



Should there be an “AI” in TEAM? Embryologists selection of high implantation potential embryos improves with the aid of an artificial intelligence algorithm

V. W. Fitz¹ · M. K. Kanakasabapathy² · P. Thirumalaraju² · H. Kandula² · L. B. Ramirez³ · L. Boehnlein⁴ · J. E. Swain⁵ · C. L. Curchoe⁶ · K. James⁷ · I. Dimitriadis¹ · I. Souter¹ · C. L. Bormann¹ · H. Shafiee⁸

Received: 18 June 2021 / Accepted: 9 September 2021 / Published online: 17 September 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Purpose A deep learning artificial intelligence (AI) algorithm has been demonstrated to outperform embryologists in identifying euploid embryos destined to implant with an accuracy of 75.3% (1). Our aim was to evaluate the performance of highly trained embryologists in selecting top quality day 5 euploid blastocysts with and without the aid of a deep learning algorithm.

Materials and methods A non-overlapping series of 200 sets of day 5 euploid embryo images with known implantation outcomes was distributed to 17 highly trained embryologists. One embryo in each set was known to have implanted and one failed implantation. They were asked to select which embryo to transfer from each set. The same 200 sets of embryos, with indication of which embryo in each set had been identified by the algorithm as more likely to implant was then distributed. Chi-squared, t-test, and receiver operating curves were performed to compare the embryologist performance with and without AI.

Results Fourteen embryologists completed both assessments. Embryologists provided with AI results selected successfully implanted embryos in 73.6% of cases compared to 65.5% for those selected using visual assessments alone ($p < 0.001$). All embryologists improved in their ability to select embryos with the aid of the AI algorithm with a mean percent improvement of 11.1% (range 1.4% to 15.5%). There were no differences in degree of improvement by embryologist level of experience (junior, intermediate, senior).

Conclusions The incorporation of an AI framework for blastocyst selection enhanced the performance of trained embryologists in identifying PGT-A euploid embryos destined to implant.

Keywords Embryo selection · Assisted reproductive technology · Artificial intelligence · Implantation

✉ C. L. Bormann
cbormann@partners.org

✉ H. Shafiee
hshafiee@bwh.harvard.edu

¹ Division of Reproductive Endocrinology and Infertility, Department of Obstetrics and Gynecology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

² Department of Medicine, Harvard Medical School, Boston, MA, USA

³ Northwell Health Fertility, Manhasset, NY, USA

⁴ Division of Reproductive Endocrinology and Infertility, Department of Obstetrics and Gynecology, University of Wisconsin, Madison, WI, USA

⁵ CCRM Fertility Network, Lone Tree, CO, USA

⁶ CCRM Fertility Orange County, Newport Beach, CA, USA

⁷ The Deborah Kelly Center for Outcomes Research, Department of Obstetrics and Gynecology, Massachusetts General Hospital, Boston, MA, USA

⁸ Division of Engineering in Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

Introduction

Optimization of single embryo transfer outcomes while reducing multiple gestation rates has remained a primary goal of the reproductive medicine field in recent years. Determinants of pregnancy success following an IVF cycle include age, diagnosis, comorbidities, endometrial receptivity, gamete, and embryo quality. One of the key actionable determinants is selection of a high-quality embryo for transfer [1]. Morphologic assessment, preimplantation genetic testing for aneuploidy (PGT-A), metabolomics, time lapse imaging, and morphokinetics have been investigated and employed as tools to improve outcome of single embryo transfers (SETs) with varying success [2].

Use of PGT-A for embryo screening results in a reduced miscarriage rate among women > 35 years old, increases the implantation rate per transfer, and leads to fewer multiple gestations [3]. However, transfer of a euploid embryo does not guarantee pregnancy. According to 2018 SART data, implantation rates range from 57 to 66% per retrieval and when utilizing PGT-A and a live birth rate (LBR) of 50–60% per retrieval when utilizing PGT-A. Excellent quality euploid embryos have been found in retrospective studies to have higher odds of implantation and ongoing pregnancy rates compared to poor quality blastocysts [4].

When an individual has multiple euploid embryos for transfer, the decision on which to transfer is left to initial morphologic assessment and grading by a highly trained embryologist. This method has been found to be highly subjective and variable [5, 6]. Deep learning, artificial intelligence (AI) algorithms have garnered interest for application in improving the consistency and outcome of embryo selection. When presented with still images of euploid embryos at 113 h post-insemination (Day 5 of embryo development), a deep learning AI algorithm has been demonstrated to identify the embryo most likely to implant in 75% of cases compared to 67% with morphological assessment by an embryologist [7].

