Diagnosis of human preimplantation embryo viability

David K. Gardner¹,* , Marcos Meseguer², Carmen Rubio³, and Nathan R. Treff⁴

¹School of BioSciences, University of Melbourne, Parkville, VIC 3010, Australia ²Laboratorio Fiv, IVI Valencia, Plaza Policı ´a Local, 3, Valencia 46015, Spain ³IGENOMIX and Fundación Instituto Valenciano de Infertilidad (FIVI)/INCLIVA, Valencia, Spain ⁴Reproductive Medicine Associates of New Jersey, Basking Ridge, NJ 07960, USA

*Correspondence address. E-mail: david.gardner@unimelb.edu.au

Submitted on June 15, 2014; resubmitted on November 2, 2014; accepted on December 8, 2014

TABLE OF CONTENTS

- Introduction
- Methods
- Morphokinetics/time-lapse imaging
- Comprehensive chromosome screening
  - CGH arrays
  - SNP arrays, qPCR and NGS
- Analysis of physiology and the search for biomarkers
  - Proteomics
  - Metabolomics
  - Targeted metabolomics
  - Glucose utilization
  - Amino acid utilization
- Laboratory on a chip
- Conclusions

BACKGROUND: Transfer of more than a single embryo in an IVF cycle comes with the finite possibility of a multiple gestation. Even a twin pregnancy confers significant risk to both mother and babies. The move to single-embryo transfer for all patients will be greatly facilitated by the ability to quantify embryo viability. Developments in time-lapse incubation systems have provided new insights into the developmental kinetics of the human preimplantation embryo. Advances in molecular methods of chromosomal analysis have created platforms for highly effective screening of biopsied embryos, while noninvasive analysis of embryo physiology reveals more about the embryo than can be determined by morphology alone.

METHODS: Recent developments in time-lapse microscopy, molecular karyotyping and in proteomics and metabolomics have been assessed and presented here in a descriptive review.

RESULTS AND CONCLUSIONS: New algorithms are being created for embryo selection based on their developmental kinetics in culture, and the impact of factors such as patient etiology and treatment are being clarified. Potential links between morphokinetic data and embryo karyotype are being elucidated. The introduction of new molecular methods of determining embryo chromosomal complement is proving to be accurate and reproducible, with the future trending toward CGH arrays or next generation sequencing as a rapid and reliable means of analysis, that should be suitable for each IVF clinic to adopt. A relationship between embryo metabolism and viability is established and is now being considered together with morphokinetic data to create more robust algorithms for embryo selection. Microfluidic devices have the capacity and potential to be used in human IVF clinics for the routine diagnosis of embryo biomarkers.

Key words: chromosome / metabolism / selection / time lapse / embryo viability
Introduction

The goal of assisted reproduction in the human is the delivery of a healthy singleton child through the transfer of an individual euploid embryo. With major advances in embryo culture and cryopreservation over the past 15 years, resulting in significant increases in embryo implantation rates and in cumulative pregnancy rates per retrieval, we have witnessed a significant reduction in the numbers of embryos being transferred, with several countries undertaking routine single-embryo transfer (SET) and confining high-order multiple gestations to the archives of human IVF. However, unless diagnosis of embryo viability is applied for embryo selection for transfer, we condemn patients to successive cycles of retrievals, cryopreservation and replacements. Our objective should be to assist couples to attain a healthy family sooner than later, not sooner or later.

A greater understanding of embryo developmental kinetics and physiology will facilitate the identification of the most viable embryo for transfer. However, even if we can identify those embryos within a cohort most able to implant, one cannot guarantee their ploidy, nor in cases of familial genetic diseases, can we select the unaffected conceptus. Rapid advances in genetic screening technologies are now improving our ability to screen embryonic cells, to ensure that transfer of a healthy euploid embryo becomes routine. Of significance, relationships between the kinetics of embryo development, proteomics and metabolism with the chromosomal status of the embryo are emerging. Consequently, in those clinics or countries where detailed genetic analysis is not feasible, there is at least the potential to decrease the possibility of transferring an aneuploid embryo through a more detailed analysis of embryonic developmental history.

Mastenbroek et al. (2011a, b) fashioned a challenging argument that ‘the path of embryo selection is turning into a dead-end in the quest for optimal IVF success rates’. Furthermore, it was proposed that ‘Embryo selection techniques will not only be unable to improve live birth rates but also they could even lower the success rate of IVF’, and that in the ‘perfect world’ the best that embryo selection can do is shorten the time to pregnancy. Even if the only reason to perform embryo selection is to decrease the time it takes for a couple to conceive, then that in itself is a very good reason. Patients do not consider a reduction in the time to achieve a pregnancy as nonconsequential. Furthermore, reducing the number of transfers a patient must undergo also carries with it a reduction in financial and emotional pain for the patient. Every time a patient undergoes a transfer and subsequent pregnancy test, they experience a dramatic rollercoaster of emotions that creates an unparalleled toll on their ‘infertility’ psychology. Routine IVF is already believed to be associated with a 17% drop out rate (Verberg et al., 2008). However, without the ability to screen the embryos of patients for genetic and chromosomal abnormalities, patients will endure either countless cycles of miscarriage, or the challenges associated with a subsequent amniocentesis. Here we discuss how developments in the fields of time-lapse microscopy (mophokinetic analysis), genetic diagnosis, and in proteomics and metabolomics could facilitate the identification and transfer of an individual viable euploid embryo, thereby reducing the time to a healthy pregnancy for our patients.

Methods

The available evidence regarding the clinical use of time-lapse microscopy, molecular karyotyping and of proteomics and metabolomics has been assessed and presented here. We consider the use and implementation of new technologies and their potential impact on facilitating SET and their ability to increase the take-home baby rate. No systematic review has been carried out.

Morphokinetics/time-lapse imaging

Morphological evaluation has been the method of embryo selection for over 30 years and remains the primary approach of embryo assessment during IVF cycles. However, this evaluation method poses limitations, not only arising from the subjectivity of the embryologist, but also because of the evaluation system itself, which views embryo development statically. Current embryo evaluation is typically based on the morphological appearance of embryos at just a few discrete time points, which limits the frequency of observations made thereby concealing what happens during the intervals between observations (Cruz et al., 2012). In addition, it is important to bear in mind that the status of an embryo (and thus its grading) can change markedly within just a few hours (Mio and Maeda, 2008; Meseguer et al., 2011). These limitations can be overcome with time-lapse systems by which one cannot only analyze embryo morphology but also provide information about the dynamic changes during the preimplantation period. The key advantages of time-lapse technology include the provision of an improved culture environment based on embryo evaluation without removal from the incubator, and the objective and accurate information obtained, not only quantitatively, but also qualitatively. Time-lapse observations have also been useful in defining new or poorly described concepts of embryology such as the duration of the first three cell cycles (Lemmen et al., 2008; Mio and Maeda, 2008; Meseguer et al., 2011) and furthermore, redefining traditionally used morphology parameters used for embryo selection such as pronuclear score (Azzarello et al., 2012; Aguilar et al., 2014).

