Improving embryo selection using a computer-automated time-lapse image analysis test plus day 3 morphology: results from a prospective multicenter trial

Joe Conaghan, Ph.D.,[a] Alice A. Chen, Ph.D.,[b] Susan P. Willman, M.D.,[c] Kristen Ivani, Ph.D.,[c] Philip E. Chenette, M.D.,[a] Robert Boostanfar, M.D.,[d] Valerie L. Baker, M.D.,[e] G. David Adamson, M.D.,[f] Mary E. Abusief, M.D.,[f] Marina Gvakharia, M.D., Ph.D.,[f] Kevin E. Loewke, Ph.D.,[b] and Shehua Shen, M.D.[b]

[a] Pacific Fertility Center, San Francisco; [b] Auxogyn, Inc., Menlo Park; [c] Reproductive Science Center of the Bay Area, San Ramon; [d] HRC Fertility, Encino; [e] Stanford Fertility and Reproductive Medicine Center, Palo Alto; and [f] Fertility Physicians of Northern California, San Jose, California

Objective: To assess the first computer-automated platform for time-lapse image analysis and blastocyst prediction and to determine how the screening information may assist embryologists in day 3 (D3) embryo selection.
Design: Prospective, multicenter, cohort study.
Setting: Five IVF clinics in the United States.
Patient(s): One hundred sixty women $\geq$18 years of age undergoing fresh IVF treatment with basal antral follicle count $\geq$8, basal FSH $<$10 IU/mL, and $\geq$8 normally fertilized oocytes.
Intervention(s): A noninvasive test combining time-lapse image analysis with the cell-tracking software, Eeva (Early Embryo Viability Assessment), was used to measure early embryo development and generate usable blastocyst predictions by D3.
Main Outcome Measure(s): Improvement in the ability of experienced embryologists to select which embryos are likely to develop to usable blastocysts using D3 morphology alone, compared with morphology plus Eeva.
Result(s): Experienced embryologists using Eeva in combination with D3 morphology significantly improved their ability to identify embryos that would reach the usable blastocyst stage (specificity for each of three embryologists using morphology vs. morphology plus Eeva: 59.7% vs. 86.3%, 41.9% vs. 84.0%, 79.5% vs. 86.6%). Adjunctive use of morphology plus Eeva improved embryo selection by enabling embryologists to better discriminate which embryos would be unlikely to develop to blastocyst and was particularly beneficial for improving selection among good-morphology embryos. Adjunctive use of morphology plus Eeva also reduced interindividual variability in embryo selection.
Conclusion(s): Previous studies have shown improved implantation rates for blastocyst transfer compared with cleavage-stage transfer. Addition of Eeva to the current embryo grading process may improve the success rates of cleavage-stage ETs. (Fertil Steril® 2013;100:412–9. ©2013 by American Society for Reproductive Medicine.)
Key Words: Embryo selection, time-lapse imaging, image analysis, noninvasive assessment, cell division

Discuss: You can discuss this article with its authors and with other ASRM members at http://fertstertforum.com/conaghanj-embryo-selection-time-lapse-imaging/

Use your smartphone to scan this QR code and connect to the discussion forum for this article now.*

* Download a free QR code scanner by searching for "QR scanner" in your smartphone's app store or app marketplace.

Received December 11, 2012; revised April 10, 2013; accepted April 11, 2013; published online May 28, 2013.
J.C. has nothing to disclose. A.A.C. is an employee and shareholder at Auxogyn Inc. S.P.W. has nothing to disclose. K.I. has nothing to disclose. P.E.C. has served as a clinical consultant to Auxogyn Inc. R.B. has nothing to disclose. V.L.B. has nothing to disclose. G.D.A. has served as a clinical consultant to Auxogyn Inc.; is chairman of and owns stock in Advanced Reproductive Care; and receives research funding from IBSA, LabCorp, and Schering Plough. M.E.A. has nothing to disclose. M.G. has nothing to disclose. K.E.L. is a shareholder at Auxogyn, Inc. S.S. is an employee and shareholder at Auxogyn Inc.
This study was supported by Auxogyn, Inc., (www.auxogyn.com).
The Eeva Test is CE Mark approved in Europe and limited to investigational use only in the United States as of the date of publication.
Reprint requests: Shehua Shen, M.D., Auxogyn, Inc., 1490 O'Brien Drive, Suite A, Menlo Park, California 94025 (E-mail: sshen@auxogyn.com).

Fertility and Sterility® Vol. 100, No. 2, August 2013 0015-0282/$36.00

http://dx.doi.org/10.1016/j.fertnstert.2013.04.021

Improvements in methods to select embryos for transfer would potentially enable further increases in IVF pregnancy rates and facilitate broader acceptance and adoption of single ET (1). Over the last decade, ET at the blastocyst stage has been used more frequently to maximize implantation rates and decrease the number of embryos transferred (2, 3). Although this strategy has almost doubled the implantation rate compared with cleavage-stage transfer (4–6), it involves prolonged culture, which currently results in blastocyst development in approximately half of all good-quality day 3 (D3) embryos (7–9). Consequently, blastocyst transfer is often avoided, especially when patients have only a few embryos or when it is deemed that embryos may not tolerate extended culture conditions (3). Further, evidence suggests that prolonged embryo culture may also increase the risk of epigenetic disorders, monozygotic twinning, preterm delivery, low birth weight, and other long-term health issues (7–11). Thus, reliable prediction of blastocyst formation by D3 may be useful, especially for IVF patients whose embryos are transferred on D3 after egg retrieval.

Time-lapse imaging is an emerging tool that allows the identification of parameters that may noninvasively predict the developmental potential of a cleavage-stage embryo through continuous monitoring (12–15). In previous studies, cell cycle timing parameters that were retrospectively extracted from large data sets of time-lapse videos have consistently shown strong correlations with human embryo development, including blastocyst formation, blastocyst quality, and implantation potential (12, 13, 16). However, the use of these retrospective parameters has not yet been validated in a prospective, multicenter clinical study and compared with traditional embryo selection methods that are based on morphology. Further, no time-lapse parameters have been tested in adjunct use with traditional morphological assessment to determine whether they may assist embryologists in selecting embryos that have the greatest potential to develop into usable blastocysts (i.e., blastocysts suitable for transfer or freezing).