While it is clear that AI can perform simple, repetitive tasks well and results from its application in embryo selection are promising, it is impractical and inappropriate to make a complete transition to solely AI-selected embryo transfer with our current level of evidence. There is significant value in the clinical experience of the embryologist informing the decision of which embryo to transfer. However, that selection may be improved by the addition of an AI algorithm to the existing embryologist knowledge.

The goal of this study was to evaluate the performance of highly trained embryologists in selecting top quality day 5 euploid blastocysts with and without the aid of a deep learning algorithm. We hypothesized that there would be no difference in rate of embryologist selection

of successfully implanted euploid embryos with the addition of AI algorithm results.

Materials and methods

This is a study of embryologist evaluation of previously graded and transferred Day 5 euploid blastocysts with known implantation outcomes. All of the images and clinical outcomes were collected retrospectively. Data were collected at the Massachusetts General Hospital (MGH) fertility center in Boston, Massachusetts.

Embryo image acquisition

We used recorded videos of embryos collected from 160 patients with informed consent for research and publication, under an institutional review board approval for secondary research use. Videos were collected for research after institutional review board approval by the Massachusetts General Hospital Institutional Review Board (IRB#2017P001339 and IRB#2019P002392). The videos were collected using a commercial time-lapse imaging system (EmbryoScope, Vitrolife). The imaging system used a Leica 20× objective that collected images at 10-min intervals under illumination from a single 635 nm LED. Video processing and still image acquisition has been described previously [7]. In brief, videos were broken into respective frames to extract still images, and only images obtained at 113 ± 0.05 h post insemination were used in this study.

For inclusion in the study, embryos were required to have a euploid result from a trophectoderm biopsy, using the same modified FAST-SeqS next-generation sequencing platform (Invitae, San Francisco, CA) and known implantation outcome. All blastocysts included in the study were fertilized with ICSI and had undergone trophectoderm biopsy and therefore the blastocysts had met biopsy criteria with grading of 3BC/3CB or greater using Gardner morphology criteria [8]. Images used in the study were taken prior to biopsy and vitrification at ~ 113 h post-insemination. Blastocysts were all vitrified and warmed using the same vitrification and thawing protocol (Irvine Scientific Vitrification and Warming, FujiFilm).

Implantation outcomes

All blastocysts included in this study were transferred individually and had a known implantation outcome. Successful implantation was defined as a gestational sac visualized on ultrasound following embryo transfer. There were no fresh embryo transfers given PGT-A evaluation of all embryos. Images were obtained from embryos between August 2014 and March 2018 at Massachusetts General Hospital Fertility

Center. Still images of 400 embryos resulting from 160 different patients were used to generate 200 unique successful/ unsuccessful implantation pairs of similar grade and quality blastocysts.

Artificial intelligence algorithm

The convolutional neural network (CNN) used in the artificial intelligence algorithm for embryo assessment has been previously described [7, 9, 10]. Briefly, the CNN(Xception architecture) was pre-trained with 1.4 million ImageNet images and transfer learned using 2440 Day 5 static human embryo images recorded at a single time point ~ 113 h post-insemination (hpi) which helped in identifying top-quality embryos capable of implantation. CNN was retrained and evaluated in identifying euploid embryos capable of implantation with known implantation outcomes and embryo images. The final classification layer was replaced with two classes making it a binary classification of “negative” for implantation and “positive” for implantation. The softmax probability values given by the network were used as the embryo’s implantation potential. It was shown to be effective in classifying successful and unsuccessful implantation outcomes in a prior test set of euploid embryos with an accuracy of 75.3% [7, 9, 10]. This AI network produces a probability score for implantation, and the embryo with the higher implantation probability was highlighted in the sets distributed to embryologists. Embryologists were blinded to the implantation probability score for each embryo.

Embryologist assessment

The de-identified embryo image pairs were distributed to 17 highly trained embryologists from five fertility centers across the USA. Embryologists were provided with the following instructions: “Each slide has 2 images of Day 5 Euploid Blastocysts with known implantation outcomes: one embryo that implanted and one embryo that did not implant. Please enter the embryo you would select as your top choice for transfer.” (Fig. 1a).

One year later, the same set of embryos was distributed to the embryologists, and their responses were compared to

the first distribution. Embryologists were not provided with the results for their embryo selections. Seven days following the second distribution, embryologists were provided with the same set of embryos in a different order, but with the image of the blastocyst with higher likelihood of implantation according to the AI algorithm highlighted (Fig. 1b). Embryologists were informed that the AI system was not perfect and had an accuracy of 75% in a prior test. Embryologists were again asked to indicate which embryo they would select as their top choice for transfer.