Here the most relevant studies performed on time-lapse in the last few years are considered; Lemmen et al. (2008) reported that embryos resulting in pregnancies displayed not only a significantly higher cleavage synchrony but also a higher synchrony in nuclear appearance at the two-blastomere stage compared with nonimplanting embryos. Wong et al. (2010) suggested that embryos that reached the blastocyst stage could be predicted by the timing of early developmental stages: first cytokinesis 0–33 min, time interval between first and second mitoses of 7.8–14.3 h, and time interval between second and third mitoses of 0–5.8 h. Finally, it was observed that in embryos which did not form a blastocyst, fragmentation was rarely reversed, whereas it could be in those reaching the blastocyst stage. Using morphokinetics it has been possible to demonstrate associations between various cleavage stage kinetic parameters and the ability of the embryos to reach the blastocyst stage (Cruz et al., 2012). One of the largest data sets on transferred ICSI-generated embryos analyzed by time-lapse was presented by Meseguer et al. (2011), where results obtained were used to design a hierarchical classification for the selection of embryos with the highest implantation potential. The timing of the cleavage to 5 cells (tS) had the best correlation with implantation success. The algorithm started with a morphological screening of the embryos to discard nonviable embryos. The next step was to exclude embryos fulfilling at least one of the following exclusion criteria: (i) uneven blastomere size after the first cleavage; (ii) abrupt division from one to three cells or (iii) multinucleation at the...
4-cell stage. After evaluation for these criteria, the remaining levels of the hierarchy were defined according to the exact timings of some of the variables (t5, the duration of the transition from a 2-cell to a 4-cell embryos (s2) and the duration of the second cell cycle (cc2)). Using this algorithm and logistic regression to analyze the predictive abilities of time lapse, a comparison using a receiver operating characteristic curve (ROC) was made with standard morphology which gave an area under the curve (AUC) of 0.64, whereas a logistic regression on the time-lapse selection gave an AUC of 0.72. The higher AUC supports the possibility of improved embryo selection using time-lapse. In the largest retrospective study on pregnancy outcomes compared with conventional incubation there was a relative 20% improvement in pregnancy rates in the time-lapse system (Meseguer et al., 2012). These results have since been confirmed in a prospective RCT (Rubio et al., 2014). Data reflect improvements of the clinical outcome as a consequence of both improved culture conditions and a better embryo selection process by using a pre-defined morphokinetic algorithm.

Once methods for embryo selection by using time-lapse were established, issues arose of whether there could be universal application of these algorithms, or whether different culture conditions may affect their efficacy and restrict their use only to those IVF units where the models were generated. Basile et al. (2013) compared different culture media, by a RCT on sibling oocytes, and observed identical timings and proportion of optimal embryos for the two media analyzed. However, in contrast, Ciray et al. (2012) reported different timings with different culture media using a similar experimental design. Furthermore, oxygen has been shown to have profound effects on mammalian preimplantation embryo development to clinical outcomes (Wale and Gardner, 2010; Kirkegaard et al., 2013a). Consequently, it is important that both media type and oxygen concentration be reported in clinical studies, as both can affect the timing of key developmental events.

Chamayou and colleagues defined two sets of morphokinetic parameters (MKP). Firstly, they established optimal ranges for some parameters which were considered predictive of the embryo’s ability to become a viable blastocyst on Day 5 (D5, time of pronuclei formation; t1, time of 2, 4, 7 and 8 cell cleavage, t(n); t2, t4, t7, t8) (Chamayou et al., 2006). They then suggested that cc3 was the MKP significantly associated with implantation and a viable pregnancy (t5-t3, cc3). With these results they suggested D3 embryo transfer for those embryos with MKP outside optimal ranges. Analysis of initial events of development revealed that PN breakdown occurred significantly later in embryos resulting in live birth and never earlier than 20 h 45 min (Azzarello et al., 2012), while timings at which second polar body extrusion, pronuclear fading and length of S-phase occurred were linked successfully to embryo implantation (Aguilar et al., 2014).

Potential inter- and intra-observer variability may impact time-lapse marker interpretation, similar to what has been found with manual embryo morphology grading (Chen et al., 2013). Of note, a prospective study has recently demonstrated the ability of a noninvasive, computer-automated test to improve the prediction of blastocyst formation by D3 and provide clinical value to embryologists for D3 selection without any manual analysis (Conaghan et al., 2013). The test, termed Eeva (early embryo viability assessment), is facilitated by dark-field microscopy and cell-tracking software algorithms (based on P2 (t3-t2) and P3 (t4-t3) as relevant parameters categorizing embryos into groups with either high or low likelihood of forming ‘usable blastocysts’). Adding Eeva to traditional D3 morphology significantly improved the ability of experienced embryologists to predict usable blastocyst outcomes and reduced the variability among embryologists (Conaghan et al., 2013). Eeva significantly improved the specificity (84.7 versus 52.1%; \( P < 0.0001 \)) and the positive predictive value (54.7 versus 34.5%; \( P < 0.0001 \)) of usable blastocyst predictions compared with morphology evaluation alone (Conaghan et al., 2013). Subsequently, the transferability of this model in four different clinics has been demonstrated retrospectively (Kirkegaard et al., 2014). Even though results obtained showed a 30% higher implantation rate between the entire cohort and embryos categorized as usable, 50.6% of the embryos that resulted in pregnancy were categorized as nonusable according to the algorithm. Consequently, these data highlight the importance of increasing implantation rates with the application of a model, but with the need to bear in mind the importance of a low rejection rate of viable embryos. A further study has since confirmed cases where poor conventional morphological score and/or suboptimal morphokinetic embryos have resulted in live births, highlighting the risk of discarding so-called ‘viable embryos’ (Stecher et al., 2014).

The transfer of blastocysts yields higher implantation rates than transfers at the cleavage stage, but this outcome must be balanced against the possible disadvantages of longer culture such as the risk of canceled cycles and concerns over the possible epigenetic effects of prolonged in vitro culture (Kirkegaard et al., 2012a). An analysis of blastocyst transfer combined with time-lapse analysis found that human embryo cleavage rates are suggestive of their ability to develop to the blastocyst stage and to implant (Dal Canto et al., 2012). It was concluded that cleavage from the 2- to 8-cell stage occurs progressively earlier in embryos with the ability to develop to blastocysts and to implant. In this line of research, Kirkegaard et al. (2013b) suggested that high-quality blastocysts could be predicted within the first 48 h of culture by a short duration of the first cytokinesis, duration of the 3-cell stage and absence of direct cleavage to 3 cells. The duration of the 2-cell stage either in Kirkegaard et al. (2012a, b) or the study by Cruz et al. (2012) is not found to differ between high- and low-quality embryos, in contrast to the findings by Wong et al. (2010). Further studies may reveal the correlation of key developmental parameters with subsequent blastocyst score. However, a caveat here is that human embryos can have the same embryo grade and yet differ with respect to their proteome and metabolism (discussed further on), and that transfer of the cleavage stage embryo to the uterus is asynchronous, which results in compromised fetal development in animal models (Barnes, 2000; Walker et al., 2015).