Here we present results from a prospective, multicenter clinical study that demonstrate the ability of a noninvasive, computer-automated test to improve the prediction of usable blastocyst formation by D3 and provide clinical value to embryologists for D3 embryo selection. The test, termed Eeva (Early Embryo Viability Assessment), is enabled by dark-field, time-lapse imaging and cell-tracking software algorithms. Because the blastocyst stage is a key indicator of a viable embryo (3) and also the outcome for seminal basic research in in vitro human embryo time-lapse research (12), this clinical study used blastocyst formation as its endpoint.

MATERIALS AND METHODS

Overall Study Design

This was a prospective, single-arm, nonrandomized clinical study conducted at five IVF clinical sites in the United States between June 2011 and February 2012. Institutional Review Board approval was obtained at each site, and the study was registered at ClinicalTrials.gov (no. NCT01369446). The study was designed to develop and test Eeva (Auxogyn, Inc.), the first computer-automated platform for time-lapse image analysis and blastocyst prediction, and compare two methods of predicting which cleavage-stage embryos would develop successfully to a usable blastocyst: D3 morphological assessment alone or morphology supplemented by Eeva. Additional details of ovarian stimulation, fertilization, and embryo culture processes are presented in the Supplemental Material (Supplemental Methods).

Subjects and Recruitment

Written informed consent was obtained from all study participants. Patients were required to be at least 18 years of age and to be undergoing fresh IVF treatment using their own or donor eggs. The study allowed for ET on D3 or D5 and was divided into a development phase and a test phase (Supplemental Table 1). The inclusion criteria for the development phase were antral follicle count (AFC) of at least 8 and basal FSH of $<$10 IU/mL. The inclusion criteria for the test phase were stricter to enroll D5 candidates and were AFC of at least 12, FSH of $<$10 IU/mL, and at least eight normally fertilized oocytes (2PN). Patients who used a gestational carrier, used surgically retrieved sperm, used reinseminated oocytes, planned preimplantation genetic testing, had history of cancer treatment, were concurrently participating in another clinical study, or had previously enrolled in this study were excluded from the study.

A total of 160 patients at five IVF clinical sites met eligibility criteria and consented to have their embryos imaged. The 160 patients enrolled were allocated to training (12 patients with 98 embryos), a development phase (63 patients with 577 embryos), or a test phase (85 patients with 1,150 embryos). For the analyses presented here, only patients with D5 embryos with complete outcome and image data were included, resulting in a development phase with 45 patients with 292 embryos and a test phase with 74 patients with 941 embryos (Supplemental Fig. 1).

Embryo Imaging and Culture

Images of developing embryos were captured with Eeva, an integrated time-lapse imaging system that fits into a standard incubator and includes a dish that facilitates group culture while maintaining the ability to track and identify embryos individually; a digital, inverted time-lapse microscope with dark-field illumination, auto-focus, and digital camera; and image acquisition software. The image acquisition software captures a single, high-resolution, single-plane image of all of the microwells in the petri dish once every 5 minutes. The analysis is performed separately for each embryo; and the computations are performed so that all embryos can be processed individually. Eeva was designed to record embryo development with minimal light exposure to embryos using a light-emitting diode at 625 nm, producing only 0.32 joules/cm^2 over 3 days of imaging, or approximately the equivalent of 21 seconds total exposure from a traditional IVF bright field microscope (Supplemental Methods).

To maintain a continuous and uninterrupted imaging process from D1 through D3, no media changes or excursions

from the incubator were permitted. Imaging was discontinued at the time of the clinic's routine D3 embryo grading procedure. After the completion of Eeva imaging on D3, embryos were cultured and tracked individually to maintain their identities, and the imaging phase and remainder of the IVF process were completed according to the standard operating procedures at each site. Each clinic, for example, used its own preferred culture medium, incubators, and gas phase throughout the culture period and performed intracytoplasmic sperm injection (ICSI) or conventional insemination at its own discretion. D3 and D5 embryo grading was performed according to the clinic's standard protocols, using the Society for Assisted Reproductive Technology (SART) standard grading system (17, 18). Traditional D3 morphology grading was performed at an average time of 66.2 ± 2.0 hours postinsemination across sites. All embryos remaining after transfer were cultured to D5 in standard culture dishes according to the protocol of each clinical site, and traditional D5 morphology was assessed. Traditional D5 morphology grading occurred at an average time of 114.9 ± 2.6 hours postinsemination across sites.

In total, 2,901 oocytes were retrieved and fertilized by IVF or ICSI. After fertilization, 1,727 were confirmed as 2PNs and transferred to Eeva dishes immediately after the fertilization check, where they remained until the morning of D3. The overall blastocyst development rate was 58.4% (720/1,233), and the formation of usable blastocysts (i.e., blastocysts that were selected for transfer or freezing on D5) was 30.6% (377/1,233; Fig. 1). Since the purpose of this work was to develop and independently test an automated blastocyst prediction algorithm, the imaging data and software predictions were not provided to the embryologists, physicians, or patients at the time of ET.

