



OPEN MAIA platform for routine clinical testing: an artificial intelligence embryo selection tool developed to assist embryologists

Mariana Nicolielo¹, Catherine Kuhn Jacobs¹, Bruna Lourenço¹, Murilo Costa Maffei², Dóris Spinosa Chéles³, Matheus Búbola Duarte², Bruno Araújo Mendes², Vinícius Casado Moraes², Maurício Barbour Chehin^{3,4}, José Roberto Alegretti¹, Eduardo Leme Alves da Motta^{4,5}, Aline Rodrigues Lorenzon^{3,7}, Marcelo Fábio Gouveia Nogueira^{6,7} & José Celso Rocha^{2,7}✉

The need to reduce the number of embryos transferred in assisted reproductive care to prevent multiple gestations has led to a stronger emphasis on selecting embryos with the highest morphological quality. Although this evaluation has traditionally been performed by trained embryologists, the increasing use of time-lapse incubators has introduced a greater volume of data and subjectivity in decision-making. Artificial intelligence (AI)-based tools can support embryologists by offering objective, standardized embryo assessments. In Brazil, like other countries, where imported embryo selection technologies may not account for local demographic and ethnic profiles, an AI model — Morphological Artificial Intelligence Assistance (MAIA) — was developed through a collaboration between a university and a private fertility clinic in São Paulo. The model was trained using 1,015 embryo images and prospectively tested in a clinical setting on 200 single embryo transfers. In clinical testing, MAIA achieved an overall accuracy of 66.5%. In elective embryo transfers, where there were more than one embryo eligible for transfer, MAIA achieved 70.1% accuracy for predicting clinical pregnancy. Designed with a user-friendly interface tailored by embryologists, MAIA provides real-time embryo evaluations to support decision-making in routine care.

Over the last few decades, in vitro fertilization (IVF) has revolutionized reproductive therapy for infertile patients and has provided several approaches to achieve a successful pregnancy^{1,2}. Previously, this practice was based on the transfer of more than one embryo, sometimes leading to multiple pregnancy and, consequently, resulting in higher maternal and neonatal risks than those in singleton pregnancies³. With the technological development of IVF laboratories, single embryo transfer (SET) is now recommended to achieve a healthy pregnancy, but maintaining success rates is challenging^{1,4}. The selection of the embryo with the highest probability of implantation still relies largely on the subjective judgment of embryologists⁵.

Historically, the selection of embryos has been based on the evaluation of their morphology, and for embryos in the blastocyst stage, three parameters, namely, the degree of embryo expansion, the homogeneity of the inner cell mass (ICM), and the trophoctoderm (TE), are usually evaluated. The Gardner classification^{5–7} is one of the most widely used criteria in clinical practice. Although well established and used worldwide, these parameters are not sufficiently precise to accurately predict potential for success, as they depend on the visual observation of the embryologist. Therefore, subjectivity and consequent divergences (inter- and intra-embryologists variations) are intrinsic factors of this evaluation system^{5,8}.

¹Embryology Department, Huntington Reproductive Medicine-Eugin Group, São Paulo, SP, Brazil. ²Laboratory of Applied Mathematics, Department of Biological Sciences, São Paulo State University (UNESP), Assis, SP, Brazil.

³Research and Development Department, Huntington Reproductive Medicine-Eugin Group, São Paulo, SP, Brazil.

⁴Clinical Department, Huntington Reproductive Medicine-Eugin Group, São Paulo, SP, Brazil. ⁵Department of Gynecology, School of Medicine, Federal University of São Paulo, São Paulo, SP, Brazil. ⁶Laboratory of Embryonic Micromanipulation, Department of Biological Sciences, São Paulo State University (UNESP), Assis, SP, Brazil.

⁷These authors contributed equally: Aline Rodrigues Lorenzon, Marcelo Fábio Gouveia Nogueira and José Celso Rocha. ✉email: jose.celso@unesp.br

The introduction of the time lapse system (TLS), with the acquisition of multiple images at different times of embryonic growth, quickly became popular because of its non-invasive nature and maintenance of ideal culture conditions⁹. The TLS allows the monitoring of embryonic development, from the zygote stage to full blastocyst expansion, without the need to remove the culture dish from its ideal conditions for morphological evaluation. The images obtained from embryos in various planes and at regular intervals can be used in digital processing programs to determine quality, thus enhancing the prediction of clinical pregnancy^{7,10,11}.

Initially, some studies extracted images obtained with the TLS and analysed them with mathematical variables representative of known morphological attributes of the blastocyst^{12–15}. Chéles et al.¹⁶ developed an automated processing protocol based on images from TLS in EmbryoScope® (Vitrolife, Sweden) and Geri® (Genea Biomedx, Australia) incubators, which generated 33 variables categorized as texture, mean grey level, grey level standard deviation, modal value, ICM area and diameter, ET thickness, and light level, among others. These variables have the potential to be used as inputs for the development of an artificial intelligence (AI) program to predict the viability of the embryo produced by IVF^{17,18}.

Artificial intelligence (AI) aims to replicate human cognitive processes in order to address complex problem-solving tasks. It comprises a broad spectrum of computational approaches, including multilayer perceptron artificial neural networks (MLP ANNs), genetic algorithms (GAs), deep learning (DL), convolutional neural networks (CNNs), fuzzy logic systems, and machine learning (ML) techniques¹⁹.

In studies related to IVF, AI has wide application potential for sperm selection²⁰ oocyte selection²¹ embryo classification^{12,22,23} pregnancy prediction²⁴ live birth^{13,14,23} and embryonic ploidy²⁵.

Although these examples are based on different methods, they demonstrate the potential of AI methods to automate the embryonic evaluation while considering the complexity of the embryological variables²⁶ and as a complementary system to the current performance of manual and subjective selection²⁷.

Although there is still some resistance to full confidence in AI programs²⁸, there are currently commercially available AI-based software programs on the market that play important roles in the clinical analysis of embryos in a retrospective/prospective manner²⁹. Examples of this are those developed for application in reproductive medicine, such as the iDAScore (Vitrolife) implemented in the EmbryoScope® incubator³⁰, the AI Chloe of the Fairtility group³¹ and the AI EMA of the AIVF group³² among other existing software. Thus, there is a clear technological demand in assisted reproduction for an automated and objective methodology for the embryonic evaluation aiming, ultimately, to increase the probability of healthy live births and to decrease the complications of multiple pregnancy in couples undergoing treatment by assisted reproduction technology (ART)³³.

Designing an AI model specifically for a Brazilian population enables results that more accurately reflect the country's distinctive genetic diversity, with the largest population in Latin America³⁴. Disparities in health outcomes across ethnic groups are particularly evident in reproductive and endocrine health, with infertility showing some of the most marked differences. For instance, ovarian reserve has been observed to differ among populations; women of latina and chinese descent between the ages of 40 and 45 may have lower antimüllerian hormone (AMH) levels than African american women. In this regard, a steepest rate of decline was observed among Chinese women (10.5%), whereas African American women showed the lowest decrease (6.3%)³⁵. Moreover, data from the Society for Assisted Reproductive Technology (SART) database indicate that clinical pregnancy and live birth rates are lower among Black, Asian, and Hispanic women compared to White women³⁶. These considerations highlight the critical importance of accounting for population diversity in the development of a new reproductive technology.

