Barrie, A, Homburg, R, McDowell, G, Brown, J, Kingsland, C and Troup, S

Examining the efficacy of six published time-lapse imaging embryo selection algorithms to predict implantation to demonstrate the need for the development of specific, in-house morphokinetic selection algorithms.

Running Title
The efficacy of five published ESAs

Title
Examining the efficacy of five published time lapse imaging embryo selection algorithms to predict implantation to demonstrate the need for the development of specific, in house morphokinetic selection algorithms.

Authors
Barrie, A[a]
Homburg, R.[a]
McDowell, G.[b]
Brown, J.[c]
Kingsland, C.[a]
Troup, S.[a]

Affiliation address
a. The Hewitt Fertility Centre, Liverpool Women's NHS Foundation Trust, Liverpool, L8 7SS, United Kingdom.
b. Manchester Metropolitan University, Minshull House, 47-49 Chorlton Street, Manchester, M1 3FY, United Kingdom
c. Edge Hill University, St Helens Road, Ormskirk, Lancashire, L39 4QP, United Kingdom

Email address
amy.barrie@lwh.nhs.uk
royhomburg@gmail.com
garry.mcdowell@manchester.ac.uk
brownjm@edgehill.ac.uk
charles.kingsland@lwh.nhs.uk
stephen.troup@lwh.nhs.uk

Corresponding author
Barrie, A.
Amy.barrie@lwh.nhs.uk
(0044)1517024122
(0044)7738677680
The Hewitt Fertility Centre
Liverpool Women's NHS Foundation Trust
Crown Street
Liverpool
L8 7SS
United Kingdom

Capsule
A selection of currently available, published embryo selection algorithms are unable to be applied externally and encourage the development of in house, specific embryo selection algorithms.

Abstract

Objective
To study the efficacy of five embryo selection algorithms (ESAs) at predicting implantation when applied to a large, exclusive set of known implantation embryos with the aim of demonstrating the need to develop specific, in-house ESAs.
Design
A retrospective, observational analysis.
Setting
Fertility Treatment Centre
Patients
Nine hundred and eighty embryos derived from 887 treatment cycles performed between September 2014 and September 2015. Patients undergoing treatment by either IVF or ICSI were included. Embryos were cultured using GTL™ (Vitrolife) at 5% O_2, 89% N_2, 6% CO_2, 37°C in EmbryoScope® instruments.
Main outcome Measure
The difference in implantation rates (IR) of the categories of embryo classification in each ESA defined using specificity, sensitivity and positive predictive value, area under the receiver operating characteristic curve (AUC) and likelihood ratio. The differences in implantation rates (IR) in the categories defined by each ESA were also analysed using Fisher's exact and Kruskall Wallis statistical tests.
Results
Each ESA specified time ranges into which embryos must fall to be identified as having the highest potential for the stated end point. The ESAs comprised a variety of observable events including time to pronuclear fading, 2, 3, 4, 5 and 8 cell, cell cycle durations and time to blastulation and blastocyst. When applied to an exclusive cohort of known implantation embryos, the PPV for IR were 42.57%, 38.14%, 44.07%, 38.79%, and 40.45%. The sensitivity was 16.70%, 51.19%, 72.94%, 98.67%, and 62.33%, respectively. Finally, the specificity was 85.90%, 48.09%, 42.12%, 2.65% and 42.62%, respectively. The AUC were 0.535, 0.512, 0.575, 0.546 and 0.583, respectively. There were no significant differences in IR between the categories in four of the five ESAs ($p>0.05$). One of the ESAs resulted in statistically significant differences in the embryo classifications in terms of IR ($p<0.0001$).
Conclusion
The results from the examination of the published ESAs examined highlight the need for the development of in house, patient, treatment and environment specific ESAs. These data suggest that currently available ESAs may not be clinically applicable and lose their diagnostic value when externally applied.

Key Words: morphokinetics, embryo selection algorithm, embryo development

Introduction

Traditional methods for embryo selection have been utilised for over twenty years. There are numerous morphological parameters that are thought to be useful for correct embryo selection; pronuclear morphology (z scoring) [1,2], polar body alignment and appearance[3, 4], appearance of cytoplasm and zona pellucida[5], early cleavage[6, 7], multinucleation[8, 9, 10], and blastomere morphology[11, 12, 13]. Basic embryo grading, including the number of blastomeres, eveness in the size of the blastomeres and the level of fragmentation remains the gold standard for embryo selection. However, using this method in a traditional sense (with a standard bench top incubator) has two limitations; a restricted overview of an embryo's development and the exposure of the embryo to suboptimal temperatures and gas concentrations. With the introduction of time lapse imaging, where an image of each embryo is taken every 10 to 20 minutes, more intricate embryo parameters can be viewed whilst leaving the embryos in an undisturbed environment. As the availability of time lapse technologies increased, attention was first focused on assessing their clinical safety. Once this had been established and the available technologies validated for clinical use[14, 15, 16, 17, 18], research then turned to determining how the time lapse imaging systems could be utilised to increase pregnancy rates through in depth embryo analysis and an undisturbed culture system.