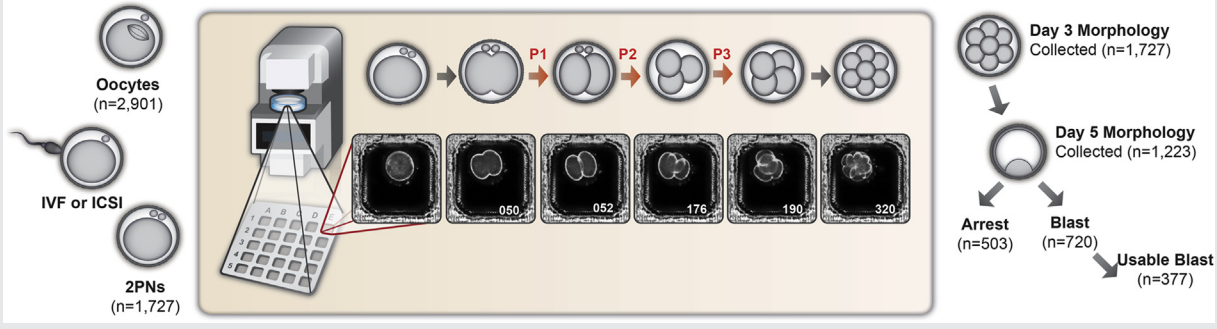
Development Phase

For all embryos in the development phase, three observers reviewed embryo videos and recorded the start/stop times of specific cell division time intervals from the 1- to 4-cell stage: the duration of first cytokinesis (P1), time between cytokinesis 1 and 2 (P2), and time between cytokinesis 2 and 3 (P3). The P1, P2, and P3 parameters were candidates for automation as they were previously reported to predict human blastocyst formation and to correlate with normal gene expression profiles (12). A classification tree model was built to determine timing windows for the P1, P2, and P3 cell cycle measurements that predicted blastocyst formation. Evidence from us and other investigators suggests that slower blastocyst formation is associated with poorer embryo viability (19–22); therefore, the model was further refined to predict usable blastocyst outcome, which was defined as blastocysts formed on D5 and suitable for transfer or freezing. Next, automated image analysis software was built and implemented in C^{++}, a programming language running in real time on a standard PC, to track cell divisions from the 1-cell to 4-cell stage. The primary features tracked by the software are cell membranes, which exhibit high image contrast through the use of dark-field illumination. By using a data-driven probabilistic framework and computational geometry, the software generates an embryo model that includes an estimate of the number of blastomeres, as well as blastomere size, location, and shape, as a function of time. Parameter measurements from the embryo models are fed through the classification tree that predicts usable blastocyst formation on D3. To illustrate the automated cell-tracking results, colored rings are overlaid on an original image of the embryo at each cell stage, for each frame of a time-lapse sequence (Fig. 2).

Test Phase

The test phase was performed using patient and embryo data that were independent from the data used for the development phase. Eeva performance was assessed by comparing its predictions with usable blastocyst formation and calculating diagnostic measures (e.g., specificity, sensitivity, positive predictive value [PPV], negative predictive value [NPV], and

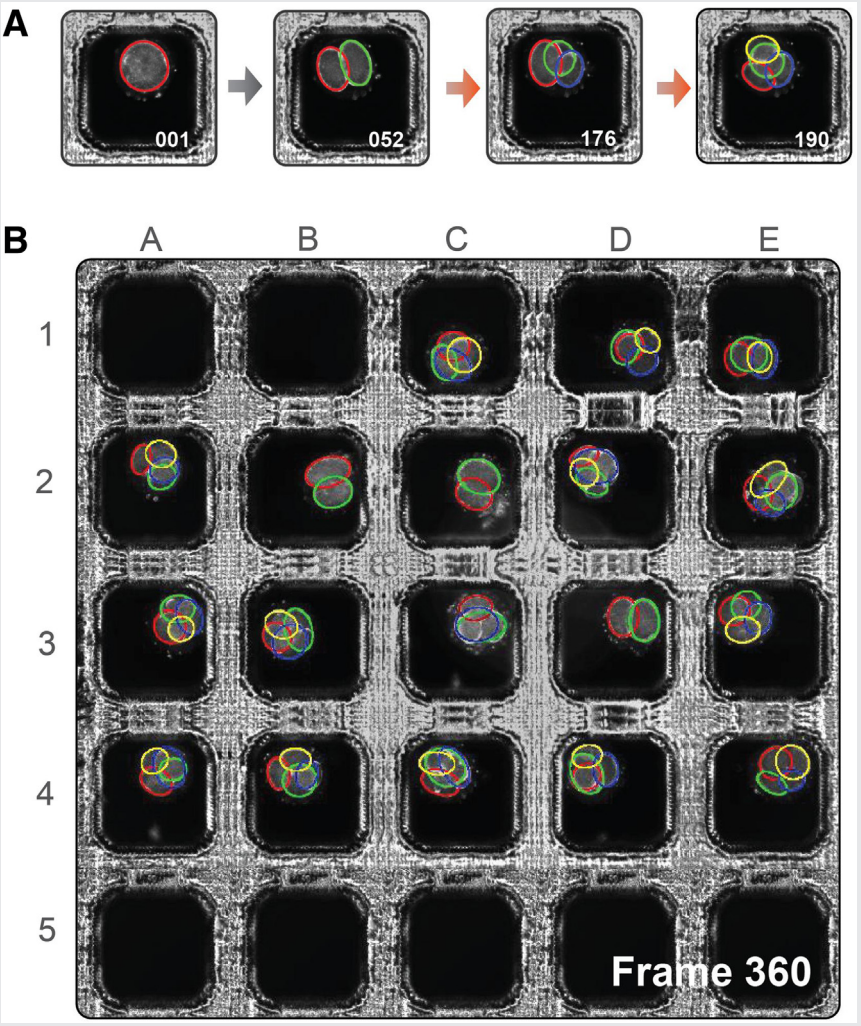
FIGURE 1



Schematic representation of the clinical study workflow at each of five IVF sites. Oocytes were retrieved and fertilized by IVF or ICSI as per each clinic's standard protocol. Successfully fertilized oocytes (2PNs) were cultured in a multiwell dish and imaged in a standard incubator using Eeva, which captured one dark-field image every 5 minutes for 3 days (insets show embryo development and frame numbers from the 1-cell to 8-cell stage). Following imaging, key cell division timing parameters (P1 = duration of first cytokinesis; P2 = time interval between cytokinesis 1 and 2; P3 = time interval between cytokinesis 2 and 3) were manually measured and used to develop and independently validate a model that could predict usable blastocyst outcome at the cleavage stage.

Conaghan. Validation of a time-lapse screening tool. Fertil Steril 2013.

FIGURE 2



Cell-tracking software developed and validated for enabling image analysis and automated prediction. Shown are the representative cell-tracking results for 18 human embryos captured at various developmental stages in a single-well (A) and multiwell dish (B). Colored rings represent the cell-tracking software's automatic delineation of cell membranes and cell divisions.

Conaghan. Validation of a time-lapse screening tool. Fertil Steril 2013.

associated 95% confidence intervals [CI]). In addition, a prospective, double-blinded comparison study compared P2 and P3 measurements and predictions generated automatically by Eeva to P2 and P3 measurements and predictions generated manually by a panel of observers who reviewed embryo videos. Agreement between the observer panel and Eeva was defined as both Eeva and manual methods having "high" (in window) or "low" (outside window) likelihood of usable blastocyst formation. The test phase also assessed whether the ability of experienced embryologists to select usable blastocysts on D3 was impacted by adding the Eeva prediction information to their traditional D3 morphological assessments (adjunct assessment).

Adjunct Assessment

Three clinical embryology laboratory directors (separate from the observer panel used to develop and test the Eeva prediction model) reviewed data in two independent sessions that assessed their prediction of usable blastocyst formation. During the first prediction session, embryologists were given D3 morphology (SART) data, including number of cells, fragmentation (0%, <10%, 11%–25%, >25%), symmetry (perfect, moderately asymmetrical, severely asymmetrical), and age of patient or egg donor. Each embryologist was blinded to the predictions of other embryologists. One week later, during a second prediction session, the same embryologists were given D3 morphology (SART) data as above and Eeva data for the same embryos. In this session, each embryologist was blinded to the predictions of other embryologists and the predictions from the first session. Eeva data included the cell cycle parameter values (P2 and P3) and a prediction score of "high" or "low" probability of usable blastocyst formation, based on the classification tree cutoffs determined in the development phase. To quantify the embryo selection

performance of the two methods, predictions made in each session (using morphology only or morphology plus Eeva) were compared with the usable blastocyst outcome.