Embryologist experience level was defined as junior, intermediate, or senior according to the following criteria. Junior embryologists were proficient at embryo assessments and transfers; intermediate embryologists were proficient at cryopreservation and ICSI; senior embryologists were proficient at performing embryo biopsy.

Statistical analysis

Paired t-tests, chi-squared, and receiver operating curve analysis were performed in Stata/IC 16.1. *P* values of < 0.05 were considered statistically significant. Analysis was stratified by embryologist experience level as defined above.

Results

Table 1 includes a list of patient and cycle characteristics of the successful and failed implantation embryos used for this study. Maternal age at time of egg retrieval was similar, and the most common infertility diagnosis was male infertility among both the implantation and failed implantation embryos. There were no significant differences in cycle characteristics among the failed and implanted embryos.

Paired embryo sets were distributed to a total of 17 embryologists. Fifteen embryologists responded to the initial distribution, 14 embryologists responded to all three distributions. The 14 complete responses were included in the analysis. Among the embryologists, there were 3 junior level, 4 intermediate level, and 7 senior level embryologists who participated in this study.

Fig. 1 Embryo implantation selection modules (with and without the aid of an AI algorithm). The highlighted embryo indicates the with higher likelihood of implantation as assessed by AI

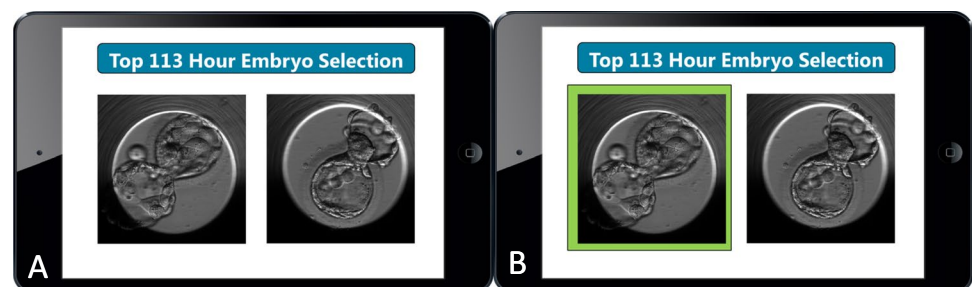


Table 1 Cycle characteristics for the 400 euploid embryo images used in the test set

Variable	Cycles of failed embryos N=80	Cycles of implanted embryos N=80
Oocyte yield mean (SD)	14.3 (4.7)	13.8 (4.5)
Number of M2	11.6 (4.0)	11.0 (4.0)
Number of 2PN	9.3 (3.8)	9.0 (3.6)
High-quality blasts	5.4 (1.8)	5.1 (2.4)
Age at egg retrieval	35.8 (3.3)	35.1 (4.7)
BMI	24.0(4.0)	23.2 (3.5)
Day 3 FSH	7.4 (1.8)	7.1 (1.9)
AMH	3.6 (2.8)	3.5 (2.6)
Infertility diagnosis (not exclusive)		
Male	22 (27.5%)	19 (23.8%)
Endo	0	1 (1.3%)
PCOS	8(10%)	6(7.5%)
DOR	10 (12.5%)	13 (16.3%)
Tubal	8 (10%)	6 (7.5%)
Uterine	2 (2.5%)	0
Ovulation Disorder	14 (7.5%)	17 (21.3%)
RPL	2 (2.5%)	5 (6.3)
Other	30 (37.5%)	39 (48.8%)
Unexplained	14 (17.5%)	10 (12.5%)

At the time of the first distribution of images, the 14 embryologists had a mean rate for selecting successfully implanted embryos of 64.7% (SD 2%). This was not significantly different from the performance of 14 embryologists 1 year later when the mean was 65.5% (SD 5.2%), $p=0.58$. Additionally, as shown in Fig. 2, AUC analysis revealed no change in accuracy with repeated evaluation of the 200 embryo sets by the embryologists between the first and second distributions ($p=0.89$). This demonstrated no significant change with repeated exposure to the embryo images. The results of the second distribution were used as the baseline assessment to compare post-AI results. Therefore, with standard morphological assessment without AI input, embryologists select the successfully implanted embryo in 65.5% of the presented image sets.

Presented with the series of embryo images, the AI algorithm selected the successfully implanted embryo in 78.5% of the embryo pairs. After providing embryologists with the highlighted image results from the AI algorithm predictions, the rate of selection of successfully implanted embryos increased from a mean of 65.5% (SD 5.2%) to 73.6% (SD 5%), $p<0.001$ (Table 2).