Through both the research that followed and that performed previously, many morphokinetic parameters were identified that correlated with the embryo's ability to create a pregnancy both in humans and animals; the appearance and disappearance of pronuclei and nuclei at each cell stage[3, 19, 20, 21], the length of time between early cytokineses[22, 23, 24, 25, 26, 27, 28, 29, 30] and initiation of blastulation[31]. Further embryological phenomena have been observed using time lapse imaging including the reabsorption of fragments[32], direct cleavage of embryos from one to three cells[33] and reverse cleavage[34]. These phenomena have been shown to affect an embryos implantation potential to varying degrees however, their discovery could lead to more effective embryo selection within a laboratory utilising time lapse technology.

Single embryo parameters, such as those named above, have been linked to embryo viability (see reference 18 for review) and now these parameters have been used to develop embryo selection algorithms (ESAs). These ESAs seek to combine a number of morphokinetic parameters that have been linked to an embryo's viability expressed either as formation of a blastocyst, implantation or a live birth. This study aims to examine the efficacy of five published ESAs for predicting an embryo's viability, expressed as implantation rate (IR), in a clinically applicable setting[21, 27, 31, 35, 36] aiming to demonstrate the need to develop specific, in-house ESAs. Examined ESAs were selected based on their clinical applicability to the test site, assessed superficially prior to analysis.

Materials and Methods

This investigation was a single site, retrospective observational design approved by the North West Research Ethics Committee (ref: 14/NW/1043). All procedures and protocols complied with UK regulation (Human Fertilisation and Embryology Act,

1990, 2008). Data were obtained from 887 treatment cycles between September 2014 and December 2015. Clinical pregnancy was confirmed by the presence of a fetal heartbeat at ultrasound scan at 6 weeks gestation. All treatments included in this analysis were from known implantation embryos i.e. a single embryo transfer or a double embryo transfer where the transfer of two embryos resulted in either a negative test or two fetal heartbeats.

Ovarian Stimulation

Pituitary down regulation was achieved using either a gonadotrophin releasing hormone agonist (buserelin, Suprecur®, Sanofi Aventis, UK) or antagonist (cetrorelix acetate, Cetrotide®, Merck Serono, Germany). Ovarian stimulation was performed using urine derived or recombinant follicle stimulating hormone (Progynova (Bayer, Germany), Fostimon, Merional (IBSA, Switzerland), Menopur® (Ferring Fertility, Switzerland), Gonal f® (Merck Serono). Doses were adjusted based on patient demographic and response. Patients were given 5000IU of subcutaneous hCG (Gonasi® HP, IBSA Pharmaceuticals, Italy) 36 hours prior to oocyte collection. Luteal support was provided using 400mg of progesterone pessaries twice daily (Cyclogest®, Actavis, UK) until the pregnancy test was taken.

Oocyte retrieval and embryology

Ultrasound guided oocyte collection was performed transvaginally under sedation (Diprivan, Fresenius Kabi, USA). Collected oocyte cumulus complexes were cultured in 4 well dishes (Nunc™, Thermo Scientific, USA) with each well containing 0.65ml GIVF™ (Vitrolife, Gothenburg, Sweden) covered with 0.35ml OVOIL™ (Vitrolife) in a standard incubator (Sanyo Multigas MCO 18M). Sperm preparation was performed using a standard gradient separation at 0.3 relative centrifugal force (rcf) for ten minutes (ISolate®, Irvine Scientific, USA) followed by two washes at 0.6rcf for ten minutes using GIVF™. Those oocytes destined for ICSI were prepared using enzymatic (HYASE 10X™, Vitrolife) and mechanical digestion. ICSI was performed approximately four hours following collection after which time all injected oocytes were placed in individual culture drops of GTL™ (Vitrolife) and cultured in the EmbryoScope® (Vitrolife). Those oocytes destined for standard insemination had this performed approximately four hours after collection and replaced into a standard incubator until fertilisation check the following day. Oocytes were then checked for fertilisation approximately 16 to 18 hours post insemination (hpi) and all fertilised oocytes along with all unfertilised metaphase II oocytes were placed in individual culture drops of GTL™ and cultured in the EmbryoScope®. Embryo selection was performed using the national grading scheme[37] along with an internally derived, ESA. This ESA was used as an additive to morphology at the test site with the latter remaining the gold standard. This ESA included three morphokinetic parameters; s2 (time between t3 and t4), cc3 (time between t4 and t5) and t5 with embryos graded in one of eight categories from A+ to D-. Embryo transfer was performed using the highest grade embryo(s) either three or five days post collection depending on the number of good quality embryos the patient had on day three as well as how many were to be transferred. Selected embryos were cultured in EmbryoGlue® (Vitrolife) for 10 to 30 minutes in a standard incubator prior to embryo transfer. All embryos were cultured at 37°C, 6% CO_2, 5% O_2, 89% N_2 throughout.

