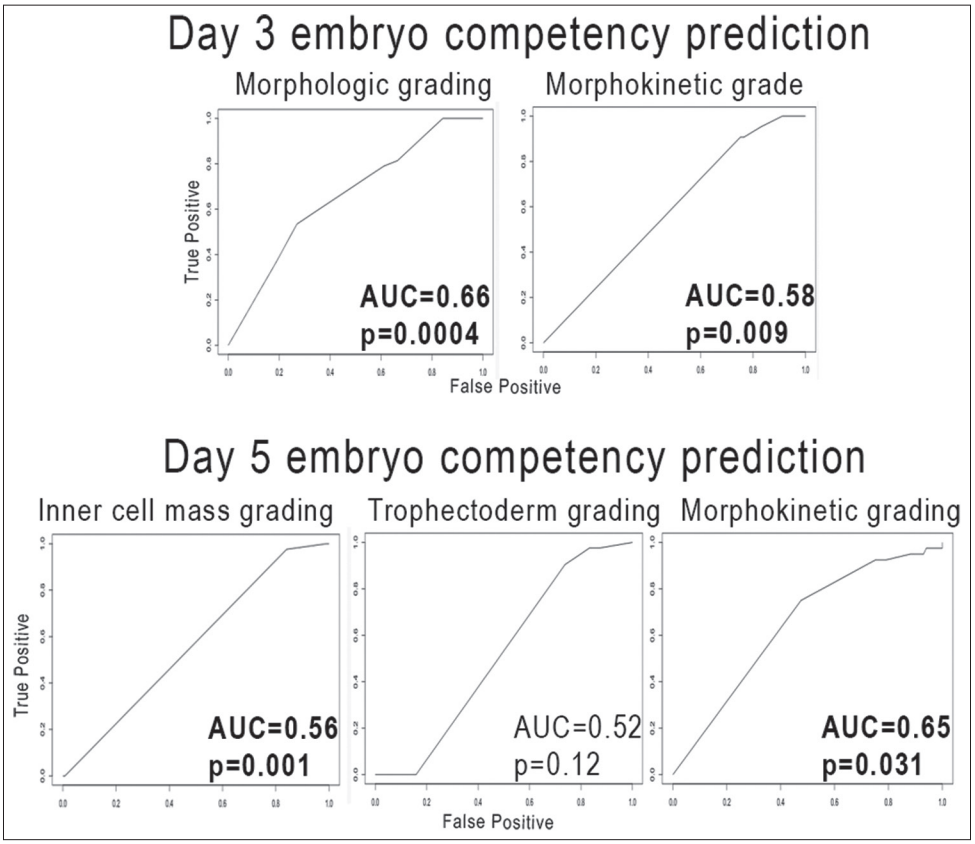



**Figure 2:** The receiver operating characteristic curves are shown for morphologic and morphokinetic grading on day 3 and day 5. AUC: Area under the curve

with an AUC closer to 1 (0.66 vs. 0.58 respectively). An AUC shows how well the test optimises sensitivity or specificity, with a value approaching 1 indicating a more balanced test, optimising both sensitivity and specificity. Although both models were statistically significant indicating each can be taken into account when selecting an embryo.

Furthermore, although blastocyst transfer is considered the standard of care, cleavage stage transfers are still performed with success.[18] Cleavage transfer may be recommended if the embryo's morphokinetic pattern is behind the anticipated division rate or the morphologic characteristics are less than ideal. Therefore, our practice still relies heavily on embryo grading earlier in culture, when there is less data for the morphokinetic model to take into account. In these instances, we see that the morphological scoring by embryologists is crucial.

For day 5 fresh embryo transfers, the morphokinetic grading was superior to morphologic grading, especially when considering the score for trophectoderm morphology. The morphokinetic formula does have more variables to consider for calculation on cycle day 5, likely contributing to its accuracy. The trophectoderm characteristics were not predictive of competency with a $P = 0.12$, for the ROC curve of this test. However, both inner cell mass grading ($P = 0.02$) and morphokinetic grading ($P = 0.03$) were predictive of embryo competence. A recent randomised control trial showed that morphokinetic or morphologic grading resulted in live birth rates that were not statistically different.[19] This is in accordance with our findings that both morphologic and morphokinetic grading has a role in embryo selection.