All 14 embryologists improved in their ability to select embryos most likely to implant after the addition of AI (Table 2). The mean percent improvement was 11.1% (range 1.4% to 15.5%). There was a trend toward greater

improvement in embryo selection in the junior level embryologists with a mean improvement of 14.1% (SD 5.8%), compared to 12.6% (SD 1.1%) in the intermediate level and 8.7% (SD 5.9%) in senior level embryologists. However, there was not a statistically significant difference globally ($p=0.22$) or between junior and senior level embryologists ($p=0.27$).

Accuracy of embryo selection was examined among the selections pre- and post-introduction of AI results. Figure 3 reflects the AUC (95% CI) for the most commonly selected embryo in each embryo pair, or the consensus selection, before and after AI. The accuracy of successfully implanted embryo selection improved after the addition of AI from 0.69 (95% CI: 0.62–0.75) to 0.76 (95% CI: 0.70–0.82), ($p=0.016$).

The embryologists agreed on which blastocyst to transfer in a majority of the image pairs. To estimate consensus among the embryologists who were all evaluating the embryos independently, we tabulated the number of times at least 75% (11 or more) of the embryologists chose the same embryo to transfer (Fig. 4). This occurred in 124/200 or 62% of the embryo pairs prior to sharing the AI results. With the addition of AI, the proportion of consensus episodes increased to 74% (147/200) which was a significant difference, $p=0.01$.

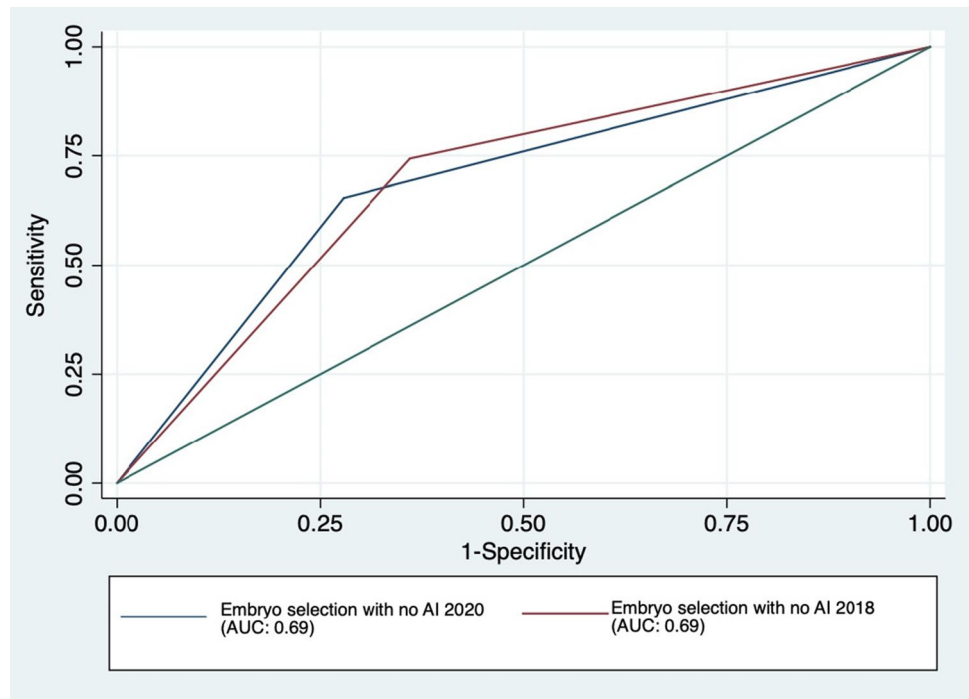
We then examined how AI performed compared to these consensus selections among embryologists. In the first distribution with 124 episodes of ≥ 11 embryologists selecting the same embryo, embryologist consensus was correct in 95/124 (77%) pairs and AI was correct in 97/124 (78%), $p=0.9$. After revealing the AI results, there was an increase in the number of consensus episodes among embryologists to 147 and the rate of implanted embryo selection among the consensus group was 86% (126/147) compared to an AI rate of 84% in those 147 embryo sets ($p=0.63$).

As AI is not 100% accurate in its prediction of successful blastocyst implantation, we also calculated the total number of times that embryologists changed from selecting an implanted embryo to a non-implanted embryo after the addition of AI. Among all 14 embryologists, each evaluating 200 embryo pairs there were 89/2800 instances of selecting a failed embryo after having selected a successful one on the first evaluation without AI assistance for an overall rate of 0.03%. Between the 2018 and 2020 assessments without AI, there was a rate of change from correct to incorrect of 0.04% (127/2800).