Analysis of time lapse information
The image interval on the EmbryoScope® was set to 15 minutes with seven focal planes. Images were collected for the duration of culture immediately following ICSI or fertilisation check (for IVF derived embryos) to utilisation. Images were assessed by a single embryologist for morphokinetic parameters described in table 1 with t0 defined as the time of insemination/ injection. Accuracy of annotation was corroborated bu the participation of the embryologist in an internal quality assurance scheme for morphokinetic analysis. Each of the ESAs (table 2) were then retrospectively applied to the same cohort of known implantation embryos.

Statistical Analysis
Positive predictive value, specificity and sensitivity were used to determine the efficacy of each of the ESAs. These methods of measurement were chosen for analysis due to their relationship to validity and predictive power. Positive predictive value was defined as the percentage of embryos creating a fetal heartbeat as well as a favourable ESA outcome. Sensitivity was defined as the ability of the ESA to correctly classify an embryo as viable. Specificity was defined as the ability of the ESA to correctly classify an embryo as non viable.

Each of the test measures were determined using the following calculations:
Positive predictive value = true positives / (true positives + false positives)
Sensitivity = true positives / (true positives + false negatives)
Specificity = true negatives / (true negatives + false positives)

The likelihood ratio was determined using the following calculation:
Likelihood ratio = sensitivity / (1 – specificity)

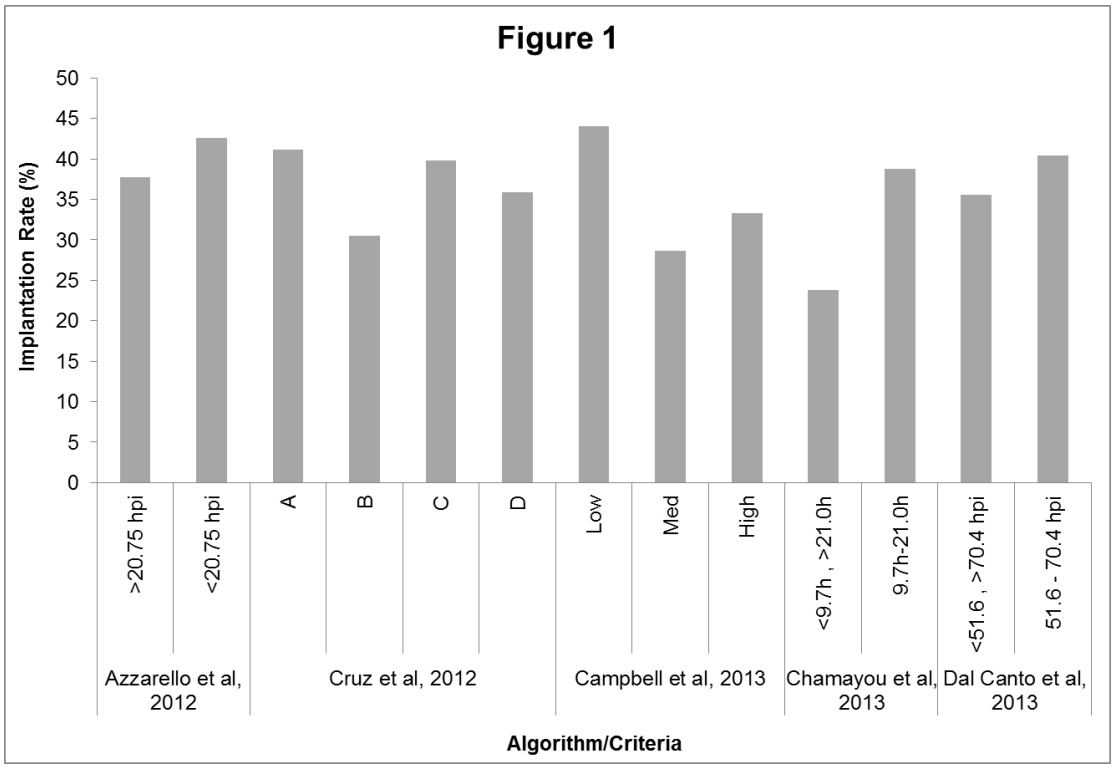
The area under the receiver operating characteristic curve (AUC) was calculated for each ESA. The IR in each category of the ESA was compared using Fisher's exact test (for ESAs with two outcome categories i.e. true, false) and Kruskall Wallis test (for ESAs with more than two outcome categories i.e. A, B, C and D). Results were considered statistically significant at $p<0.05$. Statistical analysis was performed using the statistical package Prism® 5 (GraphPad Software©, USA).