Ultimately, this study adds to the growing body of literature examining the use of technology as a tool to enhance clinical outcomes. However, it is not without limitations such as retrospective data collection at a single centre. In addition, this study does not account for the endometrium. For an embryo that results in a live birth, competency is clear. However, embryos that fail to result in live birth could be due to embryo incompetency, uterine factors, endometrial factors, immunologic factors or environmental factors or mechanical transfer factors. This is a confounder that is difficult to quantify. Despite this limitation, using live birth as a marker for competence is still a reasonable assumption. When selecting cutoffs for use in clinical practice, sensitivity for competence should be prioritised over specificity. Furthermore, results may vary based on laboratory technique and different patient populations. Whereas different morphokinetic models can be used, only one morphokinetic model was assessed. Patients included in this study underwent fresh embryo transfers, which is standard for all patients at this institution, except when patients choose to pursue pre-implantation genetic testing or fresh embryo transfer cannot be performed due to the risk of ovarian hyperstimulation syndrome. Future prospective, multicentre studies should be considered to evaluate live birth outcomes based on morphokinetic grading versus morphologic grading with a variety of morphokinetic models. In a prospective study, morphologic characteristics could both be assessed by both computer automation and the embryologist and scores could be directly compared.

## Conclusion

Traditional morphology was noted to be a better diagnostic tool on culture day 3 while a morphokinetic model was superior on day 5. Although embryoscope morphokinetic grading is helpful in selecting the best blastocyst embryos for implantation, the skill of the embryologist is essential in selecting embryos for transfer, especially earlier in embryo development, with less morphokinetic data.

### Financial support and sponsorship

Nil.

### Conflicts of interest

There are no conflicts of interest.

### Data availability statement

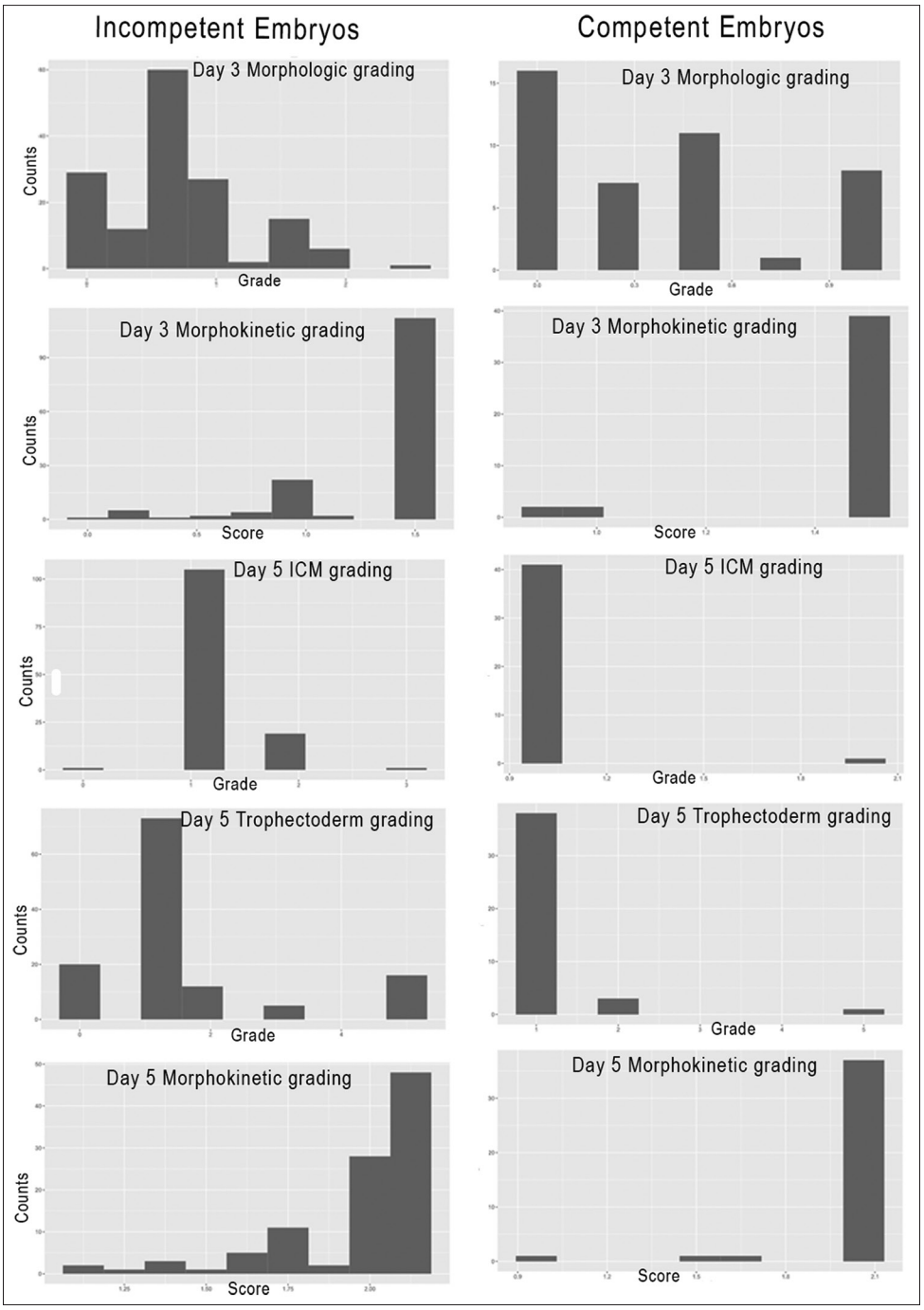
Our data are available for the journal if requested by the editor(s).