Statistical Analyses

All data and statistical analyses were carried out using SAS software version 9.2 and Matlab version R2010a. To test non-inferiority between manual and software measurements, methods of Blackwelder were used with power $(1-\beta)\% = 0.8$ and significance $\alpha\% = 0.05$. Overall percent agreement between the two methods was determined using a method agreement analysis. A proportions test was used to statistically compare prediction method performances, and a value of $P<.05$ was considered statistically significant.

RESULTS

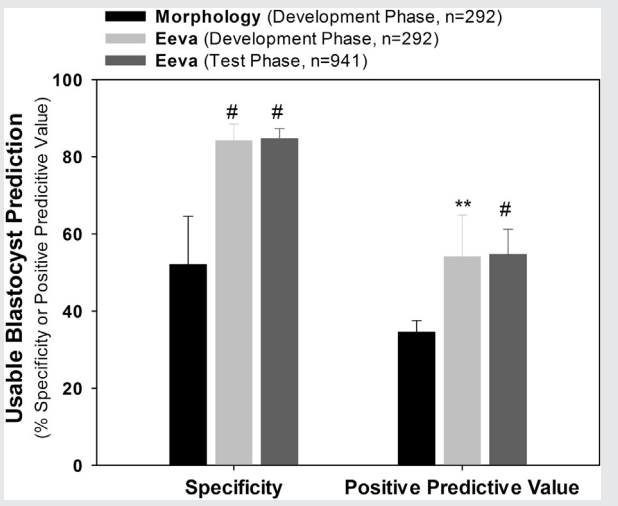
Development Phase

The Eeva prediction and cell-tracking software results indicated a high probability of usable blastocyst formation when both P2 and P3 are within specific cell division timing ranges (P2, 9.33–11.45 hours; and P3, 0–1.73 hours) and a low probability when either P2 or P3 are outside the specific cell division timing ranges. The time between cytokinesis 1 and 2 (P2) and the time between cytokinesis 2 and 3 (P3) cell cycle timings dominated the prediction model, and the duration of first cytokinesis (P1) was of lesser statistical value; therefore, prediction and cell-tracking software were based on P2 and P3 only. In the development phase, the Eeva prediction and cell-tracking software correctly predicted by D3 the embryos that became usable blastocysts with a specificity of 84.2% (95% CI = 78.7%–88.5%), sensitivity of 58.8% (95% CI = 47.0%–69.7%), positive predictive value (PPV) of 54.1% (95% CI = 42.8%–64.9%), and negative predictive value (NPV) of 86.6% (95% CI = 81.3%–90.6%). By comparison, morphology-based methods correctly identified those embryos that became usable blastocysts with a specificity of 52.1% (95% CI = 39.7%–64.6 %), sensitivity of 81.8% (95% CI = 70.6%–92.9%), PPV of 34.5% (95% CI = 31.5%–37.5%), and NPV of 90.9% (95% CI = 87.3%–94.5%). Compared with morphology evaluation alone, Eeva significantly improved the specificity (84.2% vs. 52.1%; $P<.0001$) and PPV (54.1% vs. 34.5%; $P<.01$) of usable blastocyst predictions in the development phase (Fig. 3 and Supplemental Table 2).

Test Phase

In the test phase, the Eeva prediction and cell-tracking software correctly predicted the embryos that became usable blastocysts with a specificity of 84.7% (95% CI = 81.7%–87.3%), sensitivity of 38.0% (95% CI = 32.7%–43.5%), PPV of 54.7% (95% CI = 48.0%–61.2%), and NPV of 73.7% (95% CI = 70.4%–76.8%). As in the development phase, Eeva significantly improved the specificity (84.7% vs. 52.1%; $P<.0001$) and PPV (54.7% vs. 34.5%; $P<.0001$) of usable blastocyst predictions in the test phase, compared with morphology evaluation alone (Fig. 3 and Supplemental Table 2). In addition, the overall agreement between the predictions made based on Eeva software timing measurements versus the predictions made based on manual timing measurements was 91.0% (95% CI = 86.0%–94.3%).

FIGURE 3



Usable blastocyst prediction performance (% Specificity or % PPV) of Eeva compared with D3 morphology for two independent data sets in the development and test phases. Error bars represent upper 95% CI. $^{**}P<.01$, $^{\#}P<.0001$.

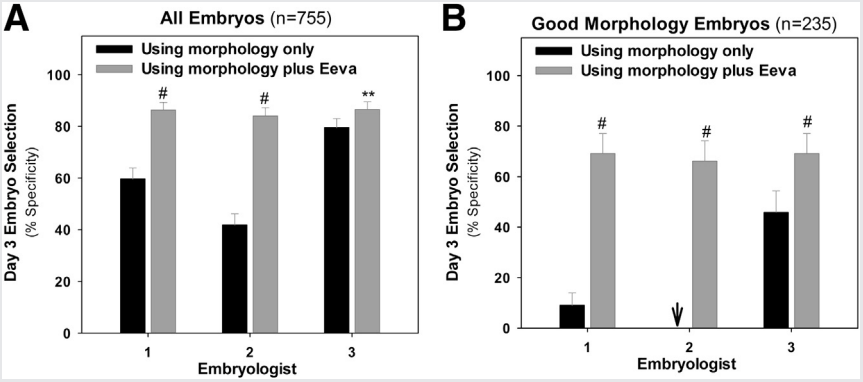
Conaghan. Validation of a time-lapse screening tool. Fertil Steril 2013.

Adjunct Assessment

The utility of combining Eeva with traditional morphology assessment for D3 embryo selection was examined using a subanalysis of patients with full cohorts of D5 embryos. For this phase, three clinical embryology lab directors from separate laboratories used two D3 assessments (morphology only or morphology plus Eeva) to select which cleavage-stage embryos would become usable blastocysts. Using D3 morphology only, embryologists 1, 2, and 3 selected embryos with a baseline specificity of 59.7%, 41.9%, and 79.5% and a baseline PPV of 45.5%, 41.5%, and 50.5%. When Eeva information was added to morphology on D3, each embryologist improved their selection of usable blastocysts to a specificity of 86.3% ($P<.0001$), 84.0% ($P<.0001$), and 86.6% ($P<.01$; Fig. 4A) and a PPV of 56.3% ($P<.05$), 52.1% ($P<.05$), and 55.5% ($P=.34$). The improvement for all embryologists was also accompanied by a reduction in variability among embryologists. Using D3 morphology alone, there was a 37.7% maximum difference in specificity and 8.9% maximum difference in PPV among embryologists. In contrast, using D3 morphology plus Eeva, there was a 2.5% maximum difference in specificity and 4.2% maximum difference in PPV.