The objective of this study was to use AI methods, particularly MLP ANNs associated with GAs, to predict gestational success (presence of a gestational sac and foetal heartbeat) from morphological variables automatically extracted from digital processing of images of blastocysts produced by IVF. Furthermore, the establishment of a technological domain in this theme – an innovation in Brazil and South America – represents the acquisition of competence with the differential AI training with a customized image bank, i.e., specific to the patients treated by the collaborating clinic in São Paulo, Brazil, who have specific demographic and ethnic characteristics typical of the country^{37,38}. As a secondary objective, multicentre clinical tests were conducted as part of a prospective observational study (i.e., the model evaluation). This study aimed to evaluate the performance of the MAIA algorithm in a real clinical setting, using embryo images and associated clinical outcomes obtained in routine care.

Results

Single and mode accuracy of the MLP ANNs

MAIA (an acronym for Morphological Artificial Intelligence Assistance) was developed based on the five best-performing multilayer perceptron artificial neural networks. During the learning process, the models were trained and validated using a dataset of morphological images, with the aim of predicting clinical pregnancy (CP) outcomes. The data were divided into two distinct subsets: training and validation, as detailed in Table 1.

To further assess the models' generalization capability, internal validation was performed. In this evaluation, the MLP ANNs exhibited more consistent performance, achieving accuracies of 60.6% or higher (Table 2).

The analysis of the area under the curve (AUC) and the receiver operating characteristic (ROC) curves for both the training and for internal validation results are presented in Table 3; Fig. 1a,b, respectively.

When all the results presented by the MLP ANNs were normalized (Supplementary Text 1, Supplementary Tables 1–8) and the mode between the ANNs was applied, the MAIA software result was 77.5% correct in the prediction of clinical pregnancy positive and 75.5% correct for the prediction of clinical pregnancy negative. The confusion matrix³⁹ is presented in Fig. 2. The AUC for this case is presented in Table 4 and ROC curves in Fig. 3.

A graphical user interface was developed for MAIA and designed to facilitate its use in the daily routine of assisted reproduction clinics. This interface is shown in Supplementary Text 2, Supplementary Fig. 1a-1b,

MLP ANN	Training		Validation	
	CP +	CP –	CP +	CP –
1	85.5	87.5	55.5	64.4
2	92.5	93.4	54.5	55.9
3	84.7	79.0	60.0	61.0
4	93.7	90.4	59.1	60.2
5	87.1	87.1	52.7	61.0

Table 1. Results of the training and validation data of the 5 best MLP ANNs obtained by gas. CP + clinical pregnancy positive, CP – clinical pregnancy negative.

MLP ANN	Internal validation	
	CP +	CP –
1	63.7	62.8
2	67.5	61.7
3	63.7	60.6
4	61.3	62.8
5	62.5	61.7

Table 2. Performance of MLP ANNs in the internal validation. CP + clinical pregnancy positive, CP – clinical pregnancy negative.

MLP ANN	Training		internal validation	
	AUC CP +	AUC CP –	AUC CP +	AUC CP –
1	0.82	0.85	0.56	0.54
2	0.96	0.95	0.57	0.59
3	0.89	0.90	0.53	0.59
4	0.90	0.92	0.58	0.65
5	0.94	0.96	0.56	0.66

Table 3. AUC values for training results and internal validation for responses to whether there is a clinical pregnancy (CP +) or no clinical pregnancy (CP –). CP + clinical pregnancy positive, CP – clinical pregnancy negative.

Supplementary Fig. 2–3 and Supplementary Text 3. A video demonstrating the entire operation of the graphical interface is presented in Supplementary Video 1.

MAIA performance in evaluation of the prediction model

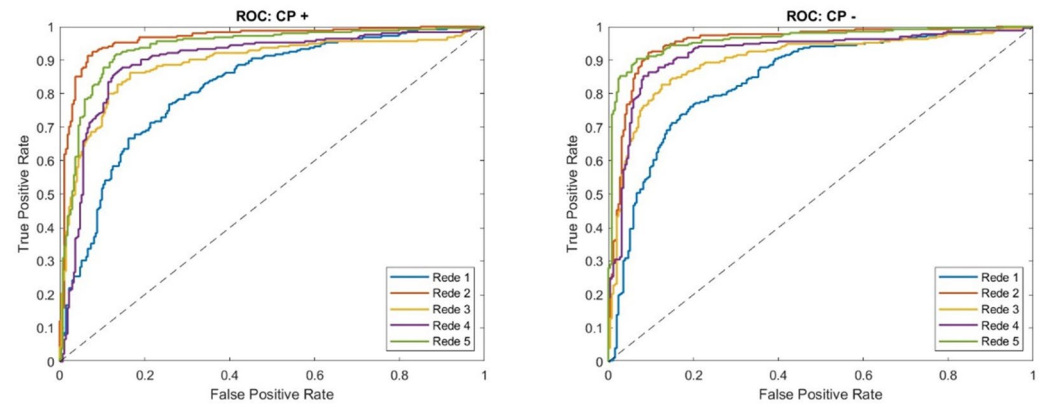
In model evaluation (i.e., tests performed under a real multicentre clinical routine), using the MAIA, version 4.0, graphical interface, the clinical pregnancy rate of all patients who underwent embryo transfer was 53% ($n = 106/200$). In centres A, B and C, the clinical pregnancy rates were 51.8%, 61.3% and 40.9%, respectively. Detailed patient and cycle characteristics are summarized in Table 5.

MAIA scores (Supplementary Fig. 1b) between 0.1 and 5.9 were considered negative predictors of clinical pregnancy, and scores between 6.0 and 10.0 were considered positive predictors (Supplementary Text 4 and Supplementary Fig. 4). The AUC, considering all cases, was 0.65. In general, the accuracy of MAIA between positive and negative clinical pregnancies was 66.5%. Among the centres that independently evaluated the embryos ($n = 200$) with the aid of MAIA, centres A, B and C obtained correct rates of 67.9% (40.5% participating, $n = 81$, of the total number of analyses), 69.3% (37.5%, $n = 75$, of the total cases) and 59.1% (22.0%, $n = 44$, of the total cases), respectively. Linear regression analysis showed that MAIA's predictions (both correct and incorrect) were strongly correlated with clinical pregnancy outcomes (CP + and CP –), with R values ranging from 0.65 to 1.0 ($P < 0.001$). In contrast, embryologists' selections across the three centres yielded R values ranging from 0.053 to 0.685, with P values varying from non-significant ($P = 0.792$) to statistically significant ($P = 0.001$) (Table 6). The graphs representing the comparison of R and P values between centres A, B and C and MAIA are presented in Supplementary text 5 and Supplementary Fig. 5–8.

Among the patients, 93 had only one embryo to be transferred (nonselective cases), and considering only those cases, the AUC was 0.65, with an accuracy of 62.4%.