There are several studies that have correlated time-lapse with probability of selecting chromosomally normal embryos (Montag, 2013; Swain, 2013). Davies et al. (2012) observed that embryos with complex aneuploidies presented delays on the first two cleavages as well as prolonged transitions between 2 and 4 cells. They also observed irregular divisions and an asynchronous PN disappearance in abnormal embryos. Chavez’s group observed precise timings in all euploid embryos to the 4-cell stage, whereas only 30% of aneuploid embryos presented parameter values within normal timing windows (Chavez et al., 2012). Campbell and colleagues went on to develop a model to categorize the risk of aneuploidy using time-lapse. They found no differences in the early stages, but defined as parameters relevant to ploidy the initiation of blastulation and the timing of the formation of a full blastocyst (Campbell et al., 2013a, b). Similarly, it has been reported that there is a significant correlation between early times to cavitation, from the first cytokinesis and from 5 cells, and reduced prevalence of aneuploidy.
(Hong et al., 2013). Basile and colleagues studied the differences in the cleavage time between chromosomally normal and abnormal embryos in order to elaborate an algorithm to increase the probability of noninvasively selecting chromosomally normal embryos. A different kinetic behavior was observed, being able to identify, using a logistic regression analysis, t5-t2 and cc3 (t5-t3) as the most relevant parameters related to normal chromosomal content (Basile et al., 2014). However, in a retrospective analysis Kramer et al. (2014) have been unable to segregate aneuploid from euploid embryos using MKP. Consequently, selection of embryos through time-lapse technology should not be considered as a replacement for preimplantation genetic screening (PGS). As reported by Basile et al. a ROC analysis to determine the predictive properties of morphokinetics with respect to chromosomal normality gave an AUC value of 0.634 (95% confidence interval, 0.581—0.687), a value which is below what we should expect from a diagnostic tool that is able to forecast an abnormal embryo (with an abnormal chromosome content) (Basile et al., 2014). However, it does represent an excellent selection tool for good prognosis patients who are not indicated for PGS, or for patients who are indicated but who for any legal, social, or economic reasons do not wish to, or cannot, have PGS performed. A summary of the papers published in time-lapse research are summarized in Table I and the graphic representation of the morphokinetic models developed by diverse groups and using different outcomes are shown in Figure 1.

In conclusion, we have tried to convey the issues around the use of time-lapse technology as a clinical tool. We are conscious that standardization in embryo annotation and nomenclature is necessary in order to allow a more accurate interpretation of the data available, and that the existing literature does not yet provide any certainty on the improvement in live birth rates permitted by time-lapse monitoring (TLM). However, we should emphasize that studies do exist which are based on large sample sizes and include several centers from different countries (Rubio et al., 2012), and that recently a RCT has been published which strongly supports time-lapse as a strategy for embryo selection (Rubio et al., 2014). Finally, we wish to convey that conventional morphological assessment only allows moderate prediction of the embryo’s implantation ability and suffers from relatively limited specificity and sensitivity. In addition, it is prone to inter- and intra-observer variations (Baxter Bendus et al., 2006) and thus can never represent a truly objective quantifiable marker.

### Comprehensive chromosome screening

Despite the identical objective of determining the molecular karyotype of an embryo, not all methods of comprehensive chromosome screening (CCS) are the same (Treff, 2012). In fact, the specific methods used, performance characteristics, and level of supportive evidence of each can be quite specific (Treff and Scott, 2012). For example, there are many methods of DNA amplification, which likely represents the most critical element in the success of CCS. The most common method, whole genome amplification (WGA), can itself be performed using any number of commercially available kits, such as RepliG (QiAgen), Genomiphi (GE Healthcare), GenomePlex (Sigma Aldrich), PicoPlex (Rubicon Genomics) or MALBAC (Yikon Genomics). There are many factors involved in the consideration of which WGA method to use. For example, when the objective is accurate genotyping a multiple displacement amplification (MDA)-based approach (i.e. RepliG) may be more accurate than PCR-based methods (i.e. GenomePlex) (Treff et al., 2011a). Consistent with this observation, methods such as karyomapping (Handyside et al., 2010) and parental support (Johnson et al., 2010a, b) have used MDA-based WGA as they rely upon genotype information to make diagnoses. In contrast, PCR-based WGA methods have demonstrated superior performance for copy number based analyses such as with array comparative genomic hybridization (CGH) (Harper and Harton, 2010), quantitative single nucleotide polymorphism (SNP) array methods (Treff et al., 2010a; Konings et al., 2012), and next-generation sequencing (NGS) (Baslan et al., 2012; Lu et al., 2012; Kohn et al., 2013; Voet et al., 2013; Yin et al., 2013; Zhang et al., 2013). An alternative to WGA involves targeted multiplex PCR, a method that is less expensive and more rapid than WGA, and that has been applied to quantitative real-time (q)PCR (Treff et al., 2012a) and NGS-based CCS (Treff et al., 2013a, b). Ultimately, it is difficult to evaluate which method of initial amplification from the embryo is optimum, as cross-comparison studies have proven to be virtually impossible due to significant commercial conflict of interest.

In addition to the wide variety of available methods for amplification, there are also many platforms for evaluating the amplified DNA. Here the focus is on CGH and SNP arrays, qPCR, and NGS-based CCS. In all cases, platforms of analysis were developed for applications that do not require extensive pre-amplification of DNA and instead were designed for large quantities of starting material [i.e. as in cancer (Nannya et al., 2005), genome-wide association studies (Glessner and Hakonarson, 2011) and prenatal diagnosis (Rajan-Separovic, 2012; Shaffer et al., 2012)]. As a result, performance when starting from only a single or few cells is an important component to consider before application to embryonic CCS. Despite the obvious benefit of the positive control, few studies exist that have rigorously evaluated performance on cells with known aneuploidies. Some argue that even cell lines are poor controls as they may possess significant levels of mosaicism. However, there are many stable cell lines with little to no evidence of mosaicism by conventional testing methodologies, which provide an adequate positive control to evaluate the performance of new testing methods. Examples of the use of blinded analysis of cell lines exist for the development of SNP arrays (Treff et al., 2010a), qPCR (Treff et al., 2012a) and targeted NGS (Treff et al., 2013b).

### CGH arrays

Aneuploidies are extremely common in the early human embryo (Harper et al., 1995; Munne and Cohen, 1998). Trisomic and monosomic embryos account for at least 10% of human pregnancies and, for women nearing the end of their reproductive lifespan the incidence can exceed 50% (Nagaoka et al., 2012). Age-related defects result in higher aneuploidy rates in offspring and an increase in spontaneous abortions, thereby reducing ongoing implantation rates (Boue et al., 1985). Aneuploidy may also be a contributing factor in other infertility populations. For example, an abnormal embryonic karyotype was found to be the most frequent cause of recurrent miscarriage (RM) (Sugiyama-Ogasawara et al., 2012). Recurrent implantation failure (RIF) remains a clinical challenge and embryonic aneuploidy has been implicated as one of the leading embryonic causes (Margalioth et al., 2006). In male factor (MF) infertility, an increase in sperm chromosomal abnormalities due to impairment of the meiotic process was described (Rubio et al.,...
# Table I  Summary of studies published from 2010 that have used time-lapse research on human embryos in a clinical setting.