Because standard morphological grading can identify good-morphology embryos, we assessed whether Eeva could help embryologists discriminate on D3 which good-morphology embryos would most likely develop to the usable blastocyst stage. For this analysis, embryos with good morphology were defined as having 6–10 cells, <10% fragmentation, and perfect symmetry. Using morphology only,

FIGURE 4



Adjunct assessment. D3 embryo selection by individual embryologists (1, 2, and 3) using morphology only versus morphology plus Eeva for (A) all embryos (n = 755) and (B) good-morphology embryos (n = 235). Good morphology is defined by 6–10 cells, <10% fragmentation, and perfect symmetry. Note that embryologists 1 and 2 were very conservative in their morphology assessments and expected that almost all D3 good-morphology embryos would become usable blastocysts. An alternative definition of good morphology defined by 7–8 cells, <10% fragmentation, and perfect symmetry is presented in Supplemental Figure 2 and shows similar results. Error bars represent upper 95% CI. $^{**}P<.01$, $^{\#}P<.0001$.

Conaghan. Validation of a time-lapse screening tool. Fertil Steril 2013.

embryologists 1, 2, and 3 varied considerably in their selection of which good embryos would become usable blastocysts (specificity 9.0%, 0.0%, and 45.9%, respectively). Using morphology plus Eeva, each embryologist improved their D3 selection to a specificity of 69.2% ($P<.0001$), 66.2% ($P<.0001$), and 69.2% ($P<.01$), respectively (Fig. 4B). For embryos with poor morphological criteria on D3, the selections of all embryologists were also improved (specificity, 77.5% vs. 92.3%, $P<.0001$ for embryologist 1; 56.5% vs. 90.3%, $P<.0001$ for embryologist 2; 91.3% vs. 92.6%, $P=.54$ for embryologist 3). These data show that, combined with D3 morphological assessment, Eeva provides valuable information to help embryologists identify which embryos that are favored by morphology are likely to subsequently arrest.

DISCUSSION

Recent studies have demonstrated that morphological and temporal assessment of human embryos through time-lapse imaging can identify parameters that correlate with developmental outcomes (12, 13–15). However, clinical adoption of these parameters requires additional documentation including scientific and clinical validation, prospective demonstration of clinical utility, and assessment of compatibility with the fast-paced and high-volume workflow of the IVF laboratory (23).

Here we report results of a prospective, multicenter clinical study that aimed to characterize the effectiveness and utility of Eeva, a noninvasive, computer-automated test of blastocyst formation based on validated cell cycle parameters and cell-tracking software algorithms. Embryos were transferred according to the usual procedure of participating clinics, and Eeva data were not made available at the time of transfer. Our objectives focused on assessment of Eeva's ability to predict usable blastocyst formation and assessment of the utility of Eeva predictions to aid embryologists in the selection of embryos on D3.

Development Phase

Eeva prediction parameters. Using Eeva to validate the usefulness of cell cycle parameters captured by time-lapse imaging in multiple IVF clinics, we first observed that usable blastocyst formation could be predicted at the cleavage stage using cell cycle timings similar to those previously discovered for frozen embryos (12). In particular, the timings for P2 (the time between the first and second cytokinesis) and P3 (the time between the second and third cytokinesis) were found to be predictive. Compared with the originally reported range for P1, the duration of the first cytokinesis broadened in the clinical data set but still fell within a relatively narrow range of approximately 30 minutes. P2 and P3 were found to statistically dominate P1 in the prediction model; therefore, a simple classification tree incorporating P2 and P3 only was used for developing Eeva prediction and cell-tracking software. Other groups have gone on to independently observe P2 and P3 as predictors of high-quality blastocyst development and implantation (13, 16, 24), even under variable stimulation protocols (25), fertilization methods (26), and culture environments (27). These studies support the conclusion of Wong et al. (12) that embryo fate is established, in large part, by the four-cell stage and may be primarily inherited from the oocyte; however, the contribution of the sperm remains to be determined. Results also support the conclusion that reported cell cycle parameters are reliable across independent data sets, as well as studies using different methodologies and measures of developmental potential (15).

Time-lapse image analysis and benefits of automation. The clinical and scientific promise of the P2 and P3 cell cycle parameters inspired the development of automated cell-tracking software that could be incorporated into a busy laboratory. While there have been a few recent technical reports on automated image analysis of human embryo microscope images

(28, 29), to our knowledge, there has been only a single successful demonstration of predictive software applied to time-lapse imaging of human embryos to predict development (12). We extended the cell-tracking framework introduced for 14 cryopreserved embryos in Wong et al. (12) and robustly validated the tracking and prediction accuracy of the Eeva software on a large, independent data set of fresh human embryos. Eeva software predictions had very high (>90%) agreement with manual predictions made by human observers and disagreed largely in cases where embryos exhibited complex dynamic behaviors that were also difficult to assess manually (data not shown).

Test Phase: Independent Validation of Eeva

Distinct from previous retrospective studies, we prospectively and blindly tested Eeva's integrated prediction and cell-tracking capabilities on clinical data that were independent from the development data set and collected from multiple, independent clinics. Our findings indicated that the parameters discovered by Wong et al. using cryopreserved and thawed supernumerary embryos (12) could be extended to develop an automated blastocyst prediction model that is standardized to multiple clinics using diverse culture protocols. Together with other reports using manual time-lapse analysis for outcomes of blastocyst formation and implantation (13, 16, 24), the reproducible science underpinning these predictive parameters gives confidence that automated time-lapse assessment of key embryo developmental events (P2, P3) may add value to current embryo selection techniques.

Adjunct Assessment: Using Morphology Plus Eeva to Aid Embryologists in D3 Embryo Selection

A demonstration of clinical utility is essential before any new tool is introduced into IVF laboratories. Therefore, an adjunct assessment subanalysis was conducted to assess whether adding automated Eeva predictions to traditional morphological methods could aid experienced embryologists in D3 embryo selection.