For elective cases, where 107 patients had more than one embryo eligible for transfer (total of 284 embryos), of the 107 transferred embryos, the AUC was 0.60, with an accuracy of 70.1% of the clinical pregnancy

a)



b)

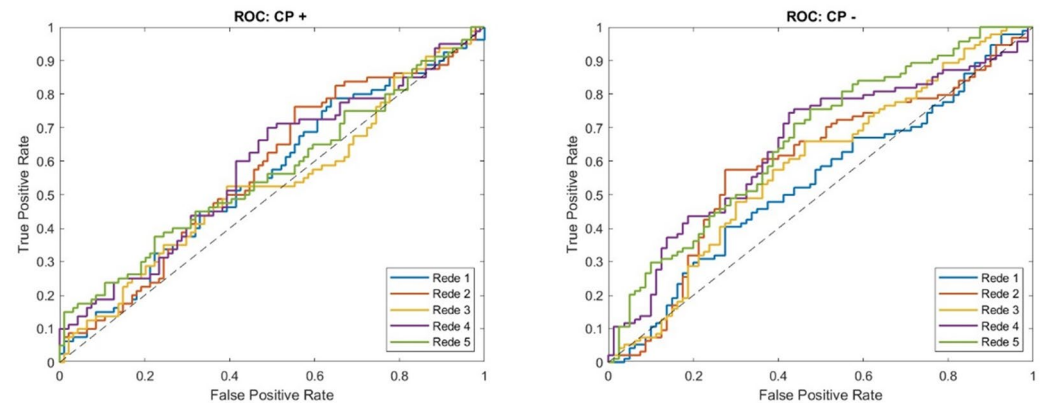


Fig. 1. (a) ROC curves for the training results for the positive and negative clinical pregnancy cases; (b) ROC curves for the internal validation results for clinical pregnancy positive and negative cases.

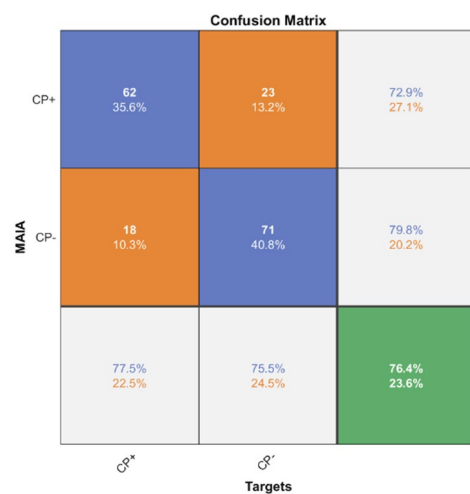
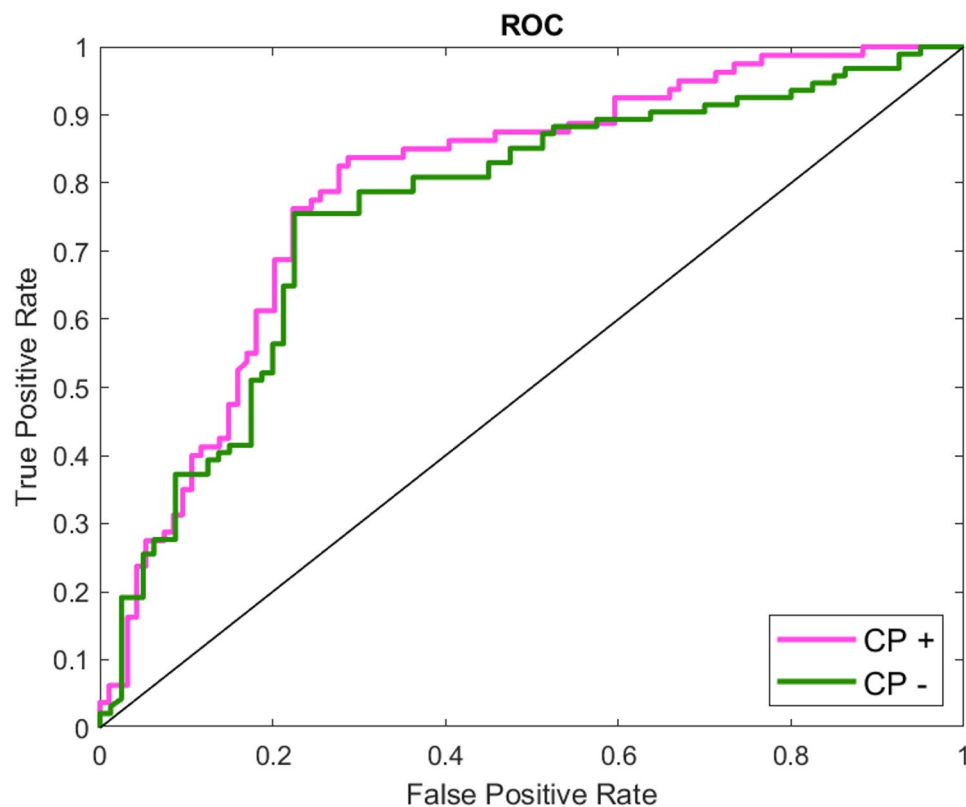


Fig. 2. Confusion matrix.

AUC for the combined data	
CP +	CP -
0.74	0.77

Table 4. AUC after normalization and mode application.**Fig. 3.** ROC curves for the normalized data and after applying the mode for the 5 MLP ANNs for the internal validation.

result. In cases in which the MAIA was responsible for the choice of embryo, the embryologist would choose another embryo ($n=44$), the accuracy was 81.8%, and the clinical pregnancy rate was 75%. In cases in which the embryologist was responsible for the choice of MAIA and disagreed with the choice of MAIA ($n=38$), the clinical pregnancy rate was 47.4%, and the success rate of MAIA was 60.5% in cases of single embryos selected for transfer. In the remaining elective cases ($n=25$), the choices of the embryologists and MAIA were in agreement, and the clinical pregnancy and MAIA accuracy rates were 64.0%.

Most of the analysed embryos reached the blastocyst stage on the 5th day of development ($n=163/200$, 81.5%), and MAIA correctly predicted 68.1% of the embryos. For the embryos that were analysed on the 6th day ($n=36/200$, 18%), the accuracy of MAIA was of 61.1%. There was only one embryo, on Day 4, in which the MAIA prediction was negative and the embryo resulted in a clinical pregnancy. The combined AUC for Days 5 and 6 was 0.64.

For euploid embryos ($n=122/200$, 61% of the transferred embryos), the accuracy of MAIA was 69.7%, with an AUC of 0.67, whereas for non biopsied embryos ($n=78/200$, 39%), the accuracy was 61.5%, and the AUC was 0.62.

Discussion

In summary, in this work, we propose an AI-based platform (named MAIA) to aid embryologists in their clinical routine for embryo assessment. MAIA was developed entirely by a Brazilian partnership (university-private clinic) and involved a database from three reference centres in the city of São Paulo that provide assistance to patients from all over the country.