<table>
<thead>
<tr>
<th>Study</th>
<th>Aim</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aguilar et al. (2014)</td>
<td>Correlate timings of fertilization events with embryo implantation</td>
</tr>
<tr>
<td>Aparicio et al. (2013)</td>
<td>Time-lapse review</td>
</tr>
<tr>
<td>Athayde Wirka et al. (2014)</td>
<td>Characterize atypical dynamic embryo phenotypes identified by time-lapse, and determine their association with embryo development</td>
</tr>
<tr>
<td>Azzarello et al. (2012)</td>
<td>Study PN development in embryos after ICSI</td>
</tr>
<tr>
<td>Basile et al. (2012)</td>
<td>Time-lapse review</td>
</tr>
<tr>
<td>Basile et al. (2013)</td>
<td>Evaluate the effect of two types of culture media on early markers of embryo development</td>
</tr>
<tr>
<td>Basile et al. (2014)</td>
<td>Elaborate an algorithm to increase the probability of noninvasively selecting a chromosomally normal embryo</td>
</tr>
<tr>
<td>Bellver et al. (2013)</td>
<td>Assess morphokinetic evaluation of embryos derived from obese women</td>
</tr>
<tr>
<td>Campbell et al. (2013a)</td>
<td>Develop a model to categorize the risk of aneuploidy in embryos based on morphokinetics</td>
</tr>
<tr>
<td>Campbell et al. (2013b)</td>
<td>Evaluate the effectiveness of the previously established, morphokinetic-based aneuploidy risk classification model</td>
</tr>
<tr>
<td>Chamayou et al. (2013)</td>
<td>Determine morphokinetic parameters predictive of embryo development and, competence in producing a clinical pregnancy after DS transfer</td>
</tr>
<tr>
<td>Chavez et al. (2012)</td>
<td>Analyze the chronology of early mitotic events in 4-cell embryos</td>
</tr>
<tr>
<td>Chen et al. (2013)</td>
<td>Time-lapse review</td>
</tr>
<tr>
<td>Ciray et al. (2012)</td>
<td>Comparison of different culture media on morphokinetics</td>
</tr>
<tr>
<td>Conaghan et al. (2013)</td>
<td>Assess a computer-automated platform for time-lapse image analysis and blastocyst prediction</td>
</tr>
<tr>
<td>Cruz et al. (2012)</td>
<td>Analyze associations between embryo division kinetics and ability to reach blastocyst stage</td>
</tr>
<tr>
<td>Cruz et al. (2012)</td>
<td>Demonstrate that time-lapse incubation conditions are comparable to conventional incubator</td>
</tr>
<tr>
<td>Cruz et al. (2013)</td>
<td>Determine if IVF or ICSI techniques influence embryo morphokinetics</td>
</tr>
<tr>
<td>Dal Canto et al. (2012)</td>
<td>Analyze cleavage timings in relation to blastocyst potential and implantation</td>
</tr>
<tr>
<td>Ergin et al. (2014)</td>
<td>To compare the detection rate of multinucleation with the time-lapse system and conventional control timing proposed by European Society of Human Reproduction and Embryology (ESHRE) consensus and evaluate its impact on pregnancy rates</td>
</tr>
<tr>
<td>Fréour et al. (2013)</td>
<td>Evaluate early embryo morphokinetic parameters according to female smoking status</td>
</tr>
<tr>
<td>Hashimoto et al. (2012)</td>
<td>Assess the development kinetics of embryos and their ability to develop to blastocysts</td>
</tr>
<tr>
<td>Herrero et al. (2013a)</td>
<td>Time-lapse review</td>
</tr>
<tr>
<td>Herrero et al. (2013b)</td>
<td>Updated timings based on morphokinetic studies developed by Meseguer et al. and implementation of morphokinetic selection criteria</td>
</tr>
<tr>
<td>Iwata et al. (2014)</td>
<td>Analyze the timing of initiation of compaction in human embryos</td>
</tr>
<tr>
<td>Joergensen et al. (2014)</td>
<td>Analyze the cleavage patterns in dipronuclear (2PN) and tripronuclear (3PN) embryos in relation to fertilization method</td>
</tr>
<tr>
<td>Kaser and Racowsky (2014)</td>
<td>Time-lapse review</td>
</tr>
<tr>
<td>Kirkegaard et al. (2012b)</td>
<td>Evaluate the effect of blastomere biopsy on early embryonic development using time-lapse</td>
</tr>
<tr>
<td>Kirkegaard et al. (2013a)</td>
<td>Evaluate the effect of culture oxygen concentrations on embryo development using time-lapse</td>
</tr>
<tr>
<td>Kirkegaard et al. (2013b)</td>
<td>Apply Wong’s model to a large set of transferred embryos to test this hypothesis and the correlation between published time intervals and clinical outcome</td>
</tr>
<tr>
<td>Kirkegaard et al. (2013b)</td>
<td>Test proposed time-lapse parameters during the first 48 h and to analyze if timing during the entire culture period differed between embryos destined to implant and to fail</td>
</tr>
<tr>
<td>Knez et al. (2013)</td>
<td>Elucidate if the presence of sperm vacuoles impacts on early embryo developmental dynamics</td>
</tr>
<tr>
<td>Liu et al. (2014)</td>
<td>To investigate the prevalence and potential causes of reverse cleavage by human early cleavage embryos and its associations with embryonic development and implantation after transfer</td>
</tr>
<tr>
<td>Machtinger and Racowsky (2013)</td>
<td>Time-lapse review</td>
</tr>
<tr>
<td>Meseguer et al. (2011)</td>
<td>Generate and evaluate an embryo selection tool based on morphokinetics</td>
</tr>
<tr>
<td>Meseguer et al. (2012)</td>
<td>Compare pregnancy outcomes of treatments in an incubator with time-lapse versus tissue culture chamber</td>
</tr>
<tr>
<td>Meseguer et al. (2012)</td>
<td>Describe current efforts made to standardize different steps of ART such as time-lapse</td>
</tr>
<tr>
<td>Muñoz et al. (2012)</td>
<td>Evaluate if the type and dose of gonadotrophin influences embryo morphokinetics, and if the estradiol and progesterone concentrations in serum has an impact on embryo</td>
</tr>
<tr>
<td>Muñoz et al. (2013)</td>
<td>Evaluate if type of GnRH analog used during controlled ovarian stimulation influences early embryo developmental kinetics</td>
</tr>
<tr>
<td>Montag (2013)</td>
<td>Current embryo selection strategies and their clinical usefulness, particularly to enhance successful single-embryo transfer</td>
</tr>
<tr>
<td>Montag (2013)</td>
<td>Review of studies that attempt to correlate timings with embryonic aneuploidy</td>
</tr>
</tbody>
</table>