Results demonstrated that when Eeva was used in combination with D3 morphology, embryologists experienced significant improvement in the likelihood of selecting embryos that would develop to usable blastocysts. In particular, combining the high specificity of Eeva with traditional morphology methods dramatically improved the ability to determine the developmental potential of good-morphology embryos (where good morphology was defined as embryos with 6–10 cells, <10% fragmentation, and perfect symmetry; see Supplemental Fig. 2 for a variation on the good-morphology definition). Notably, there is strikingly high variability in the morphology-based selections of embryologists reviewing good embryos, as their specificities spanned from 0% (because one embryologist considered that all of these embryos would develop to usable blastocyst) to 45.9% (because of the less conservative approach of another embryologist).

Using morphology plus Eeva, the average of the three embryologists' prediction specificities were significantly improved (68.2% ± 1.7% for morphology plus Eeva vs. 18.3% ± 23.3% for morphology alone; $P<.05$). The embryologists' performances were also more consistent, as the SD among embryologists was reduced. It is widely accepted that morphological grading is accompanied by significant intra- and interoperator variability, which can impact IVF success rates (30, 31). Here we have built a generalized prediction algorithm based on multiclinic data and demonstrated that the automated prediction information can be added to embryologists' morphological evaluations to improve their interoperator variability. Combining the noninvasive, automated Eeva measurements with traditional morphology is therefore likely to provide embryologists with more consistent and objective data that may make embryo assessment on D3 more standardized, reproducible, and successful.

CONCLUSIONS

Eeva is the first clinically validated assessment platform that integrates time-lapse imaging, predictive parameters that are rooted in the molecular physiology of embryos, and automated cell-tracking software. Our study focuses on the ability of noninvasive computer-automated measurement of cell cycle times to predict usable blastocyst formation and the utility of these predictions for D3 embryo selection. This study is not without limitations and will be followed by studies assessing the embryo characteristics of a poorer prognosis patient population with AFC <12; statistical models focused on evaluation of additional outcomes of developmental competence including implantation; and whether using Eeva to select embryos for transfer on D3 improves pregnancy rates over embryo selection by morphology alone. In further development of this platform, future studies may also assess the ability to incorporate additional morphological features and timings, such as multinucleation, later cell divisions, blastocyst grading, and other prospectively validated parameters. Overall, the current results showed that adding Eeva to traditional D3 morphology significantly improved the ability of experienced embryologists to predict usable blastocyst outcomes and reduced the variability among embryologists. Use of this novel test alongside the current embryo grading process may improve the success of cleavage-stage transfers and has the potential to augment overall IVF pregnancy rates.

Acknowledgments: The authors thank Kelly Wirka, Vaishali Suraj, Farshid Moussavi, and Mahnaz Maddah for their technical contributions; Shawn Chavez, Yi-Wu Eve Shen, and Yimin Shu for assistance with embryo video measurements; and Professors Renee Reijo Pera and Michael Diamond for critical review of this manuscript and insightful discussions.

REFERENCES

1. van Montfoort AP, Dumoulin JC, Land JA, Coonen E, Derhaag JG, Evers JL. Elective single embryo transfer (eSET) policy in the first three IVF/ICSI treatment cycles. Hum Reprod 2005;20:433–6.
2. Gardner DK, Lane M, Stevens J, Schlenker T, Schoolcraft WB. Blastocyst score affects implantation and pregnancy outcome: towards a single blastocyst transfer. Fertility and sterility 2000;73:1155–8.