Our results indicate that the use of morphological parameters from time-lapse images of embryos^{7,11} strongly correlates with embryonic morphology and clinical pregnancy. In addition, the time-lapse system provides a predictive improvement in terms of morphological and morphokinetic information compared with standard

	Clinic A	Clinic B	Clinic C	Overall
# Patients	81 (40.5%)	75 (37.5%)	44 (22.0%)	200
# IVF cycles/SET	81 (40.5%)	75 (37.5%)	44 (22.0%)	200
Mean age (years)	36.9 ± 3.7	40.0 ± 5.4	38.0 ± 3.9	38.3 ± 4.7
Mean BMI	23.1 ± 3.4	24.0 ± 4.6	23.9 ± 3.9	23.5 ± 3.9
CP rate (%)	51.8% (42/81)	61.3% (46/75)	40.9% (18/44)	53.0% (106/200)
Autologous cycles (%)	88.9% (72/81)	52.0% (39/75)	81.8% (36/44)	73.5% (147/200)
Mean age	36.2 ± 3.1	36.6 ± 3.1	36.7 ± 2.9	36.4 ± 3.0
BMI	22.7 ± 3.2	24.0 ± 5.4	23.6 ± 3.5	23.1 ± 3.8
PGT-A (%)	67.9% (55/81)	61.5% (24/39)	72.2% (26/36)	71.4% (105/147)
Frozen embryo transfer (%)	95.1% (77/81)	89.7% (35/39)	94.4% (34/36)	94.6% (139/147)
CP rate (%)	51.8% (42/81)	66.7% (26/39)	44.4% (16/36)	55.8% (82/147)
Donor egg cycles (%)	11.1% (9/81)	48.0% (36/75)	18.2% (8/44)	26.5% (53/200)
Mean age	42.0 ± 4.5	43.8 ± 4.9	43.8 ± 2.8	43.5 ± 4.5
BMI	26.0 ± 3.9	24.0 ± 3.1	25.1 ± 5.9	24.9 ± 3.7
PGT-A (%)	11.1% (1/9)	13.9% (5/36)	50.0% (4/8)	18.9% (10/53)
Frozen Embryo transfer (%)	77.8% (7/9)	38.9% (14/36)	100.0% (8/8)	54.7% (29/53)
CP rate (%)	22.2% (2/9)	55.5% (30/36)	25.0% (2/8)	45.3% (24/53)

Table 5. Patient and cycle characteristics used for MAIA clinical model evaluation.

	Centre A	MAIA	Centre B	MAIA	Centre C	MAIA	All centres	MAIA
R	0.0533	0.65	0.368	0.702	0.685	Perfect fit	0.325	0.747
P	0.792	< 0.001	0.146	< 0.001	0.001	/	0.009	< 0.001

Table 6. Comparison of *R* and *P* values between centres A, B and C (performed by embryologists) and MAIA. *R* linear regression coefficient, *P* statistical significance.

incubators^{40–42}. In the present study, the standardization of image quality (pixels, lighting, etc.) obtained at different centres using the same time-lapse system⁴³ enabled the use of a multicentre approach. Corroborating the aforementioned studies, the present study used static images of blastocysts obtained from time-lapse technology for the digital processing described by Chéles et al.¹⁶ to obtain mathematical variables predictive of the morphological quality of the blastocyst. The results of the software developed in the internal validation showed predictions of 77.5% (CP +) and 75.5% (CP -), which indicate high potential for continuing its daily use in assisted reproductive care clinics. This fact is supported by the prospective analysis performed in three assisted reproductive care clinics, where MAIA was able to achieve an accuracy for clinical pregnancy of 70.1% in elective embryo transfers and 62.4% in nonelective embryo transfers.

As in the present study, the quantitative analysis of embryo characteristics on the basis of digital images has already been tested^{44,45} and has become an active area of research for application in AI. The acquisition of variables representative of the embryo by image processing was also used by Chavez-Badiola et al.⁴⁶ who developed an algorithm to predict biochemical pregnancy (β-hCG positive). These authors used machine learning to extract data on morphometric characteristics from embryo images. Unlike our study, these authors did not use time-lapse images but rather conventional microscopy. On the basis of staining of the blastocyst image performed by the embryologist, the algorithm was able to extract 24 image attributes that were related to pixel intensity, area and perimeter, resulting in accuracies of 0.75 and 0.62 (for the support vector machines method, respectively) and for the random forest method, per Chavez-Badiola et al.⁴⁶. The method can be considered semiautomated and has been previously proposed for murine embryos⁴⁷. In our study, all processing (i.e., digital image processing, mode prediction of the MLP ANNs, the MAIA score and the fold change) occurs automatically from the moment the blastocyst image is selected for analysis by MAIA.

Images in multiple focal planes of the embryo have been used in the development of automated models for the evaluation of embryonic quality. Wang et al.⁴⁸ applied a deep learning model to automate the evaluation of the quality of “good-quality” or “bad-quality” blastocysts and used more than 10,000 images of embryos from SLT in 11 focal planes. The training of the model developed was based on the classification by Gardner and Schoolcraft⁶ and was performed by 5 experienced embryologists. They obtained an accuracy of 0.91 and an AUC of 0.93. Similar to what was performed in this study, Wang et al.⁴⁸ removed images in which the blastocyst was incomplete or the image was blurred from the database. These data are not yet related to implantation or live birth data and were derived from only one reproductive care centre⁴⁸. Depending on the image processing method used, the use of multiple focal planes may be more useful than the use of a single image in a focal plane for automating the morphological evaluation of the blastocyst. However, it can lead to inefficient AI, as there may be focal planes in which the embryo is blurred, leading to misinterpretation by AI. In our study, 33 input

variables derived from processing¹⁶ were obtained from a single image (i.e., a single focal plane) and are directly correlated with the clinical outcome of the embryos.

Evaluating the entire video of embryonic development captured in time-lapse culture using EmbryoScope®, Tran et al.²⁴ applied an AI model based on deep learning, called IVY. In this proposal, more than 8,000 embryos from 8 clinics were used. The model obtained an average AUC of 0.93 in a 5-fold stratified cross-validation for the prediction of a successful pregnancy. Interestingly, although the dataset used for training is considered unbalanced (mostly negative CP data compared with positive CP data), in the experiments used to define this stage, the best result was achieved — possibly through training involving the entire embryonic development process, from zygote to blastocyst. The AUC was the only measure described in the study. Additionally, although multicentre data were used, the study was limited to a retrospective analysis²⁴.

In a subsequent study in which the database used for publication by Tran et al.²⁴ was expanded, Berntsen et al.⁴⁹ applied deep learning, called iDAScore v1.0, to predict a foetal heartbeat. They used 115,832 image sequences from EmbryoScope® Plus, in partnership with 18 clinics. Unlike the present study, the authors considered not only transferred embryos but also nontransferred embryos (termed discarded) in the model training process, and these embryos were pseudo-labelled as negative predictors in an attempt to make the evaluation more automated considering all the embryos in a cohort. The model obtained an AUC of 0.95 considering the entire cohort of embryos and 0.67 in the test set for embryos with known implantation⁴⁹.