Discussion

In summary, this pre-post study demonstrates that there is significant improvement in embryologist rate of selection of high-quality euploid blastocysts destined to implant with the

Fig. 2 Receiver operating curve comparing the result of consensus embryo selection among embryologists without the use of AI after repeated exposure to the same embryo set



	AUC (95% CI)	P-value comparing ROC	Percent correct	P-value comparing paired t-test for percent correct
<i>Embryo selection with no AI 2020</i>	0.6876 (0.621-0.754)	0.8902	69.5%	0.6321
<i>Embryo selection with no AI 2018</i>	0.6915 (0.627-0.756)		68%	

Generated using Stata

Table 2 Percent improvement with AI selection by embryologist experience level

	Percent improvement (mean, SD)	Global <i>p</i> value	Junior vs Senior <i>p</i> value
Junior (<i>n</i> = 3)	14.1% (5.8%)	0.27	0.22 (no post-hoc adjustment)
Intermediate (<i>n</i> = 4)	12.6% (1.1%)		
Senior (<i>n</i> = 7)	8.7% (5.9%)		

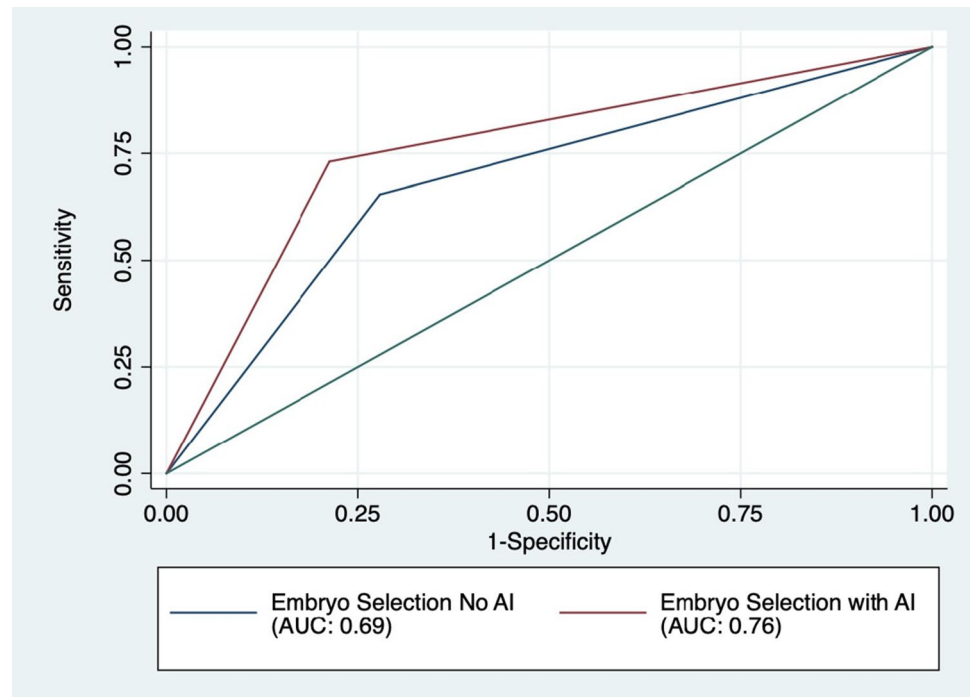
addition of our AI algorithm. Using archived still images of PGT-A euploid blastocysts with known successful or unsuccessful implantation following frozen euploid single embryo transfer, we have demonstrated that embryologists select the successfully implanted embryo in an average of 65.5% of embryo pairs. This average rate increased to 73.6% with the addition of AI.

There was no significant difference in rate of improvement when analyzed by embryologist level of experience, though there was a notable trend toward greater improvement in those with less experience. This suggests that the addition of AI may offer greater assistance to embryologists with less experience and improve selection of HQB for transfer overall. Additionally, there were no situations

when the rate of successful selection was reduced after providing the AI algorithm results.

The AI algorithm outperformed embryologists on both assessments, before and after they were informed of algorithm results. This agrees with published literature demonstrating a different algorithm outperforming embryologists for prediction of embryo euploidy [11]. Several publications have proposed using AI for embryo selection as a stand alone approach, not in combination with embryologist assessment [7, 12]. While that may be the ultimate method of embryo selection, the field is not currently poised to do this. There are barriers including infrastructure of individual clinics, validation, and generalizability of AI models [13, 14].