Continued
2001; Rodrigo et al., 2010). PGS by fluorescence in situ hybridization (FISH) for a limited number of chromosomes was widely applied for almost two decades. However, this technique lacked sensitivity and was far from comprehensive with regards to the coverage of chromosome complement. Consequently, there was a clear need for a technique capable of CCS, which could also produce reliable and faithful results in a short period of time. Among other technologies CGH arrays have been identified as a robust and accessible diagnostic approach to assess 24-chromosome aneuploidy, and consequently IVF programs are moving toward CCS using CGH arrays (Simpson, 2012; Handyside, 2013). This platform was first validated by reanalyzing the same embryos with FISH, confirming the high efficiency of the platform for aneuploidy detection with only 2.9% of embryos having no results, and the error rate with discrepancies when compared with FISH was 1.9% (Gutierrez-Mateo et al., 2011). Furthermore, CGH arrays platform were validated in single cells from embryos previously diagnosed as abnormal by FISH, with high concordance rates (Mir et al., 2013). There were no differences in efficiency and accuracy when comparing cleavage and blastocyst stage biopsies with whole embryo CGH array analysis (Mir et al., 2013). Additionally, CGH arrays can be applied to identify unbalanced embryos in carriers of Robertsonian and reciprocal translocations, with a resolution as high as 6 Mb for trophectoderm biopsies (Alfarawati et al., 2011). Therefore, CCS with CGH arrays can be accurately applied at different embryo biopsy stages, and currently, blastocyst biopsy is the most common approach, with some groups also describing successful results with D3 biopsies. For D3 biopsies, euploid embryos are commonly transferred at blastocyst stage, in the same cycle as retrieval. For blastocyst biopsies in most of the centers, blastocysts are biopsied and vitrified for transfer in a subsequent cycle. For both approaches, DNA amplification from a single cell or 4–5 cells is performed, followed by DNA labeling with Cy3 and Cy5 fluorophores and cohybridization onto the arrays for 4–12 h (Fig. 2). Fluorescence intensity is detected using a laser scanner and specific software is used for data processing (Fig. 2).

In cleavage stage biopsies, a retrospective study of 2858 cycles showed the clinical benefits of embryo selection based on CGH array analysis in poor reproductive prognosis couples. The most frequent indication was advanced maternal age (AMA), followed by RIF, MF and couples with a previous trisomic pregnancy. The percentage of informative embryos was 97.9%, with only 2.1% of samples noninformative mostly due to cell loss or damage during manipulation, resulting in defective amplification. Aneuploidy rates were 67.2% in women <38 years and 86.3% in women ≥38 years. Differences in the percentage of chromosomal abnormalities could be attributed to the significant increase in embryos with complex aneuploidies with AMA (9.6 versus 23.5%, P < 0.05); with no differences in the percentages of embryos with chaotic pattern (15.1 versus 16.7%) or in embryos with segmental aneuploidies (6.9 versus 5.1%), in women <38 years and ≥38 years, respectively. Mean pregnancy rate per transfer and implantation rate in women <38 years were 59.0% and 49.0%, respectively. In women ≥38 years, pregnancy rates per transfer and implantation rates stand as high as 51.1% and 46.2%, respectively. Miscarriage rates were also comparable in the <38 years and ≥38 years age groups (11.9 versus 9.9%). The increase in chromosomal abnormalities related to female age was reflected in a higher percentage of cycles without embryo transfer in women ≥38 years, showing 62% of cycles with all embryos chromosomally abnormal. However, when a euploid embryo was transferred, outcome was not affected by maternal age (Rubio et al., unpublished). Another retrospective case–control study comparing the outcome in poor prognosis patients who underwent cleavage stage biopsies with fresh embryo transfer cycles and similar patients undergoing regular IVF cycles showed double implantation and ongoing pregnancy rate with a decrease in multiple pregnancy and miscarriage rates in CCS cases (Keltz et al., 2013).

At the blastocyst stage, Wells et al. (2009) found that the probability of an individual analyzed by CGH arrays of generating a pregnancy was 66.7% compared with 27.9% without CGH array testing. A multicenter retrospective study described an increase in the incidence of aneuploid embryos, which correlated with increased maternal age, observing similar implantation and ongoing pregnancy rates per transfer after CCS in patients up to 42 years of age, after which these rates dramatically declined (Harton et al., 2013). Another retrospective study of 106 blastocyst biopsy cycles showed an increase in aneuploidy rates with AMA that was associated with the higher percentage of complex abnormalities in embryos from AMA patients (Rubio et al., unpublished). The percentages of chaotic embryos and embryos with segmental...
Figure 1 Graphical representation of the published algorithm that, using morphokinetics, suggests a method for embryo selection based on implantation or chromosomal euploidy as the final outcome. Abbreviations are related with timings in hours: t; timing of cleavage from ICSI until the number of cells considered 2, 3, 4 etc.; cc, cell cycle duration; s, synchrony of the cell cycle. In the figure, the calculations of each of the variables used in the algorithms are described graphically.
aneuploidies were similar in both groups. There was a decrease in the percentage of abnormal embryos compared with cleavage stage embryos that could be mostly attributed to the arrest of embryos with chaotic pattern and complex abnormalities before reaching blastocyst stage. Focusing on clinical results, pregnancy rates per transfer and implantation rates in women <38 years were similar to those obtained in cleavage stage biopsies, at 54.3% and 42.4%, respectively (Rubio et al., unpublished). More recently, PGS by CGH array with single euploid blastocyst transfer appears to be a successful strategy for patients with multiple failed IVF attempts (Greco et al., 2014). Moreover, CCS using CGH arrays has been also applied in good prognosis patients with a high potential to increase overall pregnancy rates in IVF programs and to decrease multiple pregnancies when SET is performed. The first RCT comparing SET at blastocyst stage with and without CGH arrays in good prognosis patients showed an aneuploidy rate of 44.9% among biopsied blastocysts, with a significantly higher clinical pregnancy rate in the CCS group (70.9 versus 45.8%, \( P = 0.017 \)). There were no twin pregnancies. This study revealed the limitations of SET when conventional morphology was used alone, even in patients without an increased risk for aneuploidy, as the CGH arrays group implanted with greater efficiency and yielded a lower miscarriage rate than those selected without CGH arrays (Yang et al., 2012).

In summary, the incorporation of CGH arrays appears to result in an increase in pregnancy and implantation rates, showing that aneuploidies are the main cause of reproductive failure with advancing female age. CCS with CGH arrays is a robust and highly effective approach for the assessment of embryo viability, either on D3 embryos or at blastocyst stage.

**SNP arrays, qPCR and NGS**

SNP arrays allow for both genotyping and copy number predictions at thousands of positions in the genome. Some SNP array-based CCS methods rely upon the genotypes alone (Handyside et al., 2010; Johnson et al., 2010a,b), whereas others involve quantitative analysis of copy number assignments (Treff et al., 2010a; Konings et al., 2012; van Uum et al., 2012; Tan et al., 2013). Typically, the signals at each position are compared between the embryo biopsy and those obtained from known normal samples in order to identify possible imbalances in the embryo (Fig. 3). One advantage of SNP arrays is the ability to characterize copy number neutral events, such as recombination sites (Handyside et al., 2010), uniparental disomy (Gueye et al., 2014), parental origin of aneuploidy (Handyside et al., 2010; Rabinowitz et al., 2012) and balanced translocations (Treff et al., 2011b). However, like array CGH, disadvantages include the time to obtain a result and the expense of the procedure, both of which have been overcome by the development of qPCR-based CCS (Treff et al., 2012a).