With the implantation outcome known, Fruchter-Goldmeier et al.⁵⁰ performed a retrospective study that included 608 embryos also incubated at EmbryoScope®. Unlike our study, they did not use embryos transferred after cryopreservation; embryos were analysed by preimplantation genetic testing, and only autologous cycles were considered. For training of the model developed, manual marking was performed in the parts corresponding to the ICM, trophoctoderm and blastocyst perimeter by experienced embryologists. The segmentation model applied to the manual markings of the blastocysts was the Chloe™ by Fairtility LTD, and a result of 0.70 was obtained. The authors concluded that the embryos that resulted in implantation had a lower ICM size-to-blastocyst size ratio than did those that did not result in implantation. This segmentation model is similar to our study in that it determines quantitative measurements of parts of the blastocyst, as well as its entirety, for application in AI⁵⁰ although our study is fully automated in embryo segmentation.

Although our study presents an AI model developed on the basis of retrospective data and, in addition, on the performance of MAIA in a multicentre clinical routine trial, a prospective randomized double-blind noninferiority trial has not yet been performed. A study with these characteristics was performed by Illingworth⁵¹ who demonstrated that deep learning is not inferior to standard morphology in terms of clinical pregnancy.

In this regard, our study has several limitations. Despite the high training AUC results and the lower performance in the internal validation (Table 3), there is no clear evidence of model overfitting. This is further supported by the 66.5% accuracy observed in multicentre clinical tests, which although still subject to optimization, demonstrates the model's ability to generalize and perform consistently in real-world clinical settings. Furthermore, prospective double-blind randomized analysis of any AI model is an important step after retrospective analysis and its absence can be considered one of the main limitations of the present study. The morphokinetic parameters of embryonic development and the clinical data of patients constitute additional information to be applied as inputs in our algorithm in the future.

In recent years, the use of AI in human reproduction has increased^{46,52,53} and there are commercially available options for its routine clinical use⁵⁴. An example is iDAScore, a deep learning algorithm developed for Vitrolife's EmbryoScope Plus, which uses 3D convolutional neural networks to automatically identify both spatial (morphological) and temporal (morphokinetic) patterns from raw time-lapse image sequences in order to rank embryos and predict clinical pregnancy outcomes⁵⁵. In a validation study of iDAScore v2.0, Lassen et al. (2023)⁵⁵ evaluated the algorithm's performance using data from over 100,000 embryos at different days of development. For embryos assessed on day 5 or later, the model achieved test AUC of 0.694. In our study, MAIA demonstrated a comparable performance in the multicentre test, reaching an AUC of 0.64 for embryos on days 5 and 6.

In contrast, Chloe system, from the Fairtility group, uses AI algorithms to analyse images of embryos in time during IVF and automatically notes and classifies morphokinetic and morphological events, providing information for embryo selection and clinical research and automated annotations. Evidence regarding the application of Chloe in routine clinical practice remains limited. In a retrospective study, the model reported an AUC of 0.64 for predicting clinical pregnancy following SET; however, it is unclear whether this value refers to a training or test dataset, which limits the interpretability and generalizability of the finding⁵⁶.

Developed, trained and validated entirely in Brazil, the MAIA software has the potential to become part of the daily routine of assisted reproductive care clinics, given its prospective predictive performance in 3 different IVF clinics. MAIA may provide support for the appropriate selection of embryos to be transferred on the basis of the automatic evaluation of a single image of the blastocyst obtained without interrupting its culture, improving the gestational success rate and reducing the number of cycles required for a blastocyst to yield a healthy pregnancy.

To our knowledge, no other study has proposed the use of predictive variables derived from morphological quality for application in AI software (using MLP and GA ANNs) and use in clinical practice with a user-friendly interface. Additionally, the methodology proposed in our study (through image processing prior to AI analysis) is original to our group¹⁶. Because it is fully automated, together with the application of AI methods, it makes the process less dependent on the subjectivity and experience of the embryologist in the evaluation and annotation of morphokinetic or morphological variables, which are normally included in traditional embryonic analysis. This fact was observed in a study using a methodology that was extremely similar to that of the present study⁴⁵ with in vitro-produced bovine blastocysts, in which the agreement (Cohen's kappa statistic) between

	Clinic A	Clinic B	Clinic C	Overall
# Patients	701 (78.7%)	98 (11%)	92 (10.3%)	891
# IVF cycles/SET	817 (80.5%)	102 (10.0%)	96 (9.5%)	1015
Mean age (years)	38.9 ± 4.5	38.9 ± 4.8	37.9 ± 3.8	38.8 ± 4.5
Mean BMI	23.1 ± 3.2	23.8 ± 3.6	22.3 ± 2.8	23.0 ± 3.2
CP rate (%)	49.3% (403/817)	43.1% (44/102)	46.9% (45/96)	48.5% (492/1015)
Autologous cycles (%)	74.8% (611/817)	61.8% (63/102)	82.3% (79/96)	74.2% (753/1015)
Mean age	37.5 ± 3.6	36.5 ± 4.0	37.7 ± 3.7	37.4 ± 3.7
BMI	22.9 ± 3.2	23.4 ± 3.5	22.1 ± 2.8	22.8 ± 3.2
PGT-A (%)	75.6% (462/611)	68.2% (43/63)	75.9% (60/79)	75.0% (565/753)
Frozen embryo transfer (%)	99.0% (605/611)	92.1% (58/63)	87.3% (69/79)	97.2% (732/753)
CP rate (%)	48.6% (297/611)	44.4% (28/63)	46.8% (37/79)	48.1% (362/753)
Donor egg cycles (%)	25.2% (206/817)	38.2% (39/102)	17.7% (17/96)	25.8% (262/1015)
Mean age	43.0 ± 4.2	42.9 ± 3.1	40.5 ± 3.4	42.8 ± 4.0
BMI	23.6 ± 3.3	24.7 ± 3.6	23.4 ± 3.0	23.7 ± 3.3
PGT-A (%)	19.9% (41/206)	17.9% (7/39)	35.3% (6/17)	20.6% (54/262)
Frozen embryo transfer (%)	26.7% (55/206)	35.9% (14/39)	35.3% (6/17)	32.8% (86/262)
CP rate (%)	51.5% (106/206)	41.0% (16/39)	47.1% (8/17)	49.6% (130/262)
# Embryos excluded	70 (8.6%)	10 (9.8%)	6 (6.25%)	86 (8.5%)

Table 7. Patient and cycle characteristics used for MAIA model training.

the 3 best-trained embryologists was lower than the agreement of the 3 best MLP ANNs when the same digital image was analysed.

In conclusion, we developed a fully automated AI-based software, MAIA, capable of ranking blastocyst images uploaded by the user and providing a robust, objective assessment to complement the embryologist's expertise. In this study, MAIA demonstrated a strong correlation with clinical pregnancy outcomes and, in some metrics, performed comparably or superior to embryologists' selections. These findings suggest that MAIA can serve as a valuable decision-support tool, enhancing consistency and objectivity in embryo selection while preserving the clinical judgment of experienced professionals.