Results

A total of 980 known implantation embryos from 887 treatment cycles were subject to retrospective analysis to determine the efficacy of five published ESAs (table 2). 531 of these embryos were created using conventional IVF while 449 were created using ICSI. The mean patient age was 33.43 ± 4.52 with an average treatment attempt number of 1.37. The primary aetiologies for infertility were male factor (32.2%), maternal age (4.1%), ovulatory disorders (9.9%), tubal disorders (6.6%), uterine disorders (4.1%), other (including genetic disorder) (0.2%), hormonal deficiency (1%) and unexplained (41.8%). Of the 887 treatment cycles, three resulted in a cleavage stage embryo transfer while all other transfers were performed on day five (blastocyst). 93 double embryo transfers and 794 single embryo transfers were performed. 50.2% of treatment cycles were an agonist protocol with the remainder, an antagonist protocol. An overall implantation rate of 39.59% was achieved with 388

of the 980 embryos implanting and 592 not implanting.
The positive predictive value for each of the ESAs did not reach above 45% in any case (table 2). The sensitivity and specificity were considerably more variable (table 2), as would be expected, identifying that one ESA had a high sensitivity[35] and another, a high specificity[21]. All of the other ESAs had sensitivities and specificities close to 50%. The likelihood ratios of all ESAs revealed that there was very little predictive power of implantation where a favourable ESA result is obtained (table 2). Likelihood ratios range from 0 to infinity and a likelihood ratio close or equal to 1 indicates a lack of diagnostic value (McGee, 2002); the furthest from 1 that any of the ESAs in this investigation reached was 0.26 indicating that an embryo has a 0.26 increased chance of creating a pregnancy if a favourable ESA outcome is achieved. Finally, the AUC analysis revealed values from 0.512 to 0.583 (table 2); a further indication of a lack of predictive power of the examined ESAs.

The IR for each category of four of the analysed ESAs did not vary significantly ($p>0.05$) (figure 1). However, the IR for the three categories of the aneuploidy risk classification ESA[31] varied significantly ($p<0.0001$). This ESA also had the strongest likelihood ratio and positive predictive value (44.07%). Incidentally, the number of embryos classified as high risk using this ESA was just three, of which one implanted giving this category an IR of 33.33%; a potentially misleading result. The absolute difference between the IR of low and medium risk embryos was 15.46% (figure 1).



Discussion

All five of the examined ESAs achieved an AUC less than 0.6, indicating reduced predictive capability. None of the ESAs achieved a PPV above 45%, also indicative of reduced diagnostic value. Worthy of note is the ESA that was found to have statistical significance between the categories of embryo classification[31] however, the number of embryos classified as high risk was just three of 980. Further validation, performed by the developers of this ESA[38] using 88 embryos, classified four as high risk. Clearly, using this ESA, the chance of an embryo being classified as high risk is low which raises issues about the specificity of the ESA especially when evidence suggests that over 50% embryos are aneuploid[39]. With an AUC of 0.575 and a 0.26 increased chance that an embryo would create a pregnancy if classified as low risk, this ESA may not represent a robust, clinically applicable embryo selection. Nonetheless, this ESA is the most effective out of the five assessed when a combination of specificity, sensitivity, PPV, AUC, likelihood ratio and differences in implantations between embryo classification category is considered.

The analyses performed indicate that ESAs available in the literature may not provide substantial, additional aid for embryo selection in a clinically relevant setting. The current investigation highlights that externally derived ESAs are developed, inevitably, under conditions different to that of the adoptive centre (table 3) encouraging the development of in house, specific ESAs.. It has been shown that the method by which embryos are created (IVF or ICSI) can affect their temporal behavior[40, 41, 42]. In addition to varying treatment types a number of the analysed ESAs excluded certain patient groups to avoid confounding factors. This includes those with endometriosis, PCOS, severe male factor infertility and maternal age over 39 years. This exclusion constitutes a proportion of patients that make up a significant fraction of patients treated in an IVF laboratory and onto which these ESAs could be critically useful. There is evidence to suggest that the reason for infertility could affect an embryo's morphokinetic profile in particular those with PCOS[43] thus their exclusion in the ESA development is understandable but reduces its clinical applicability unless a specific ESA is developed for this specific patient group. Furthermore, one group's ESA was developed using oocyte donors only, a clear confounder for the application of this ESA in other centres. The majority of the ESAs were developed on embryos created under an agonist protocol. However, one group's ESA development cohort contained a proportion of embryos created under an antagonist protocol[31] The use of agonist and antagonist protocols has yet to be shown to affect an embyro's morphokinetic profile however, they have been linked to embryo quality[44, 45] which could indicate that there is a potential for them to also have a temporal effect. Finally, and perhaps most significantly, varying culture conditions were used in the development of these ESAs. It has been shown that an embryos morphokinetic profile is significantly altered in different culture media specifically between sequential and single step media[46, 47]. This means that those developed using sequential media may not be effective in selecting embryos cultured in single step media, and vice versa. In addition, varying CO_2 and O_2 gas concentrations were used in the development of a number of these published ESAs. Oxygen tension has been specifically linked to an embryo's morphokinetic profile in both humans[48] and mice[49] where those embryos cultured at 20% O_2 have reduced developmental rates

and the completion of the third cell cycle is significantly delayed. These fundamental differences in the development of each ESA need to be seriously considered before their external adoption. It is highly unlikely that an external centre will have the same patient, treatment and environmental parameters as that of the developing centre.