One reason qPCR is faster and cheaper is the elimination of the need for WGA. Instead, a multiplex PCR reaction is performed to preamplify 96 copy number neutral positions in the genome, four per chromosome. Each of the 96 positions are then quantified in individual reactions using...
TaqMan primers and fluorescent target sequence specific probes on a 384 well plate, followed by normalization to previous data from known normal samples (Fig. 3). The process can be completed in 4 h, requires the least hands on time, is easily automatable, involves relatively low tech equipment and is one-third the cost of array based methods. Another advantage is the ease by which additional primers can be incorporated in order to characterize single gene disorders (Treff et al., 2013a), small duplications and deletions (Treff et al., 2011b), mitochondrial disorders (Treff et al., 2012b) or unbalanced translocations (Treff et al., 2013c).

Perhaps, the most ideal combination of throughput, cost and capability is afforded by NGS-based CCS. While the approach can involve WGA (Baslan et al., 2012; Lu et al., 2012; Yin et al., 2013; Zhang et al., 2013; Fiorentino et al., 2014), it is unclear what if any necessity there is to detect imbalances below the resolution of the whole chromosome, as there is still considerable debate regarding the prevalence of de novo deletions and duplications at the preimplantation stage of development. This is particularly true given that a large percentage of ‘dup/dels’ observed in the preimplantation embryo are likely artifacts of the procedures used to identify them (Van der Aa et al., 2013; Voet et al., 2013). Alternatively, targeted amplification of specific loci in the genome can also be used prior to NGS-based CCS and may provide the most cost effective methodology (Treff et al., 2013c). In either case, the amplified DNA undergoes massively parallel sequencing in order to count the number of sequence ‘reads’ which align to each of the 24 chromosomes. The read counts are then normalized to data from known normal samples in order to define the copy numbers present in the embryo biopsy (Fig. 3). Reduced cost per embryo is provided by the opportunity to perform molecular barcoding, which allows multiple samples to be sequenced in parallel in the same reaction and then segregated back out to the original sample using standard bioinformatics methods (Knapp et al., 2012).

For all methods, mosaicism is of course an important variable to consider, particularly when evaluating consistency of diagnoses from multiple samples from the same embryo (Taylor et al., 2014). The misguided interpretation of discrepant results as true mosaicism instead of...

---

**Figure 3** Representation of three contemporary methods of comprehensive chromosome screening (CCS). Quantitative single nucleotide polymorphism (SNP) arrays can involve whole genome amplification (WGA), followed by labeling and hybridization to an array of synthetic oligonucleotides representing thousands of SNPs in the genome. Embryonic DNA profiles are compared with known normal samples in order to characterize relative quantities of each chromosome and make a CCS diagnosis. A multiplex PCR reaction for four loci per chromosome are preamplified in a quantitative real-time (q)PCR-based CCS approach. Individual loci are then quantified in a 384-well plate format using TaqMan fluorescent probes. Normalization to known normal samples is then performed in order to obtain an embryo CCS diagnosis. Finally, next-generation sequencing (NGS) can involve either WGA or targeted multiplex PCR followed by sequencing and determine relative read counts on each chromosome (Chr) for an embryo CCS diagnosis against known normal samples.
technical artifact has largely overestimated the true frequency of preimplantation embryonic mosaicism, particularly when estimates were made using FISH (Northrop et al., 2010; Treff et al., 2010b). As a result of the well-established technical and clinical limitations of FISH (Mastenbroek et al., 2011a, b), its use to confirm findings of a new method of CCS should not be considered a sufficient level of validation (Gutierrez-Mateo et al., 2011; Mir et al., 2013). One way in which to prevent true mosaicism from impacting the evaluation of the technical consistency of a new method of CCS is to evaluate rebiopsies of embryos that had already been given two consistent diagnoses from two separate biopsies using an alternative CCS methodology (Treff et al., 2012a). A number of recent studies have investigated the frequency and distribution of mosaicism within the blastocyst using contemporary methods of CCS. For example, Capalbo et al. (2013a, b) found a frequency of 15.7% mosaicism in 70 blastocysts and no evidence of preferential segregation of aneuploidy in the inner cell mass or trophoectoderm. Johnson et al. (2010a, b) observed 20% mosaicism within 10 blastocysts originally diagnosed as aneuploid and again no evidence of preferential segregation to either lineage. Fragouli et al. (2011) found 32.4% mosaicism among 19 blastocysts, and Northrop et al. (2010) found 24% mosaicism among 50 blastocysts, both also demonstrating a lack of preferential segregation.

Not only is it important to develop more rigorous control for and incidence estimates of mosaicism but also it is necessary to consider the sensitivity to detection when CCS is performed on multicell trophoectoderm biopsies. This question has been evaluated by preparing artificial mixtures of two cell lines, one aneuploid and one euploid, to determine what level of aneuploidy is necessary to allow detection. One study involved the use of array CGH and indicated a sensitivity to detection at 50% aneuploid:euploid ratios (Mamas et al., 2012). Another study involved the use of SNP arrays and found that 40% aneuploid:euploid ratios were sufficient for detection (Northrop et al., 2010). A variety of NGS-based CCS methods under development are likely to involve similar studies as it is anticipated that this technology will further improve sensitivity to detection of mosaic biopsies.

While mosaicism represents an important variable to consider during preclinical development, it also represents an important challenge to the clinical application of CCS. By definition, it is impossible for any CCS method to accurately diagnose a mosaic embryo since the outcome of CCS is to diagnose the embryo as either normal or aneuploid. In a mosaic embryo, there may in fact be both normal and aneuploid cells. The question of how critical mosaicism is to the predictive value of CCS can only be adequately addressed through the study of clinical outcomes of embryos diagnosed after they have already been selected for transfer using standard morphology without regard to the CCS diagnosis (Scott et al., 2012), and only then is it possible to determine whether the CCS diagnosis would have made the correct prediction for the actual clinical outcomes. This is particularly critical to define the false-positive rate of aneuploidy diagnosis, as the standard RCT is incapable of defining it. The risk of using a method that has not undergone such an evaluation is discarding embryos with true reproductive potential. This risk becomes more relevant when applied to patients of AMA or with a single gene disorder indication, where already less genetically normal embryos will be available for transfer.

In the only reported ‘nonselection’ study to date, the predictive value of an aneuploid diagnosis for a negative clinical outcome was 96% (using a specific SNP array based CCS approach) demonstrating minimal impact of mosaicism and suggesting the safety of applying this approach to the selection of embryos for transfer (Scott et al., 2012).