Methods
Database

In this retrospective study, data from 1,015 embryos from 1,015 in vitro fertilization cycles of 891 patients who underwent single embryo transfer (fresh and frozen between November 2017 and June 2022) at three different assisted reproduction centres were used. Detailed patient and cycle characteristics are summarized in Table 7. The study was approved by the Brazilian National Council for Research Ethics (CONEP), through the Research Ethics Committee (CEP) of Hospital Heliópolis – UGA I, São Paulo/Brazil, under number CAAE 06081218.4.0000.5449. In addition, all patients signed the Free and Informed Consent Form. This study follows the guidelines stipulated in TRIPOD – AI in the development and evaluation of the prediction model⁵⁷.

The mean age of all patients was 38.8 ± 4.5 years, and the mean BMI was 23.0 ± 3.2. Patients in autologous cycles (mean age 37.4 ± 3.7 years and BMI 22.8 ± 3.2) constituted 74.2% (n = 753) of the patients, and among them, 75.0% (n = 565) had undergone preimplantation genetic testing for aneuploidy (PGT-A). A total of 25.8% (n = 262) of the patients used donated eggs (mean age 42.8 ± 4.0 years and BMI 23.7 ± 3.3) and 20.6% (n = 54) underwent PGT-A. The inclusion criterion was a single embryo transfer with clinical pregnancy results confirmed by ultrasound (positive cases) and willingness to sign of the informed consent form at the Huntington Clinic Huntington Reproductive Medicine, São Paulo, Brazil.

The embryos were cultured to the blastocyst stage in incubators fitted with an Embryoscope + time-lapse system (Vitrolife) that acquires images every 10 min in 11 focal planes in 2048 × 1088 pixels (2.2 MP) with a 12-bit monochrome CMOS camera (EmbryoScope™+ incubator user manual, 2024)⁴³.

Using the focal planes provided by the EmbryoScope®, the embryologist selected and exported a single image of the expanded blastocyst — captured at the focal plane that offered the best visualization of the inner cell mass and trophectoderm — with a resolution of 500 × 500 pixels for analysis by MAIA. Of the 1,015 blastocyst images from the retrospective cohort, 755 images (74.4%) were used for the effective learning of AI, 174 were used for the internal validation (17.1%), and 86 were excluded from the study (8.5%). The reason for exclusion was incomplete or suboptimal visualization of the blastocyst, which could compromise accurate evaluation. Specifically, embryos were excluded due to incomplete visualization of the blastocyst (n = 30), oval-shaped appearance that hindered full structural assessment (n = 11), being out of focus (n = 23), not having reached the blastocyst stage (n = 15), or incomplete data in the database (n = 7). These exclusions were necessary to ensure that only high-quality, standardized images were used for reliable analysis and consistent AI assessment, as poor image quality or incomplete development could bias the results.

Standardization of blastocyst image processing

Previously developed digital processing software was used with the MATLAB® platform, which automatically extracts 33 mathematical variables representing the morphological characteristics of the embryo from the digital image of the blastocyst¹⁶.

Algorithm for artificial intelligence

For the prediction of clinical pregnancy positive or negative (CP + or CP -), an artificial intelligence algorithm was developed on the MATLAB platform, which uses the method for multilayer perceptron artificial neural networks associated with the genetic algorithm (GA) method.

The AI algorithm included the 33 variables derived from the previous digital processing of the blastocyst images as inputs. For the training and validation phases, an MLP ANN with 1 to 3 intermediate layers was used, where the number of neurons varied between 20 and 500 in each layer. The output was a numerical CP prediction vector (between the highest probability of CP + and the highest probability of CP -). The stopping criterion for the MLP ANNs was the number of epochs, which was between 50 and 700.

The learning algorithm employed was backpropagation, which minimizes error by comparing predicted and actual outputs and adjusting the connection weights accordingly. To train the MLP ANN, the dataset (755 images) was split into 70% for training and 30% for validation. Several hyperparameters were tuned during training, including the number of hidden layers, the number of neurons per layer, the learning rate, and the transfer functions — applied randomly during the learning process — included tansig, logsig, purelin, hardlim, tribas, radbas, and satlin^{58,59}.

The GA method was used to determine the most accurate MLP ANN architecture for the prediction of CP + and CP -. Initially, a random population with different MLP ANN architectures was built according to the aforementioned specifications. This population ranged from 100 to 1,000 individuals (i.e., the individual being the specific architecture of an MLP ANN). After the initial generation, the following generations were constructed considering 20 to 30% of the most accurate individuals of the previous population; 50 to 60% of the individuals were introduced by recombination (crossing over) of the individuals selected as the most accurate; 15% came from the migration of newly created MLP ANN architectures, and 5% came from mutation (i.e., previous architectures with random point modifications). Thus, the aim was to ensure that the populations (after the initial population) had constant variability and potential for the detection of the best individuals (more accurate MLP ANNs), a process called elitism⁶⁰. As the stopping criterion for the GA method, 100 to 500 generations were adopted, i.e., epochs (illustrative flowchart in Supplementary Text 6 and Supplementary Fig. 9). As an illustration, the complete process of applying the AI method, from digital processing to the ANN MLP and the GA, is shown in Supplementary Text 7 and Supplementary Fig. 10.

Multicentric routine clinical tests

The multicentric clinical tests were conducted as part of a prospective observational study carried out across three IVF centres of the Huntington Group (named A, B and C). There was no intervention applied to modify clinical or laboratory practice; instead, the study aimed to evaluate the performance of the MAIA algorithm in a real-world clinical setting, using embryo images and associated clinical outcomes obtained under routine care. At each centre, 3 previously trained embryologists were assigned to this test. Thus, a total of 9 embryologists participated in the evaluation. The tests were performed between October 2023 and August 2024, totalling 200 single embryo transfers.

The test was performed concurrently at the 3 centres, and the following data were computed: date of single embryo transfer; the patient's ID; the transferred embryo number; whether it was an elective case (more than one embryo available for transfer) or nonelective case (only one embryo available for transfer); whether the embryo was biopsied (when biopsied, only euploid embryos were considered – PGT-A threshold < 30% aneuploidy⁶¹); the day of embryonic development (day 4, day 5 or day 6); the total number of embryos analysed; the number of the first to fifth embryos (in descending order per the MAIA score and depending on the number of embryos available for each patient); the tiebreaker in the choice of embryo to be transferred (choice on the basis of MAIA, on the embryologist or if both agreed in the choice); and observations of the result generated by the “show process” button, such as an incorrectly segmented blastocyst image; if there was an early pregnancy (β-hCG positive or negative), if the result was positive, or if the result was a clinical pregnancy (i.e., the presence of a gestational sac and foetal heartbeat). The statistical analyses of the clinical trial data were subsequently performed separately and together for all 3 centres.

Data availability

The data used for training and optimization of the MLP artificial neural network are considered personal data of patients and are protected by the Brazilian General Data Protection Law and cannot be made publicly available. The data may be requested from the corresponding author and will be made available upon approval by the national ethics committee and after signing the terms of use agreement.