3. Diamond MP, Willman S, Chenette P, Cedars MI. The clinical need for a method of identification of embryos destined to become a blastocyst in assisted reproductive technology cycles. J Assist Reprod Genet 2012;29: 391–6.
4. Papanikolaou EG, D'Haeseleer E, Verheyen G, Van de Velde H, Camus M, Van Steirteghem A, et al. Live birth rate is significantly higher after blastocyst transfer than after cleavage-stage embryo transfer when at least four embryos are available on day 3 of embryo culture. A randomized prospective study. Hum Reprod 2005;20:3198–203.
5. Blake DA, Farquhar CM, Johnson N, Proctor M. Cleavage stage versus blastocyst stage embryo transfer in assisted conception. Cochrane Database Syst Rev 2002:D002118.
6. Gelbaya TA, Tsoumpou I, Nardo LG. The likelihood of live birth and multiple birth after single versus double embryo transfer at the cleavage stage: a systematic review and meta-analysis. Fertil Steril 2010;94:936–45.
7. Niemitz EL, Feinberg AP. Epigenetics and assisted reproductive technology: a call for investigation. Am J Hum Genet 2004;74:599–609.
8. Horsthemke B, Ludwig M. Assisted reproduction: the epigenetic perspective. Hum Reprod Update 2005;11:473–82.
9. Manipalviratn S, DeCherney A, Segars J. Imprinting disorders and assisted reproductive technology. Fertil Steril 2009;91:305–15.
10. Kallen B, Finnstrom O, Lindam A, Nilsson E, Nygren KG, Otterblad Olausson P. Trends in delivery and neonatal outcome after in vitro fertilization in Sweden: data for 25 years. Hum Reprod 2010;25:1026–34.
11. Kalra SK, Ratcliffe SJ, Barnhart KT, Coutifaris C. Extended embryo culture and an increased risk of preterm delivery. Obstet Gynecol 2012; 120:69–75.
12. Wong C, Loewke K, Bossert N, Behr B, De Jonge C, Baer T, et al. Non-invasive imaging of human embryos before embryonic genome activation predicts development to the blastocyst stage. Nat Biotechnol 2010;28: 1115–21.
13. Meseguer M, Herrero J, Tejera A, Hilligsoe KM, Ramsing NB, Remohi J. The use of morphokinetics as a predictor of embryo implantation. Hum Reprod 2011;26:2658–71.
14. Kirkegaard K, Agerholm IE, Ingerslev HJ. Time-lapse monitoring as a tool for clinical embryo assessment. Hum Reprod 2012;27:1277–85.
15. Chen AA, Tan L, Suraj V, Reijo Pera R, Shen S. Biomarkers identified with time-lapse imaging: discovery, validation, and practical application. Fertil Steril 2013;99:1035–43.
16. Cruz M, Garrido N, Herrero J, Perez-Cano I, Munoz M, Meseguer M. Timing of cell division in human cleavage-stage embryos is linked with blastocyst formation and quality. Reprod Biomed Online 2012;25:371–81.
17. Vernon M, Stern JE, Ball GD, Wininger D, Mayer J, Racowsky C. Utility of the national embryo morphology data collection by the Society for Assisted Reproductive Technologies (SART): correlation between day-3 morphology grade and live-birth outcome. Fertil Steril 2011;95:2761–3.
18. Racowsky C, Vernon M, Mayer J, Ball GD, Behr B, Pomeroy KO, et al. Standardization of grading embryo morphology. Fertil Steril 2010;94: 1152–3.
19. Shapiro BS, Richter KS, Harris DC, Daneshmand ST. A comparison of day 5 and day 6 blastocyst transfers. Fertil Steril 2001;75:1126–30.
20. Levens ED, Whitcomb BW, Hennessy S, James AN, Yauger BJ, Larsen FW. Blastocyst development rate impacts outcome in cryopreserved blastocyst transfer cycles. Fertil Steril 2008;90:2138–43.
21. Barrenetxea G, de Larruzea AL, Ganzabal T, Jiménez R, Carbonero K, Mandiola M. Blastocyst culture after repeated failure of cleavage-stage embryo transfers: a comparison of day 5 and day 6 transfers. Fertil Steril 2005;83:49–53.
22. Kovacic B, Vlaisavljevic V, Reljic M, Cizek-Sajko M. Developmental capacity of different morphological types of day 5 human morulae and blastocysts. Reprod Biomed Online 2004;8:687–94.
23. Harper J, Cristina Magli M, Lundin K, Barratt CLR, Brison D. When and how should new technology be introduced into the IVF laboratory? Hum Reprod 2012;27:303–13.
24. 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.
25. Munoz M, Cruz M, Humaidan P, Garrido N, Perez-Cano I, Meseguer M. Dose of recombinant FSH and oestradiol concentration on day of HCG affect embryo development kinetics. Reprod Biomed Online 2012;25:382–9.
26. 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.
27. Ingerslev H, Hindkjaer J, Kirkegaard K. Effect of oxygen concentration on human embryo development evaluated by time-lapse monitoring. Fertil Steril 2013;99:738–44.
28. Filho ES, Noble JA, Wells D. A review on automatic analysis of human embryo microscope images. Open Biomed Eng J 2010;4:170–7.
29. Filho ES, Noble JA, Poli M, Griffiths T, Emerson G, Wells D. A method for semi-automatic grading of human blastocyst microscope images. Hum Reprod 2012;27:2641–8.
30. Baxter Bendus AE, Mayer JF, Shipley SK, Catherino WH. Interobserver and intraobserver variation in day 3 embryo grading. Fertil Steril 2006;86:1608–15.
31. Paternot G, Devroe J, Debrock S, D'Hooghe TM, Spiessens C. Intra- and inter-observer analysis in the morphological assessment of early-stage embryos. Reprod Biol Endocrinol 2009;7:105.
32. Rijnders PM, Jansen CA. Influence of group culture and culture volume on the formation of human blastocysts: a prospective randomized study. Hum Reprod 1999;14:2333–7.
33. Lemmen JG, Agerholm I, Ziebe S. Kinetic markers of human embryo quality using time-lapse recordings of IVF/ICSI-fertilized oocytes. Reprod Biomed Online 2008;17:385–91.
34. Cruz M, Gadea B, Garrido N, Pedersen KS, Martinez M, Perez-Cano I, et al. Embryo quality, blastocyst and ongoing pregnancy rates in oocyte donation patients whose embryos were monitored by time-lapse imaging. J Assist Reprod Genet 2011;28:569–73.
35. Nakahara T, Iwase A, Goto M, Harata T, Suzuki M, Ienaga M, et al. Evaluation of the safety of time-lapse observations for human embryos. J Assist Reprod Genet 2010;27:93–6.
36. Payne D, Flaherty SP, Barry MF, Matthews CD. Preliminary observations on polar body extrusion and pronuclear formation in human oocytes using time-lapse video cinematography. Hum Reprod 1997;12:532–41.
37. Kirkegaard K, Hindkjaer JJ, Grondahl ML, Kesmodel US, Ingerslev HJ. A randomized clinical trial comparing embryo culture in a conventional incubator with a time-lapse incubator. J Assist Reprod Genet 2012;29:565–72.
38. Mio Y, Maeda K. Time-lapse cinematography of dynamic changes occurring during in vitro development of human embryos. Am J Obstet Gynecol 2008; 199:660.e1–5.

SUPPLEMENTAL METHODS

Ovarian Stimulation, Fertilization, and Embryo Culture

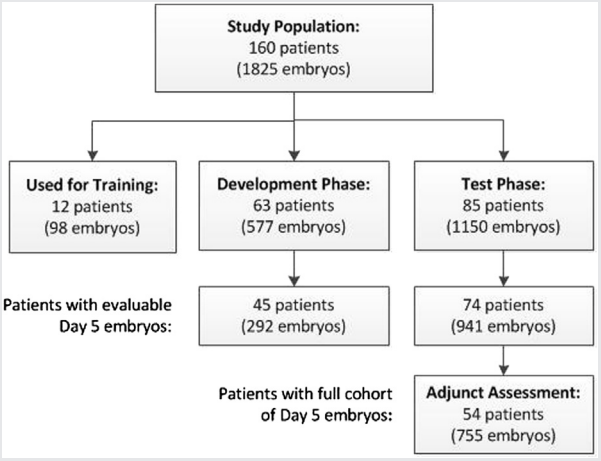
Patients underwent ovarian stimulation according to guidelines of each clinic, where protocols included agonist luteal phase, agonist microdose flare, and antagonist suppression. On the day (D) of oocyte retrieval (D0), oocytes were fertilized using conventional insemination or ICSI. Immediately after the fertilization check, successfully fertilized oocytes (2PNs) were transferred to a multiwell Eeva dish, which is a standard-sized 35-mm diameter petri dish made of conventional tissue culture plastic, with an inner ring containing a precision-molded array of 25 wells (well size 250 μm long $\times$ 250 μm wide $\times$ 100 μm deep). The microwell format holds individual embryos separately but in close proximity under a shared media droplet (40 μL overlaid with mineral oil), while reference labels provide visual orientation of each embryo's specific location in the dish array. Individual well tracking is performed under a single optical field of view, which reduces the need for motorized parts, which are often used in imaging systems to individually address and monitor each embryo (13, 24). At the same time, the shared media permit group culture, which has been suggested to promote improved blastocyst formation rates by promoting positive paracrine signaling between embryos (5, 32). Throughout embryo culture, each clinical site used its own laboratory protocols, including their standard culture media, protein supplementation, incubator at 37°C, and incubation environment.