A further consideration for the use of externally derived ESAs is the subjective nature of annotating morphokinetic parameters, the differences in image capture analysis, such as the number of focal planes, and the varying definition of t0. The subjective nature of annotations creates unreliability in the external application of ESAs. There has been some development with this due to the publication of annotation guidelines in 2014[50] however, this will not eliminate the subjectivity completely. Interestingly, there are now two commercially available 'one size fits all' ESAs that, based on the results presented here, should not perform as well as expected. Variations in image acquitision is unlikely to create significant disparity however, coupled with the variability between 'annotaters', an increasing level of inaccuracy could be created. Although undefined in some of the publications, the definition of t0 varies between groups with some using t0 as the time of insemination or injection, the inaugural and arguably the most common method, and others the mid-point of ICSI. It has now been largely accepted that the use of insemination/ injection is arbitrary and the exact moment that the sperm enters the oocyte is indeterminate for IVF cases and, where possible, time of pronuclei fading should be used as t0.

Conclusion

The development of ESAs, thus far, has not involved the control of confounding factors such as media type, patient age and treatment type, except inadvertently by virtue of availability. They are often developed under the environmental parameters available in the laboratory performing the development and thus are clinically relevant in these cases alone. For external application, the ESAs lose their predictive capabilities. The primary objective of ESAs is to allow the selection of the best embryo from a cohort in a clinical setting. Those presented here, clarify that embryo morphokinetics could be used for embryo selection however, they do not offer clinically relevant means to aid in embryo selection in other laboratories unless the development criteria are also adopted. The collective contribution of confounding factors means that derived ESAs can only be applied to that on which they were developed and when applied to a heterogeneous cohort of embryos, as would be found in an IVF laboratory, the capability of the ESA to detect the most viable embryo diminishes. Further research needs to focus on the development of ESAs that are specific to subgroups of patients, environments and treatments. At the very least, embryology laboratories should proceed with caution when implementing ESAs derived from published sources and consider thorough in house validation of such ESAs before clinical use, if at all.

Conflicts of interest: none.
Funding source: none.

References

1. Tesarik, J. and Greco, E. (1999) The probability of abnormal preimplantation development can be predicted by a single static observation on pronuclear stage morphology. Hum Reprod 14: 1318-23
2. Scott, L. (2003) Pronuclear scoring as a predictor of embryo development. Reproductive Biomedicine Online. 6: 201-14
3. Payne, D., Falherty, S.P., Barry, M.F., Matthews, C.D. (1997) Preliminary observations on polar body extrusion and pronuclear *formation* in human oocytes using time lapse video cinematography. Hum Reprod 12: 532-41
4. De Placido, G., Wilding M., Strina, I., Alviggi, E., Alviggi, C., Mollo, A. *et al,* (2002) High outcome predictability after IVF using a combined score for zygote and embryo morphology and growth rate. Hum Reprod 17: 2402-09
5. Palmstierna, M., Murkes, D., Csemiczky, G., Andersson, O., Wramsby, H. (1998) Zona pellucida thickness variation and occurrence of visible mononucleated blastomeres in pre embryos are associated with a high pregnancy rate in IVF treatment. J Assist Reprod Genet 15: 70-5
6. Lundin, K., Bergh, C., Hardarson, T. (2001) Early embryo cleavage is a strong indicator of embryo quality in human IVF. Hum Reprod 16: 2652-7
7. Isiklar, A., Mercan, R., Balaban, B., Alatas, C. Aksoy, S., Urman, B. (2002) Early cleavage of human embryos to the two cell stage. A simple, effective indicator of implantation and pregnancy in intracytoplasmic sperm injection. J Reprod Med 47: 540-4
8. Pickering, S.J., Taylor, A., Johnson, M.H., Braude, P.R. (1995) An analysis of multinucleated blastomere formation in human embryos. Hum Reprod 10: 1912-22
9. Jackson,K.V., Ginsburg, E.S., Hornstein, M.D., Rein, M.S., Clarke, R.N. (1998) Multiniucleation in normally fertilized embryos is associated with an accelerated ovulation induction response and lower implantation and pregnancy rates in in vitro fertilization embryo transfer cycles. Fertil Steril 70: 60-6
10. Yakin, K., Balaban, B., Urman, B. (2005) Impact of the presence of one or more multinucleated blastomeres on the developmental potential of the embryo to the blastocyst stage. Fertil Steril 83: 243-5
11. Shapiro, B.S., Harris, D.C., Richter, K.S. (2000) Predictive value of 72 hour blastomere cell number on blastocyst development and success of subsequent transfer based on the degree of blastocyst development. Fertil Steril 73: 582-6
12. Hardarson, T., Hanson, C. Sjogren, A., Lundin, K. (2001) Human embryos with unevenly sized blastomeres have lower pregnancy and implantation rates: indications for aneuploidy and multinucleation. Hum Reprod 16: 313-8
13. Johansson, M., Hardarson, T., Lundin, K. (2003) There is a cutoff limit in diameter between a blastomere and a small anucleate fragment. J Assist Reprod Genet 20: 309-13
14. Freour, T., Lammers, J., Splingart, C. Jean, M. Barriere, P. (2010) Time lapse (EmbryoScope™) as a routine technique in the IVF laboratory: A useful tool for better embryo selection? Gynecol Obstet Fertil 40: 476-80