Debate also remains regarding the optimum timing of embryo biopsy, not only for the safety of the procedure, but also for the predictive value of CCS for the ultimate clinical outcome. For example, evidence suggests that blastomere biopsy reduces the reproductive potential of the embryo (Scott et al., 2013a) and may be one of the most genetically unstable stages of embryonic development (Vanneste et al., 2009; Fragouli et al., 2013). This has now led to a growth in either polar body or trophoectoderm based CCS. Proponents of polar body screening often cite eliminating the impact of mosaicism, that maternal meiosis is the primary origin of aneuploidy, a reduced impact of biopsy, and government regulations as the primary factors favoring a polar body approach (Montag et al., 2009; Geraedts et al., 2010). In contrast, proponents of trophoectoderm screening suggest the opportunity to detect all origins of aneuploidy, the demonstrated safety of the procedure, multiple RCTs demonstrating clinical efficacy, and reduced cost to the laboratory and the patient as favorable factors (Kokkali et al., 2007; Forman et al., 2012; Capalbo et al., 2013a, b). Some even maintain that blastomere biopsy remains a viable approach for patients that fail to produce blastocysts making it difficult for a consensus to be reached at the present time (Wilton, 2007; Rubio et al., 2013).

Preclinical accuracy and safety of the procedure represent important components of new CCS technologies. Retrospective studies may also be useful but only as a preliminary analysis to justify the endeavor of an appropriately designed RCT. There are now at least four reported CCS RCTs across three laboratories and three different CCS technologies, all of which involved testing at the blastocyst stage of development (Schoolcraft et al., 2012; Yang et al., 2012; Forman et al., 2013; Scott et al., 2013b). In all cases, a clear clinical benefit was demonstrated, validating the strategy of aneuploidy screening to improve the efficiency of IVF. Although other studies using polar bodies or blastomeres are forthcoming, the success with, and emphasis on, trophoectoderm biopsy may mark the beginning of a paradigm shift to this stage of analysis. The impact of CCS on IVF outcomes has been assessed through several RCTs, as listed in Table II.

Despite rapid developments made in increasing the sensitivity and accuracy of the technologies involved in CCS, together with the growing number of trials showing their efficacy, there are still those who question the validity of this approach (Gleicher et al., 2014; Mastenbroek and Repping, 2014). Future studies will undoubtedly confirm which patient populations may or may not benefit from CCS-based embryo selection.

What is clear is that not all euploid blastocysts result in a healthy newborn. Many fail to implant and progress, indicating that additional factors are important to the reproductive potential of the embryo. Determining the role of the embryonic ‘omics’ or noninvasive predictions obtained by using the media, cumulus cells or time-lapse imaging, in identifying biomarkers is the next challenge. However, the lessons learned from the development and application of new methods of CCS in IVF should at minimum provide a framework for appropriate validation of new embryo selection technologies prior to routine utilization (Scott and Treff, 2010). Furthermore, appropriately designed studies to develop new biomarkers of reproductive potential should include control over the presence of aneuploidy within cases and controls, as this is one of the most well proven factors influencing the reproductive potential of the embryo.
<table>
<thead>
<tr>
<th>Authors</th>
<th>Female age (years)</th>
<th>Intervention</th>
<th>Eligibility</th>
<th>No. of cycles</th>
<th>% Abnormal embryos</th>
<th>Ongoing PR/cycle or delivery rates</th>
<th>MR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yang et al., Mol Cytogenet 2012</td>
<td>&lt; 35</td>
<td>SET after blastocyst biopsy versus blastocyst transfer (Array CGH)</td>
<td>Young good prognosis, IVF patients, first cycle, no prior miscarriage</td>
<td>55 PGS</td>
<td>44.9%</td>
<td>69.1 PGS versus 41.7 (P = 0.0009)</td>
<td>2.6 PGS versus 9.1 (NS)</td>
</tr>
<tr>
<td>Forman et al., Fertil Steril 2013 NCT01408433</td>
<td>&lt; 43</td>
<td>SET after blastocyst biopsy versus DET of unscreened blastocysts (qPCR)</td>
<td>All indications ≥ 2 blastocyst for biopsy</td>
<td>89 PGS</td>
<td>31%</td>
<td>60.7 PGS versus 65.1 (NS)</td>
<td>11.5 PGS versus 20.0 (NS)</td>
</tr>
<tr>
<td>Scott et al., 2013a, b NCT01219283</td>
<td>21–42</td>
<td>Blastocyst biopsy versus blastocyst transfer</td>
<td>All indications ≤ 1 failed IVF</td>
<td>72 PGS</td>
<td>28.6%</td>
<td>84.7 PGS versus 67.5 (P = 0.01)</td>
<td>–</td>
</tr>
<tr>
<td>Schoolcraft et al., ASRM 2012</td>
<td>&gt; 35</td>
<td>Fresh blastocyst transfer versus frozen blastocyst biopsy (SNP microarray)</td>
<td>AMA</td>
<td>47 PGS</td>
<td>–</td>
<td>74.5 PGS versus 53.7 (P &lt; 0.05)</td>
<td>–</td>
</tr>
<tr>
<td>Rubio et al., ESHRE 2014 NCT01571076</td>
<td>38–41</td>
<td>D3 biopsy with blastocyst transfer versus blastocyst transfer (Array CGH)</td>
<td>AMA &lt; 2 miscarriages &lt; 2 IVF failures</td>
<td>75 PGS</td>
<td>77.9%</td>
<td>42.7 PGS versus 25.6 (P = 0.0294)</td>
<td>3.3 PGS versus 43.6 (P &lt; 0.0001)</td>
</tr>
<tr>
<td>ESHRE Study for Oocyte Euploidy (ESTEEM) NCT01532284</td>
<td>36–41</td>
<td>Polar body biopsy (Array CGH)</td>
<td>AMA</td>
<td>86 control</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yilun Siu and Shangai Ji Ai Genetics &amp; IVF Institute NCT02223221</td>
<td>18–35</td>
<td>Blastocyst biopsy versus blastocyst transfer (Array CGH)</td>
<td>RPL ≥ 3 miscarriages Recruiting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rubio IVI NCT01571076</td>
<td>&lt; 38</td>
<td>D3 biopsy with blastocyst transfer versus blastocyst transfer (Array CGH)</td>
<td>Severe male factor &lt; 2 million sperm/ml Recruiting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Munne Reprogenetics NCT01946945</td>
<td>22–42</td>
<td>Blastocyst biopsy versus blastocyst transfer (NGS)</td>
<td>All indications Recruiting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scott RMANJ NCT02032264</td>
<td>18–42</td>
<td>DET blastocyst biopsy (NGS)</td>
<td>≤ 1 prior failed IVF Recruiting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CGH, comparative genomic hybridization; SNP, single nucleotide polymorphism; SET, single-embryo transfer; qPCR, quantitative PCR; NGS, next-generation sequencing; PGS, preimplantation genetic screening; AMA, advanced maternal age; RPL, recurrent pregnancy loss; MR, miscarriage rate.
**Analysis of physiology and the search for biomarkers**

Viability, defined as the ability of an embryo to implant and give rise to a healthy baby, varies even among euploid embryos, with factors other than chromosome complement reflecting the inherent developmental potential of the embryo. Hence the term ‘biomarker’ refers to those proteins and metabolites that are associated, either positively or negatively, with pregnancy outcome. Morphology has been used extensively for over three decades now, and through the development of elegant scoring systems for the pronucleate oocyte (Scott, 2003), cleavage (Van Royen et al., 1999) and blastocyst stage (Gardner et al., 2000) embryos, we already have useful metrics on which to base our decisions regarding which embryo has the highest developmental potential, and with the advent of robust and effective time-lapse systems, further data are emerging to create/augment selection algorithms. Using embryo-grading systems has clearly assisted in improving transfer outcome, but assessment of morphology alone is not absolute and does not necessarily reflect the physiological status of the embryo. For example, it has been established for many years that embryos of the same alphanumeric grade, from the same patient, can differ enormously in their proteomes (Katz-Jaffe et al., 2006a, b) and metabolic activity (Gardner et al., 2001). Clearly, it is time to consider how such analyses can be incorporated into the embryo selection process.