Fig. 3 Receiver operating curve comparing the result of consensus embryo selection among embryologists with and without the use of AI



	AUC (95% CI)	P-value comparing ROC	Percent correct	P-value comparing paired t-test for percent correct
<i>Embryo selection with no AI</i>	0.6876 (0.621-0.754)	0.0162	69.5%	0.0193
<i>Embryo selection with AI</i>	0.7588 (0.697-0.820)		76.5%	

Generated using Stata

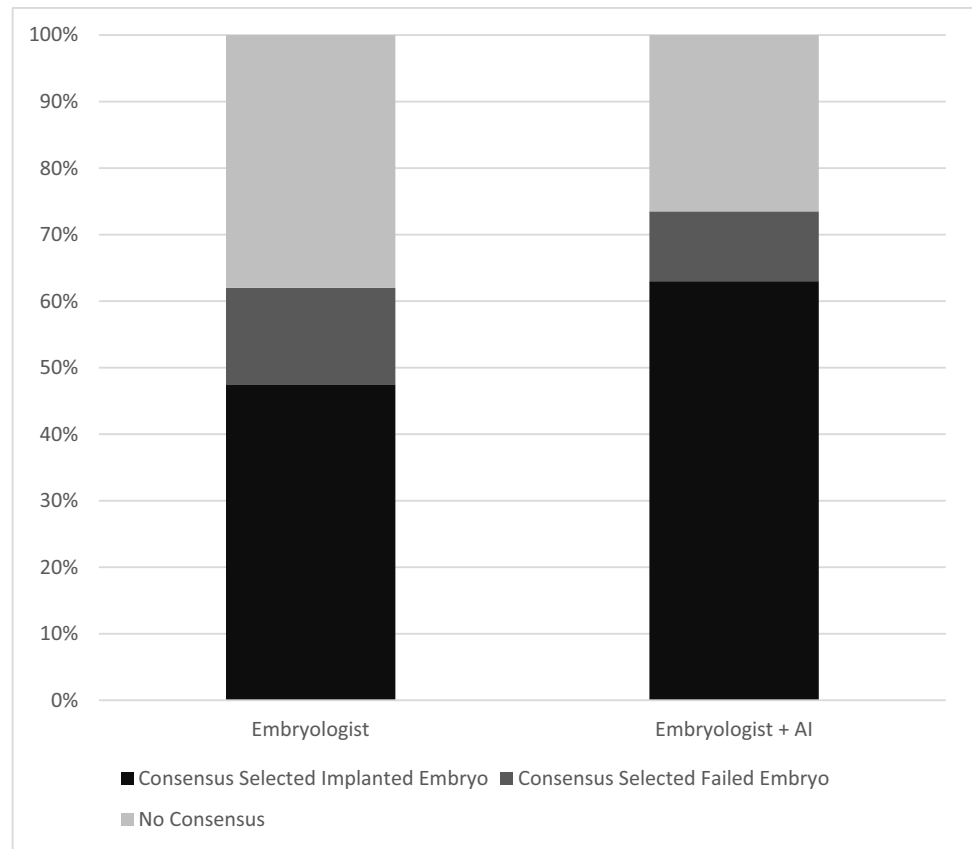
Interestingly, when limited to selections with greater than 75% consensus among embryologists, the AI algorithm and the consensus decisions had a similar rate of success in selecting the implanted embryo (84% and 86% respectively). This suggests that adding an AI algorithm as a tool in embryo selection would be similar to increasing the number of embryologists evaluating embryo morphology at one time. Importantly, there were selections that changed from an implanted embryo to a failed embryo after the addition of AI. However, the rate was very low at 0.03% of selections, and it was similar to the rate of change observed between the first and second assessments before the introduction of AI, 0.04%.

Prior analysis has demonstrated that there is no improvement in live birth or clinical pregnancy rates with morphological assessment of the already homogenous group of euploid PGT-A tested embryos [15]. However, morphological assessment in that study was performed by individual embryologists and not with the assistance of artificial intelligence. Our data suggests that improved embryo selection with the use of AI may make a difference in these rates if studied prospectively. Morphological assessment of

blastocysts has been shown to be highly variable and subjective despite rigorous training of embryologists [5, 16]. Our finding of increased rate of consensus among embryologists following addition of AI suggests that the addition of AI in the embryology lab may reduce inconsistency in embryo grading and provide additional value for morphological assessment of euploid embryos to improve the outcomes of single embryo transfer.