15. Nakahara, T., Iwase, A., Goto, M., Harata, T., Suzuki, M., Ienaga, M. *et al* (2010) Evaluation of the safety of time lapse observations for human embryos. J Assist Reprod Genet 27: 93-6
16. Basile, N., Morbeck, D., Garcis-Velasco, J., Bronet, F., Meseguer, M. (2011) Time lapse technology reveals that embryo kinetics are not affected by culture media. Fertil Steril 96 (Suppl): S108
17. Cruz, M., Gadea, B., Garrido, N., Pedersen, K.S., Martinez, M., Perez-Cano, I. *et al* (2011) Embryo quality, blastocyst and ongoing pregnancy rates in oocyte donation patients whose embryos were monitored by time lapse imaging. J Assist Reprod Genet 28: 569-73
18. Kirkegaard, K., Hindkjaer, J.J., Grondahl, M.L., Kesmodel, U.S., Ingerslev, H.J. (2012) A randomized clinical trial comparing embryo culture in a conventional incubator with a time lapse incubator. J Assist Reprod Genet 29: 565-72
19. Lemmen, J.G., Agerholm, I., Ziebe, S. (2008) Kinetic markers of human embryo quality using time lapse recordings of IVF/ICSI ferilized oocytes. Reproductive Biomedicine Online 17: 385-91
20. Scott, L. (2010) The origin and consequences of day 2 multinucleation of human embryos. Abstracts of the 26th annual meeting of ESHRE, Rome, Italy, 25: P-204
21. Azzarello, A., Hoest, T., Mikkelsen, A.L. (2012) The impact of pronuclei morphology and dynamicity on live birth outcome after time lapse culture. Hum Reprod 27: 2649-57
22. Gonzales, D.S., Pinheiro, J.C., Bavister, B.D. (1995) Prediction of the developmental potential of hamster embryos in vitro by precise timing of the third cell cycle. J Reprod Fertil 105: 1-8
23. Ramsing, N.B. and Cellesen, H. (2006) Detecting timing and duration of cell divisions by automatic image analysis may improve selection of viable embryos. Fertil Steril 86 (Suppl): S189
24. Ramsing, N.B., Berntsen, J., Callesen, H. (2007) Automated detection of cell division and movement in time lapse images of developing bovine embryos can improve selection of viable embryos. Fertil Steril 88 (Suppl 1): S38
25. Lechniak, D., Pers-Kamczyc, E., Pawlak, P. (2008) Timing of the first zygotic cleavage as a marker of developmental potential of mammalian embryos. Reprod Biol 8: 23-42
26. Herrero, J., Tejera, A., Ramsing, N., Romero, J.L., Rubio I. (2011) Establishing the optimal time ranges of key events during development using time lapse video cinematography. Fertil Steril 96(Suppl): S102
27. Cruz, M., Garrido, N., Herrero, J., Perez-Cano, I., Munoz, M., Meseguer, M. (2012) Timing of cell divisions in human cleavage stage embryos correlates with blastocyst formation and quality. Reproductive Biomedicine Online 25: 371-81
28. Hlinka, D., Kalatova, B., Uhrinova, I., Dolinska, S., Rutarova, J., Rezacova, J. (2012) Time lapse cleavage rating predicts human embryo viability. Physiol Res 61: 513-25

29. Meseguer, M., Rubio, I., Cruz, M., Basile, N., Marcos, J., Requena, A. (2012a) Embryo incubation and selection in a time lapse monitoring system improves pregnancy outcome compared with a standard incubator: a retrospective cohort study. Fertil Steril 98: 1481-9
30. Cruz, M., Garrido, N., Herrero, J., Perez-Cano, I., Munoz, M., Meseguer, M. (2012) Timing of cell division in human cleavage stage embryos is linked with blastocyst formation and quality. Reproductive Biomedicine Online 25: 371-81
31. Campbell, A., Fishel, S., Bowman, N., Duffy, S., Sedler, M., Hickman, C.F. (2013a) Modelling a risk classification of aneuploidy in human embryos using non invasive morphokinetics. Reproductive BioMedicine Online 26: 477-85
32. Hardarson, T., Lofman, C., Coull, G., Sjogren, A., Hamberger, L., Edwards, R.G. (2002) Internalization of cellular fragments in a human embryo: time lapse recordings. Reproductive Biomedicine Online 5: 36-8
33. Rubio, I., Kuhlmann, R., Agerholm, I., Kirk, J., Herrero, J., Escriba, M.J. *et al* (2012) Limited implantation success of direct cleaved human zygotes: a time lapse study. Fertil Steril 98: 1458-63
34. Liu, Y., Chapple, V., Roberts, P., Matson, P. (2014) Prevalence, consequence, and significant of reverse cleavage by human embryos viewed with the use of the EmbryoScope time lapse video system. Fertil Steril 102: 1295-1300
35. Chamayou, S., Patrizio, P., Storaci, G., Tomaselli, V., Alecci, C., Ragolia, C. *et al* (2013) The use of morphokinetic parameters to select all embryos with full capacity to implant. J Assist Reprod Genet 30: 703-10
36. Dal Canto, M., Coticchio, G., Mignini Renzini, M., De Ponti, E., Novara, P.V., Brambillasca, F. *et al* (2012) Cleavage kinetics analysis of human embryos predicts development to blastocyst and implantation. Reproductive BioMedicine Online 25: 474-80
37. Cutting R., Morroll, D., Roberts, S.A., Pickering, S., Rutherford, A., BFS and ACE (2008) Elective single embryo transfer: guidelines for practice British Fertility Society and Association of Clinical Embryologists. Human Fertility 11: 131-46
38. Campbell, A., Fishel, S., Bowman, N., Duffy, S., Sedler, M., Thronton, S. (2013b) Retrospective abalysis of outcomes after IVF using an aneuploidy risk model derived from time lapse imaging without PGS. Reproductive BioMedicine Online 27: 140-6
39. Fragouli, E. and Wells, D. (2011) Aneuploidy in the human blastocyst. Cytogenetic and Genome Research 133: 149–59.
40. Cruz, M., Garrido, N., Gadea, B., Munoz, M., Perez-Cano, I., Meseguer, M. (2013) Oocyte insemination techniques are related to alterations of embryo developmental timing in an oocyte donation model. Reproductive BioMedicine Online 27: 367-75
41. Bodri, D., Sugimoto, T., Serna, J.Y., Kondo, M., Kato, R., Kawachiya, S. *et al* (2015) Influence of different oocyte insemination techniques on early and late morphokinetic parameters: retrospective analysis of 500 time lapse monitored blastocysts. Fertil Steril 104: 1175-81
42. Liu, Y., Chapple, V., Feenan, K., Roberts, P., Matzon, P. (2015) Time lapse videography of human embryos: using pronuclear fading rather than