## References

1. Zhang JQ, Li XL, Peng Y, Guo X, Heng BC, Tong GQ. Reduction in exposure of human embryos outside the incubator enhances embryo quality and blastulation rate. Reprod Biomed Online 2010;20:510-5.
2. 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.
3. Finn A, Scott L, O'Leary T, Davies D, Hill J. Sequential embryo scoring as a predictor of aneuploidy in poor-prognosis patients. Reprod Biomed Online 2010;21:381-90.
4. Scott L, Finn A, O'Leary T, McLellan S, Hill J. Morphologic parameters of early cleavage-stage embryos that correlate with fetal development and delivery: Prospective and applied data for increased pregnancy rates. Hum Reprod 2007;22:230-40.
5. 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.
6. 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.
7. Cruz M, Garrido N, Herrero J, Pérez-Cano I, Muñoz 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.

8. 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.
9. Hashimoto S, Kato N, Saeki K, Morimoto Y. Selection of high-potential embryos by culture in poly(dimethylsiloxane) microwells and time-lapse imaging. Fertil Steril 2012;97:332-7.
10. Desai NN, Goldstein J, Rowland DY, Goldfarb JM. Morphological evaluation of human embryos and derivation of an embryo quality scoring system specific for day 3 embryos: A preliminary study. Hum Reprod 2000;15:2190-6.
11. Ziebe S, Petersen K, Lindenberg S, Andersen AG, Gabrielsen A, Andersen AN. Embryo morphology or cleavage stage: How to select the best embryos for transfer after *in-vitro* fertilization. Hum Reprod 1997;12:1545-9.
12. Gardner DK. Towards Reproductive certainty: Fertility and genetics beyond 1999: The plenary proceedings of the 11th world congress. Florida: CRC Press; 1999. p. 378.
13. 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.
14. 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.
15. 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.
16. 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.
17. Shioya M, Kobayashi T, Sugiura T, Fujita M, Takahashi K. The morphokinetics algorithm based on data from day 5 blastocyst transfer (KIDScoreD5 version 3) is also useful for embryo selection in day 6 blastocyst transfer. Reprod Med Biol 2022;21:e12484.
18. Neblett MF 2nd, Kim T, Jones TL, Baumgarten SC, Coddington CC, Zhao Y, *et al.* Is there still a role for a cleavage-stage embryo transfer? F S Rep 2021;2:269-74.
19. Ahlström A, Lundin K, Lind AK, Gunnarsson K, Westlander G, Park H, *et al*. A double-blind randomized controlled trial investigating a time-lapse algorithm for selecting Day 5 blastocysts for transfer. Hum Reprod 2022;37:708-17.




**Supplementary Figure 1:** Histograms are shown which demonstrate the distribution of morphologic grades and morphokinetic scores amongst the cohort of embryos. In the first column, we present the distribution of grades amongst our incompetent embryos. The embryos were assessed at two different time points (day 3 and day 5) and with five different scoring systems as detailed in the Materials and Methods. In the right section, we present the distribution of grades amongst our competent embryos. These embryos were assessed with the same time points and scoring systems

**Supplementary Table 1: Morphologic grading for day 3 embryos**

| Grade | Number of cells | Percentage fragmentation | Symmetry |
|---|---|---|---|
| 0 | 8 | 0 | Perfect |
| 0.25–1 | 7–9 | 0–10 | Near perfect |
| 1.25–2.0 | <7 cells or >9 cells | 11–25 | Moderate |
| >2.0 | <4 or >12 cells | 25 | Absent |

**Supplementary Table 2: Morphologic inner cell mass grading for day 5 embryos**

| Grade | Description |
|---|---|
| 1 | Large size/compacted - oblong or spherical |
| 2 | Medium size/compacted |
| 3 | Small size/compacted |
| 4 | Individual blastomeres visible, partial compaction |
| 5 | Loose cells |
| 6 | No cells visible, degraded or necrotic cells |

**Supplementary Table 3: Morphologic trophectoderm grading for day 5 embryos**

| Grade | Description |
|---|---|
| 1 | Even distribution, scalloped, many cells |
| 2 | Uneven distribution, scalloped, many cells |
| 3 | Uneven distribution, partially scalloped, some cells |
| 4 | Uneven distribution, not scalloped, few cells |
| 5 | Dead blastomeres outside, incomplete blastulation |