Code availability

The training and optimization code for MLP artificial neural networks as well as the use of the Genetic Algorithm are available at <https://gitfront.io/r/LaMAp924/R1N8PfpF2Ut/MAIA/>.

Received: 23 April 2025; Accepted: 26 August 2025

Published online: 02 September 2025

References

1. Graham, M. E. et al. Assisted reproductive technology: Short- and long-term outcomes. *Dev. Med. Child. Neurol.* **65**, 38–49 (2023).
2. Jiang, V. S. & Bormann, C. L. Artificial intelligence in the in vitro fertilization laboratory: a review of advancements over the last decade. *Fertil. Steril.* **120**, 17–23 (2023).
3. Devine, K. et al. Single vitrified blastocyst transfer maximizes liveborn children per embryo while minimizing preterm birth. *Fertil. Steril.* **103**, 1454–1460 (2015).
4. Tiitinen, A. Single embryo transfer: why and how to identify the embryo with the best developmental potential. *Best Pract. Res. Clin. Endocrinol. Metab.* **33**, 77–88 (2019).
5. Glatstein, I., Chavez-Badiola, A. & Curchoe, C. L. New frontiers in embryo selection. *J. Assist. Reprod. Genet.* **40**, 223–234 (2023).
6. Gardner, D. K. & Schoolcraft, W. B. Culture and transfer of human blastocysts. *Curr. Opin. Obstet. Gynaecol.* **11**, 307–311 (1999).
7. Sciorio, R. & Meseguer, M. Focus on time-lapse analysis: blastocyst collapse and morphometric assessment as new features of embryo viability. *Reprod. BioMed. Online.* **43**, 821–832 (2021).
8. Sundvall, L., Ingerslev, H. J., Knudsen, U. B. & Kirkegaard, K. Inter- and intra-observer variability of time-lapse annotations. *Hum. Reprod.* **28**, 3215–3221 (2013).
9. Gallego, R. D., Remohí, J. & Meseguer, M. Time-lapse imaging: the state of the Art. *Biol. Reprod.* **101**, 1146–1154 (2019).
10. VerMilyea, M. D. et al. Computer-automated time-lapse analysis results correlate with embryo implantation and clinical pregnancy: a blinded, multi-centre study. *Reprod. Biomed. Online.* **29**, 729–736 (2014).
11. Chéles, D. S., Molin, E. A. D., Rocha, J. C. & Nogueira, M. F. G. 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.* **24**, 470–479 (2020).
12. Rocha, C., Nogueira, M. G., Zaninovic, N. & Hickman, C. Is AI assessment of morphokinetic data and digital image analysis from time-lapse culture predictive of implantation potential of human embryos? *Fertil. Steril.* **110**, e373 (2018).
13. Zaninovic, N. et al. Application of artificial intelligence technology to increase the efficacy of embryo selection and prediction of live birth using human blastocysts cultured in a time-lapse incubator. *Fertil. Steril.* **110**, e372–e373 (2018).
14. Alegre, L. et al. First application of artificial neuronal networks for human live birth prediction on Geri time-lapse monitoring system blastocyst images. *Fertil. Steril.* **114**, e140 (2020).
15. Bori, L. et al. An artificial intelligence model based on the proteomic profile of euploid embryos and blastocyst morphology: a preliminary study. *Reprod. BioMed. Online.* **42**, 340–350 (2021).
16. Chéles, D. S. et al. An image processing protocol to extract variables predictive of human embryo fitness for assisted reproduction. *Appl. Sci.* **12**, 3531 (2022).
17. Jacobs, C. K. et al. Embryologists versus artificial intelligence: predicting clinical pregnancy out of a transferred embryo who performs it better? *Fertil. Steril.* **118**, e81–e82 (2022).
18. Lorenzon, A. et al. P-211 development of an artificial intelligence software with consistent laboratory data from a single IVF center: performance of a new interface to predict clinical pregnancy. *Hum. Reprod.* **39**, deae108.581 (2024).
19. Fernandez, E. I. et al. Artificial intelligence in the IVF laboratory: overview through the application of different types of algorithms for the classification of reproductive data. *J. Assist. Reprod. Genet.* **37**, 2359–2376 (2020).
20. Mendizabal-Ruiz, G. et al. Computer software (SiD) assisted real-time single sperm selection associated with fertilization and blastocyst formation. *Reprod. BioMed. Online.* **45**, 703–711 (2022).
21. Fjeldstad, J. et al. Segmentation of mature human oocytes provides interpretable and improved blastocyst outcome predictions by a machine learning model. *Sci. Rep.* **14**, 10569 (2024).
22. Khosravi, P. et al. Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization. *NPJ Digit. Med.* **2**, 21 (2019).
23. Hickman, C. et al. Inner cell mass surface area automatically detected using Chloe eq™ (fertility), an ai-based embryology support tool, is associated with embryo grading, embryo ranking, ploidy and live birth outcome. *Fertil. Steril.* **118**, e79 (2022).
24. Tran, D., Cooke, S., Illingworth, P. J. & Gardner, D. K. Deep learning as a predictive tool for fetal heart pregnancy following time-lapse incubation and blastocyst transfer. *Hum. Reprod.* **34**, 1011–1018 (2019).
25. Rajendran, S. et al. Automatic ploidy prediction and quality assessment of human blastocysts using time-lapse imaging. *Nat. Commun.* **15**, 7756 (2024).
26. Bormann, C. L. et al. Consistency and objectivity of automated embryo assessments using deep neural networks. *Fertil. Steril.* **113**, 781–787e1 (2020).
27. Kragh, M. F. & Karstoft, H. Embryo selection with artificial intelligence: how to evaluate and compare methods? *J. Assist. Reprod. Genet.* **38**, 1675–1689 (2021).
28. Cromack, S. C., Lew, A. M., Bazzetta, S. E., Xu, S. & Walter, J. R. The perception of artificial intelligence and infertility care among patients undergoing fertility treatment. *J. Assist. Reprod. Genet.* <https://doi.org/10.1007/s10815-024-03382-5> (2025).
29. Fröhlich, H. et al. From hype to reality: data science enabling personalized medicine. *BMC Med.* **16**, 150 (2018).
30. Zhu, J. et al. External validation of a model for selecting day 3 embryos for transfer based upon deep learning and time-lapse imaging. *Reprod. BioMed. Online.* **47**, 103242 (2023).
31. Yelke, H. K. et al. O-007 Simplifying the complexity of time-lapse decisions with AI: CHLOE (Fairtivity) can automatically annotate morphokinetics and predict blastulation (at 30hpi), pregnancy and ongoing clinical pregnancy. *Hum. Reprod.* **37**, deac104.007 (2022).
32. Papatheodorou, A. et al. Clinical and practical validation of an end-to-end artificial intelligence (AI)-driven fertility management platform in a real-world clinical setting. *Reprod. BioMed. Online.* **45**, e44–e45 (2022).
33. Salih, M. et al. Embryo selection through artificial intelligence versus embryologists: a systematic review. *Hum. Reprod. Open* hoad031 (2023).
34. Nunes, K. et al. Admixture's impact on Brazilian population evolution and health. *Science.* **388**(6748), eadl3564 (2025).
35. Jackson-Bey, T. et al. Systematic review of Racial and ethnic disparities in reproductive endocrinology and infertility: where do we stand today? *F&S Reviews.* **2**, 169–188 (2021).
36. Kassi, L. A. et al. Body mass index, not race, May be associated with an alteration in early embryo morphokinetics during in vitro fertilization. *J. Assist. Reprod. Genet.* **38**, 3091–3098 (2021).
37. Pena, S. D. J., Bastos-Rodrigues, L., Pimenta, J. R. & Bydlowski, S. P. DNA tests probe the genomic ancestry of Brazilians. *Braz J. Med. Biol. Res.* **42**, 870–876 (2009).
38. Fraga, A. M. et al. Establishment of a Brazilian line of human embryonic stem cells in defined medium: implications for cell therapy in an ethnically diverse population. *Cell. Transpl.* **20**, 431–440 (2011).
39. Amin, F. & Mahmoud, M. Confusion matrix in binary classification problems: a step-by-step tutorial. *J. Eng. Res.* **6**, 0–0 (2022).
40. Magdi, Y. et al. Effect of embryo selection based morphokinetics on IVF/ICSI outcomes: evidence from a systematic review and meta-analysis of randomized controlled trials. *Arch. Gynecol. Obstet.* **300**, 1479–1490 (2019).
41. Guo, Y. H., Liu, Y., Qi, L., Song, W. Y. & Jin, H. X. Can time-lapse incubation and monitoring be beneficial to assisted reproduction technology outcomes? A randomized controlled trial using day 3 double embryo transfer. *Front. Physiol.* **12**, 794601 (2022).
42. Giménez, C., Conversa, L., Murria, L. & Meseguer, M. Time-lapse imaging: morphokinetic analysis of in vitro fertilization outcomes. *Fertil. Steril.* **120**, 228–227 (2023).
43. Vitrolife EmbryoScope + time-lapse system. (2023). <https://www.vitrolife.com/products/time-lapse-systems/embryoscopeplus-ti-me-lapse-system/>.