insemination in IVF and ICSI cycles removes inconsistencies in time to reach early cleavage milestones. Reproductive Biology 15: 122-5

43. Wissing, M.L., Bjerge, M.R. Olesen, A.I., Hoest, T., Mikkelsen, A.L. (2014) Impact of PCOS on early embryo cleavage kinetics. Reproductive BioMedicine Online 28: 508-14
44. Murber, A., Fancsovits, P., Ledo, N., Gilan, Z.T., Rigo, J., Urbancsek, J. (2009) Impact of GnRH analogues on oocyte/embryo quality and embryo development in in vitro fertilization/intracytoplasmic sperm injection cycles: a case control study. Reproductive Biology and Endocrinology 25: 103
45. Vengetesh, P.M., Ramachandran, A., Kumar, P. (2015) Choosing GnRH Antagonist protocol shows improved oocyte and embryo quality, coherent with perifollicular vascularity (PFV) in assisted reproductive techniques. Journal of Clinical and Diagnostic Research 9: QC24-8
46. Ciray, H.N., Aksoy, T., Goktas, C., Ozturk, B., Bahceci, M. (2012) Time lapse evaluation of human embryo development in single versus sequential culture media – a sibling oocyte study. J. Assist. Reprod. Genet. 29: 891-900
47. Barrie, A., Taylor, E., Schnauffer, K., Kingsland, C., Troup, S. (2015) An examination of embryo morphokinetics and utilisation in single step and sequential culture media systems. Hum Reprod 30 (Suppl 1): P-282
48. Kirkegaard, K., Hindkjaer, J.J. Ingerslev, H.J. (2013) Effect of oxygen concentration on human embryo development evaluated by time lapse monitoring. Fertil Steril 99: 738-44
49. Wale, P.L. and Gardner, D.K. (2010) Time lapse analysis of mouse embryo development in oxygen gradients. Reproductive BioMedicine Online 21: 402-10
50. Ciray, H.N., Campbell, A., Agerholm, I.E., Aguilar, J., Chamayou, S., Esbert, M. *et al* (2014) Proposed guidelines on the nomenclature and annotation of dynamic human embryo monitoring by a time lapse user group. Hum Reprod 29: 2650-60

Figure 1 caption

Implantation rates (IRs) of the embryo classification categories in each of the analysed ESAs.

Azzarello *et al*, 2012; IR of those embryos where pronuclear fading (PNf) occurred after 20.75hpi (n=832, 37.74%) and those that faded before 20.75hpi (n=148, 42.57%) (p>0.05, Fisher's exact test). Cruz *et al*, 2012; IR of embryos classified as A (t5 = 48.8-56.6hpi and s2 ≤0.76h, n=365), B (t5 <48.8 OR >56.6 and s2 ≤0.76h, n=141), C (t5 = 48.8-56.6hpi and s2 >0.76h, n=354) and D (t5 <48.8 OR >56.6 and s2 >0.76h, n=120) with respective IR of 41.1%, 30.5%, 39.83% and 35.83% (p>0.05, Kruskal-Wallis test). Campbell *et al*, 2013; IR for embryos classified as low risk (tSB <92.2hpi and tB <122.9hpi, n=624), medium risk (tSB ≥96.2 and tB ≤122.9hpi, n=353) and high risk (tB ≥122.9hpi, n=3) with respective IR of 44.07%, 28.61% and 33.33% (p<0.05, Kruskall-Wallis test). Chamayou *et al*, 2013; IR of those embryos where cc3 (t5-t3) occurred between 9.7-21h (n=959, 23.81%) and those that did not (n=21, 38.79%) (p>0.05, Fisher's exact test). Dal Canto *et al*, 2012; IR of embryos where t8 occurred between 51.6-70.4hpi (n=581, 35.59%) and those that did not (n=399, 40.45%) (p>0.05, Fisher's exact test).