Strengths of this study include the inclusion of only euploid embryos, allowing for elimination of confounding factors that can impact transfer outcomes. Additionally, the assessment by 14 different embryologists located at multiple private and academic fertility centers means that our results are generalizable and not limited to the experience at a single institution or laboratory. Our AI algorithm was developed to select embryos based on morphologic characteristics at Day 5 of development, a timepoint that was chosen in order to challenge the system by only comparing high quality blastocysts at the same time of development. By using this group to test the AI algorithm and embryologists, we are able to control for impacts of timing of development, embryo stage, and uterine environment in regard to

Fig. 4 Consensus Selections: Proportion of 200 selections when > 75% of embryologists selected the same embryo for transfer. The proportion of consensus decisions increased after the addition of AI



hormonal preparation. We chose to use an image from a single point in time on Day 5 rather than time lapse images as the vast majority of embryology labs are using serial, static time point assessment for embryo morphology grading and ultimately selection, not time lapse technology; therefore it is more generalizable to current practice to use a single time point. Additionally, the algorithm was developed using still images from this same time point. Applying a similar methodology but using time lapse images would be an interesting future study.

There are limitations of this study including its retrospective nature. The images were obtained and paired based on known euploid status; therefore all blastocysts were subject to a trophectoderm biopsy procedure which may itself impact the implantation competency of an embryo [17]. There is also an assumption that all uteri and endometrium are equal in regard to providing an environment for successful implantation. This is obviously not the case, and we attempted to account for this in demonstrating no significant difference between the patient and cycle characteristics or infertility diagnoses of the cycles from which the embryos were generated. However, the best way to validate the addition of AI to embryo selection and account for confounding factors would be a prospective randomized cohort study. Additionally, AI algorithms in the future may incorporate patient level factors into predictions for likelihood

of implantation. Another limitation is the sample size of 14 embryologists participating fully in the study. A greater number of participants would have strengthened the study and will be a goal in future AI assessments. The 14 embryologists are from 5 different centers across the country which increases the generalizability of the study.

Embryos that are euploid on PGT-A testing and frozen to await future transfer represent a unique opportunity to apply AI technology to embryo selection in a prospective fashion. Given previous findings that morphologic assessment offers no significant benefit in the decision on which euploid embryo to transfer, randomizing an individual to standard embryologist selection versus selection by or with the assistance of AI algorithm based on pre-existing still images would allow for prospective assessment of this technology and confirm benefit while presenting no reduced chance at pregnancy for the patient. An aneuploid diagnosis on PGT-A has been demonstrated to be highly predictive of failed pregnancy [18]; however among euploid embryos, there is not a clearly effective method of selecting one that is of greater likelihood of implantation than another. Application of our AI algorithm for embryo selection in this scenario may improve implantation and live birth rates.

In conclusion, we have shown that the addition of a deep learning AI algorithm to standard embryologist evaluation of euploid day 5 blastocysts results in improved rates of

selection of successfully implanted embryos when compared to embryologist evaluation alone. The AI algorithm also increases consensus among embryologists, which could serve to reduce variability and improve outcomes of single euploid embryo transfer. This is clearly demonstrated in the highest prognosis transfer cycles used in this study. Future directions include validation in a prospective study and assessment of the decision making time with and without AI, application in untested blastocysts and those at other time points in development, as well as use of AI for embryo selection as an adjunct, as demonstrated here, or independently from embryologist assessment.

Authors' contributions All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by CB, KJ, and VWF. The first draft of the manuscript was written by VWF, and all authors commented on previous version of the manuscript. All authors read and approved the final manuscript.

Funding This work was partially supported by the Brigham Precision Medicine Developmental Award (Brigham Precision Medicine Program, Brigham and Women's Hospital), Partners Innovation Discovery Grant (Partners Healthcare), and R01AI118502, and R01AI138800.

Declarations

Ethics approval Approval for this study was obtained from the ethics committee of Massachusetts General Hospital (IRB #2017P001339 and #2019P002392).

Conflict of interest CB, HS, MK, and PT all have a patent WO2019068073A1 pending in relation to this work. Other authors declare no competing interests.