44. Lagalla, C. et al. A quantitative approach to blastocyst quality evaluation: morphometric analysis and related IVF outcomes. *J. Assist. Reprod. Genet.* **32**, 705–712 (2015).
45. Rocha, J. C. et al. A method based on artificial intelligence to fully automatize the evaluation of bovine blastocyst images. *Sci. Rep.* **7**, 7659 (2017).
46. Chavez-Badiola, A. et al. Predicting pregnancy test results after embryo transfer by image feature extraction and analysis using machine learning. *Sci. Rep.* **10**, 4394 (2020).
47. Matos, F. D., Rocha, J. C. & Nogueira, M. F. G. A method using artificial neural networks to morphologically assess mouse blastocyst quality. *J. Anim. Sci. Technol.* **56**, 15 (2014).
48. Wang, S., Zhou, C., Zhang, D., Chen, L. & Sun, H. A deep learning framework design for automatic blastocyst evaluation with multifocal images. *IEEE Access.* **9**, 18927–18934 (2021).
49. Berntsen, J., Rimestad, J., Lassen, J. T., Tran, D. & Kragh, M. F. Robust and generalizable embryo selection based on artificial intelligence and time-lapse image sequences. *PLoS One.* **17**, e0262661 (2022).
50. Fruchter-Goldmeier, Y. et al. An artificial intelligence algorithm for automated blastocyst morphometric parameters demonstrates a positive association with implantation potential. *Sci. Rep.* **13**, 14617 (2023).
51. Illingworth, P. J. et al. Deep learning versus manual morphology-based embryo selection in IVF: a randomized, double-blind noninferiority trial. *Nat. Med.* **30**, 3114–3120 (2024).
52. Kanakasabapathy, M. K. et al. Development and evaluation of inexpensive automated deep learning-based imaging systems for embryology. *Lab. Chip.* **19**, 4139–4145 (2019).
53. Loewke, K. et al. Characterization of an artificial intelligence model for ranking static images of blastocyst stage embryos. *Fertil. Steril.* **117**, 528–535 (2022).
54. Hengstschläger, M. Artificial intelligence as a door opener for a new era of human reproduction. *Hum. Reprod. Open* hoad043 (2023).
55. Lassen Theilgaard, J., Fly Kragh, M., Rimestad, J., Nygård Johansen, M. & Berntsen, J. Development and validation of deep learning based embryo selection across multiple days of transfer. *Sci. Rep.* **13** (1), 4235 (2023).
56. Lozano, M. et al. P-301 Assessment of ongoing clinical outcomes prediction of an AI system on retrospective SET data, *Human Reprod.* **38**(Issue Supplement_1), dead093.659. (2023).
57. Collins, G. S. et al. TRIPOD + AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* **385**, e078378 (2024).
58. Abdolrasol, M. G. M. et al. Artificial neural networks based optimization techniques: a review. *Electronics* **10**, 2689 (2021).
59. Yuzer, E. O. & Bozkurt, A. Instant solar irradiation forecasting for solar power plants using different ANN algorithms and network models. *Electr. Eng.* **106**, 3671–3689 (2024).
60. Guariso, G. & Sangiorgio, M. Improving the performance of multiobjective genetic algorithms: an elitism-based approach. *Information* **11**, 587 (2020).
61. García-Pascual, C. M. et al. Optimized NGS approach for detection of aneuploidies and mosaicism in PGT-A and imbalances in PGT-SR. *Genes* **11**, 724 (2020).

Acknowledgements

This study was financed, in part, by the São Paulo Research Foundation (FAPESP), Brasil. Processes Numbers #2023/16156-1, #2023/08159-0, #2023/05345-8, #2020/07634-9, #2019/26749-4, #2018/19053-0, #2017/19323-5, # 2012/20110-2 and #2012/50533-2.

Author contributions

ELAM, JRA, ARL, MFGN, and JCR conceived the idea and planned the study. MCM, MBD, BAM, VCM, MFGN and JCR conducted the data analysis. MN, CKJ, BL, DSC, MBC, JRA, ELAM and ARL provided the clinical cases. MN, CKJ, DSC, ELAM, ARL, MFGN and JCR conducted the analysis and interpretation of all the experiments. DSC, JRA, ELAM, ARL, MFGN and JCR contributed to critical discussions. MN, CKJ, BL, MCM, DSC, MBD, BAM, VCM, MBC, JRA, ELAM, ARL, MFGN and JCR wrote and revised the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval and consent to participate

The study was conducted in accordance with the ethical standards established by the National Council for Research Ethics (CONEP - Brazil). Ethics committee approval for this study was obtained by the medical research ethics committee of the Heliópolis UGA I hospital - São Paulo/Brazil. Ethics committee approval decision number: CAAE 06081218.4.0000.5449. All patients provided written informed consent prior to participation in the study.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-17755-y>.

Correspondence and requests for materials should be addressed to J.C.R.

Reprints and permissions information is available at www.nature.com/reprints.

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025, corrected publication 2025