Annotated morphokinetic parameter	Definition
tPNf	Time when both pronuclei had faded
t2	Time to 2-cell
t3	Time to 3-cell
t4	Time to 4-cell
t5	Time to 5-cell
t8	Time to 8-cell
tSB	Time to start of blastulation; when first signs of a cavity were visible
tB	Time to full blastocyst; when the blastocoele filled the embryo with <10% increase in it's diameter
Calculated morphokinetic parameters	
s2	Time of synchrony of second cell cycle (t4-t3)
cc3	Time of third cell cycle (t5-t3)

Table 1; summary of morphokinetic parameters used for analysis including those requiring annotation as well as those requiring calculation from the annotated values.

	Model type	Parameter	Time frame	Sensitivity (%)	Specificity (%)	PPV (%)	AUC	Likelihood ratio	Category analysis
Azarello *et al*, 2012	Selection/ deselection	PNf	>20h 45m	16.70	85.90	42.57	0.535	1.19	0.2724 (Fisher's exact)
Cruz *et al*, 2012	Hierarchical	t5 s2	48.8-56.6h ≤0.76h	51.19	48.09	38.14	0.512	0.99	0.1402 (Kruskal-Wallis)
Campbell *et al*, 2013	Risk classification model	tSB tB	*Low risk*: tSB <96.2hpi tB <122.9hpi *Medium risk*: tSB ≥96.2hpi tB <122.9hpi *High risk*: tB ≥122.9hpi	72.94	42.12	44.07	0.575	1.26	<0.0001**** (Kruskall-Wallis)
Chamayou *et al*, 2013	Selection/ deselection	cc3	9.7-21h	98.67	2.65	38.79	0.546	1.01	0.1817 (Fisher's exact)
Dal Canto *et al*, 2012	Selection/ deselection	t8	54.9 ± 5.2h	62.33	42.62	40.45	0.583	1.09	0.1415 (Fisher's exact)

Table 2; summary of embryo selection criterion and main results.

PNf; pronuclear fading. t5; time to 5-cell. s2; time between 3-cell and 4-cell. tSB; time to start of blastulation. tB; time to full blastocyst. cc3; time between 3-cell and 5-cell. t8; time to 8-cell. PPV; positive predictive value. AUC; area under the receiver operating characteristic curve.

	n (embryos)	n (cycles)	Fertilisation method	End point	Exclusion criteria	Inclusion criteria	Image capture interval (mins)	Protocol	Culture	Media change	Transfer day
Azarello *et al*, 2012	159	130	ICSI	LBR	-	Embryos transferred at 4-cell stage with equal blastomeres and <25% fragmentation, autologous gametes, female age ≤39, male factor infertility (1-5x10^5 motile sperm/ ejaculate)	20	Agonist	Cook® 5.5% CO_2, 5% O_2, 89.5% N_2	No	2 (44hpi)
Cruz *et al*, 2012	834	165	ICSI	BFR	-	Oocyte donor meeting all required criteria for donation programme	20	Agonist	Global IVF medium (LifeGlobal) 6% CO_2, 21% O_2, 37.4°C	Yes	5
Campbell *et al*, 2013	88	25	ICSI	CPR and LBR	-	Patients undergoing a cycle inclusive of PGS	20	Agonist (75%) Antagonist (25%)	Global IVF medium (LifeGlobal) 5.5% CO_2, 5% O_2, 89.5% N_2	Yes	5
Chamayou *et al*, 2013	178	78	ICSI	BFR	Severe endometriosis, premature ovarian failure, severe asthenoteratozoospermia	Fresh gametes	20	Agonist	Quinn's Advantage (SAGE) 5% CO_2, 5% O_2	Yes	5
Dal Canto *et al*, 2012	134	71	IVF (22) and ICSI (49)	IR	-	Indication for standard IVF or ICSI due to male factor, tubal factor, stage I or II endometriosis or PCOS, maternal age 27-42.	20	Agonist	ISM1 (day 1-3) BlastAssist (day 3-5) 6% CO_2, 5% O_2, 89% N_2	Yes	3 and 5

Table 3; summary of publications used for examination of efficacy of selection criteria. ICSI; intracytoplasmic sperm injection. IVF; in vitro fertilisation. LBR; live birth rate. BFR; blastocyst formation rate. CPR; clinical pregnancy rate. IR; implantation rate. PGS; preimplantation genetic screening. PCOS; polycystic ovary syndrome. CO_2; carbon dioxide. O_2; oxygen. N_2; nitrogen. hpi; hours post insemination.