References

1. Strauss JF, Barbieri RL, editors. *Yen & Jaffe's reproductive endocrinology: physiology, pathophysiology, and clinical management*. 8th ed. Philadelphia, PA: Elsevier; 2019.
2. Bromer JG, Seli E. Assessment of embryo viability in assisted reproductive technology: shortcomings of current approaches and the emerging role of metabolomics. *Curr Opin Obstet Gynecol*. 2008;20(3):234–41. <https://doi.org/10.1097/GCO.0b013e3282fe723d>.
3. Munné S, et al. Preimplantation genetic testing for aneuploidy versus morphology as selection criteria for single frozen-thawed embryo transfer in good-prognosis patients: a multicenter randomized clinical trial. *Fertil Steril*. 2019;112(6):1071–1079.e7. <https://doi.org/10.1016/j.fertnstert.2019.07.1346>.
4. Irani M, et al. Morphologic grading of euploid blastocysts influences implantation and ongoing pregnancy rates. *Fertil Steril*. 2017;107(3):664–70. <https://doi.org/10.1016/j.fertnstert.2016.11.012>.
5. Baxter Bendus AE, Mayer JF, Shipley SK, Catherino WH. Inter-observer and intraobserver variation in day 3 embryo grading. *Fertil Steril*. 2006;86(6):1608–15. <https://doi.org/10.1016/j.fertnstert.2006.05.037>.
6. Bormann CL, et al. Consistency and objectivity of automated embryo assessments using deep neural networks. *Fertil Steril*. 2020;113(4):781–787.e1. <https://doi.org/10.1016/j.fertnstert.2019.12.004>.
7. Bormann CL, et al. Performance of a deep learning based neural network in the selection of human blastocysts for implantation. *eLife*. 2020;9:e55301. <https://doi.org/10.7554/eLife.55301>.
8. Gardner DK, Surrey E, Minjarez D, Leitz A, Stevens J, Schoolcraft WB. Single blastocyst transfer: a prospective randomized trial. *Fertil Steril*. 2004;81(3):551–5. <https://doi.org/10.1016/j.fertnstert.2003.07.023>.
9. Kanakasabapathy M, et al. An inexpensive, automated artificial intelligence (AI) system for human embryo morphology evaluation and transfer selection. *Fertil Steril*. 2019;111(4): e11. <https://doi.org/10.1016/j.fertnstert.2019.02.047>.
10. Thirumalaraju P, et al. Deep learning-enabled blastocyst prediction system for cleavage stage embryo selection. *Fertil Steril*. 2019;111(4): e29. <https://doi.org/10.1016/j.fertnstert.2019.02.077>.
11. Chavez-Badiola A, Flores-Saiffe-Farias A, Mendizabal-Ruiz G, Drakeley AJ, Cohen J. Embryo Ranking Intelligent Classification Algorithm (ERICA): artificial intelligence clinical assistant predicting embryo ploidy and implantation. *Reprod Biomed Online*. 2020;41(4):585–93. <https://doi.org/10.1016/j.rbmo.2020.07.003>.
12. Chen T-J, Zheng W-L, Liu C-H, Huang I, Lai H-H, Liu M. Using Deep Learning with Large Dataset of Microscope Images to Develop an Automated Embryo Grading System. *Fertil Reprod*. 2019;01(01):51–6. <https://doi.org/10.1142/S2661318219500051>.
13. Curchoe CL. “The paper chase and the big data arms race.”. *J Assist Reprod Genet*. 2021. p. s10815–021–02122–3. <https://doi.org/10.1007/s10815-021-02122-3>.
14. Curchoe CL, Flores-Saiffe-Farias A, Mendizabal-Ruiz G, Chavez-Badiola A. Evaluating predictive models in reproductive medicine. *Fertil Steril*. 2020;114(5):921–6. <https://doi.org/10.1016/j.fertnstert.2020.09.159>.
15. Shear MA, et al. Blasts from the past: is morphology useful in PGT-A tested and untested frozen embryo transfers? *Reprod Biomed Online*. 2020;41(6):981–9. <https://doi.org/10.1016/j.rbmo.2020.07.014>.
16. Storr A, Venetis CA, Cooke S, Kilani S, Ledger W. Inter-observer and intra-observer agreement between embryologists during selection of a single Day 5 embryo for transfer: a multicenter study. *Hum Reprod*. 2017;32(2):307–14. <https://doi.org/10.1093/humrep/dew330>.
17. Paulson RJ. Preimplantation genetic screening: what is the clinical efficiency? *Fertil Steril*. 2017;108(2):228–30. <https://doi.org/10.1016/j.fertnstert.2017.06.023>.
18. Tiegs AW, et al. A multicenter, prospective, blinded, nonselection study evaluating the predictive value of an aneuploid diagnosis using a targeted next-generation sequencing–based preimplantation genetic testing for aneuploidy assay and impact of biopsy. *Fertil Steril*. 2021;115(3):627–37. <https://doi.org/10.1016/j.fertnstert.2020.07.052>.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.