Embryo Imaging and Light Exposure

Eeva was designed to record embryo development with minimal light exposure to embryos from a light-emitting diode at 625 nm wavelength. Using an optical power meter, it was determined that the power of the illuminating LED light of the Eeva microscope is ~0.6 mW/cm^2. By comparison, the power of a typical IVF inverted microscope (measured on the Olympus IX-71 and CK40 Hoffman Modulation Contrast systems) can be up to 10 mW/cm^2. Because Eeva captures a high image frequency (one single-plane image every 5 minutes) at a relatively low light intensity and exposure time (0.6 seconds for each image), Eeva produces only 0.36 mJ/cm^2 of energy per image. This is equivalent to a total energy exposure of only 0.32 J/cm^2 over 3 days of imaging, or approximately 21 seconds total exposure from a traditional IVF bright-field microscope.

Several reports comparing time-lapse imaged human embryos with nonembryos have demonstrated no negative impact on fertilization rate, embryo development, blastocyst formation, and implantation (12, 33–38). In addition, extensive gene expression analysis of mouse embryos that underwent dark-field time-lapse imaging confirmed no effect on embryonic gene expression (12). We confirmed that the blastocyst formation rate of our study (49.9% average, with a range of 16.9%–60.0% across sites) was comparable to the average rate (45.4%) and range (28.0%–60.3%) of blastocyst formation rates reported between 1998 and 2006 (5), suggesting that embryos imaged by Eeva have competence for normal development.

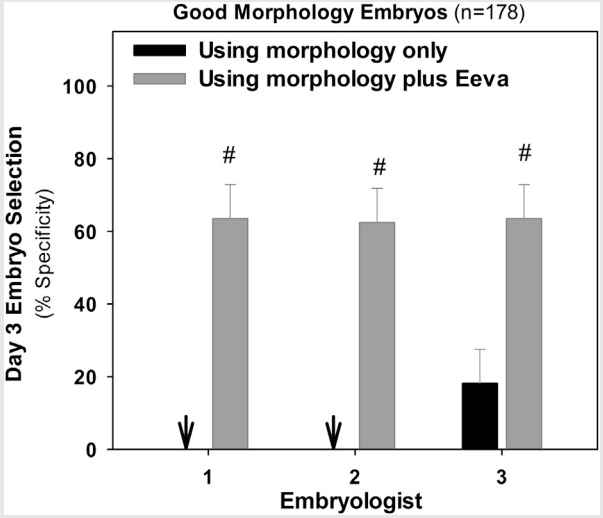
SUPPLEMENTAL FIGURE 1



Study patients and embryos in training, development phase, test phase, and adjunct assessment.

Conaghan. Validation of a time-lapse screening tool. Fertil Steril 2013.

SUPPLEMENTAL FIGURE 2



Alternative analysis of D3 embryo selection by individual embryologists (1, 2, and 3) using morphology only versus morphology plus Eeva for good-morphology embryos (n = 178). In this analysis, good morphology is defined by 7–8 cells, <10% fragmentation, and perfect symmetry. Note that embryologists 1 and 2 were very conservative in their morphology assessments and expected that all D3 good-morphology embryos would become usable blastocysts. $^{\#}P<.0001$.

Conaghan. Validation of a time-lapse screening tool. Fertil Steril 2013.

SUPPLEMENTAL TABLE 1

Clinical characteristics of study patients and embryos in development phase and test phase.

Clinical characteristics	Development phase	Test phase
Total no. of patients	63	85
Total no. of eggs	1,046	1,855
Total no. of 2PNs	577	1,150
Patient demographics (mean ± SD)		
Egg age, y	34.2 ± 4.5	32.2 ± 5.6
Recipient age, y	35.6 ± 4.4	36.2 ± 5.8
Height, in	66.0 ± 2.9	65.6 ± 2.9
Weight, lb	145.1 ± 29.7	146.8 ± 31.0
Cycle type (%)		
Patient using own eggs	58/63 (92.1)	67/85 (78.8)
Oocyte donor	5/63 (7.9)	18/85 (21.2)
Reason for assisted reproductive technology (%)		
Male factor infertility	20/63 (31.8)	16/85 (18.8)
History of endometriosis	3/63 (4.8)	1/85 (1.2)
Ovulation disorders	4/63 (6.4)	9/85 (10.6)
Diminished ovarian reserve	3/63 (4.8)	12/85 (14.1)
Tubal ligation	1/63 (1.6)	0/85 (0.0)
Tubal hydrosalpinx	1/63 (1.6)	0/85 (0.0)
Other tubal disease	1/63 (1.6)	2/85 (2.4)
Uterine	0/63 (0.0)	1/85 (1.2)
Unexplained	11/63 (17.5)	18/85 (21.2)
Multiple reasons	11/63 (17.5)	21/85 (24.7)
Other[a]	8/63 (12.7)	5/85 (5.9)
Stimulation protocol (%)		
Agonist luteal phase	15/63 (23.8)	6/85 (7.1)
Agonist microdose flare	2/63 (3.2)	4/85 (4.7)
Antagonist suppression	29/63 (46.0)	56/85 (65.9)
Other	17/63 (27.0)	19/85 (22.4)
Stimulation and retrieval counts (mean ± SD)		
AFC	16.9 ± 7.0	21.9 ± 9.3
No. of follicles	16.7 ± 7.8	21.0 ± 7.4
No. of eggs	16.6 ± 7.3	21.8 ± 7.7
Method of insemination (%)		
ICSI	39/63 (61.9)	58/85 (68.2)
IVF	21/63 (33.3)	25/85 (29.4)
Both	3/63 (4.8)	2/85 (2.4)
Fertilization count (mean ± SD): no. of 2PNs	9.6 ± 4.7	13.6 ± 4.9

[a] "Other" includes age-related subfertility (n = 3); oligoovulation (n = 2); single woman (n = 2); amenorrhea (n = 1); menopause (n = 1); recurrent pregnancy loss (n = 1); and tubal adhesions (n = 1).

Conaghan. Validation of a time-lapse screening tool. Fertil Steril 2013.

SUPPLEMENTAL TABLE 2

Usable blastocyst prediction performance (% specificity, % sensitivity, % PPV) of Eeva compared with D3 morphology for two independent data sets in the development and test phases.

	No. of patients	No. of embryos	Specificity, %	Sensitivity, %	PPV, %	NPV, %
Morphology (development phase)	45	292	52.1	81.8	34.5	90.9
Eeva (development phase)	45	292	84.2	58.8	54.1	86.6
Eeva (test phase)	74	941	84.7	38.0	54.7	73.7

Conaghan. Validation of a time-lapse screening tool. Fertil Steril 2013.