**Proteomics**

The proteome is estimated to consist of over a million proteins, and being derived from the transcriptome, is responsible for cell function. Whereas the analysis of the proteome requires cellular extraction, it is feasible to analyse a large number of proteins and peptides secreted by an individual human embryo, the secretome, noninvasively by sampling the surrounding culture medium. Such analyses have become possible with the advent of novel mass spectrometry platforms, such as surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF MS) (Katz-Jaffe et al., 2005; Katz-Jaffe et al., 2006a, b), and electrospray MS (Beardsley et al., 2010) capable of screening spent embryo culture media for biomarkers. Using SELDI-TOF, Katz-Jaffe et al. (2006a, b) were able to identify an 8.5-kDa protein whose abundance increased only in the secretome of developing blastocysts, indicating a potential relationship between this protein and developmental capability. Protein isolation and identification by tandem MS with peptide sequencing revealed the best candidate for this 8.5-kDa protein biomarker to be ubiquitin. Ubiquitin has been implicated in the implantation process during mammalian development through the control of the activities and turnover of key signalling molecules. Other approaches for the analysis of the secretome have included the use of protein microarrays containing 120 antibody targets (Dominguez et al., 2008). Similarly, a targeted proteomics platform, Luminex, has provided great insight into the proteome of the fluids of the female reproductive tract (Hannan et al., 2011). Could such proteomic analyses be of value in the diagnosis of embryo ploidy, given that the proteome reflects the transcriptome? A pioneering study by McReynolds et al. (2011) has revealed that the presence of lipocalin-1 is increased in the secretome of aneuploid human blastocysts, data obtained through MS and subsequently confirmed through enzyme-linked immunosorbent assay. Whether it is possible or indeed desirable to analyze the karyotype of an embryo through its secretome remains to be resolved. What is evident however, is that differences in the proteome exist due to differences in gene expression between male and female preimplantation embryos (Epstein et al., 1978; Kobayashi et al., 2006; Bermejo-Alvarez et al., 2011), which will reflect differences in physiology (Gardner et al., 2010, 2011).

It is evident that proteomics is currently providing a unique insight into embryo (and endometrial) physiology and signaling (Hannan et al., 2011; Binder et al., 2014). Despite of its promise however, the analysis of the embryonic secretome has yet to be adopted clinically. A key determinant for this is the sheer cost of MS platforms and that further basic research is required to further identify peptides/proteins that reflect embryo viability and ploidy. Once these biomarkers have been identified, then adopting a targeted proteomic approach becomes feasible. With rapid developments in the field of microfluidics, it is fair to assume that it will become possible to perform quantitative and targeted proteomics within an IVF laboratory.

**Metabolomics**

Defined as the systematic study of the unique chemical footprints that specific cellular processes leave behind, metabolomics is able to characterize typically thousands of metabolites (Brisson et al., 2007; Singh and Sinclair, 2007; Botros et al., 2008). Depending upon the platform technology used, it is feasible to both identify and quantify metabolites of interest. Quantitative platforms include RAMAN (Seli et al., 2007) and 1H nuclear magnetic resonance (Seli et al., 2008), but the limitation of sensitivities of these technologies means that although they are powerful research tools in human reproductive biology (Gook et al., 2014), working with individual embryos is not feasible. More sensitive is Near Infrared (NIR) spectroscopy, capable of providing detailed spectra of spent media from a single embryo (Seli et al., 2011). However, although such technologies have been used to create algorithms associated with embryo viability, NIR cannot identify nor quantify specific nutrients, although this latter characteristic should be of relatively small concern if this approach to embryo selection can be shown to provide statistical increases in implantation rates. Studies using NIR initially created a relative ‘embryo viability score’, which was intended to reflect embryo developmental potential. It was subsequently determined that this score did indeed correlate to both positive and negative implantation outcomes. Interestingly, when human embryos of similar morphology were examined using the same NIR spectral profile their viability scores varied remarkably in relation to morphology, indicating that the metabolism of embryos was not necessarily reflective of embryo score (Vergouw et al., 2008). These data are consistent with those of Katz-Jaffe et al. (2006a, b) who determined that the proteome of individual human blastocysts of the same grade differed, confirming that embryo morphology is not completely linked to its physiology. To date, however, the algorithms created using NIR have not been proven in prospective trials (Ahlstrom et al., 2011; Hardarson et al., 2012; Vergouw et al., 2012). One of the main issues encountered with the NIR system was that the threshold of signal distinguishing between a viable and nonviable embryo was susceptible to signal noise. NIR spectroscopy systems and the algorithms generated from them can create models that inadvertently conceal problems in a particular platform. Consequently, a method established and cross-validated on a larger scale can be problematic as the variation can lie within the technical platform itself. Hence, if these problems can be dealt with successfully, it is feasible that NIR could still have a role to play in embryo selection.
More recently, two other MS approaches have been piloted to analyze spent human embryo culture media; electrospray ionization mass spectrometry (ESI-MS) (Cortezzi et al., 2013) and direct injection (DI)-MS (Sheedy et al., 2014). Both represent a fast analytical approach, requiring minimal sample preparation, and the ability to produce large amounts of data from culture media obtained from a single embryo. Such approaches may yet serve as valuable tools in the identification of novel biomarkers.

**Targeted metabolomics**

An approach that has been used for over three decades for the advancement of assisted reproductive techniques (ART) is referred to here as ‘Targeted Metabolomics’, where one or several known metabolites used by the embryo can be quantified noninvasively. Examples of this include glucose and one of its metabolites lactate, and amino acids together with ammonium. However, unlike genetic diagnosis, which determines whether a specific gene is normal, or whether the correct chromosome complement is present, an accurate diagnosis of embryo viability is confounded by the fact that the proteome/secretome and consequently physiology and metabolome can be directly affected by the conditions used to sustain embryo development in the laboratory. Furthermore, there are growing data on differences between male and female embryos during the preimplantation period. Therefore, if an embryo is analyzed under one set of conditions, it is plausible that as the surrounding conditions change, so too do the parameters that are measured. Factors affecting the metabolic profile of the human embryo include: embryonic stage of development, sex of the embryo, medium composition (i.e. differences in nutrient concentrations, albumin source), frequency of medium renewal, rate of ammonium accumulation and oxygen concentration. Perhaps, one of the most extensively studied factors that regulate embryonic function is oxygen, having been shown to affect gene expression, the proteome and metabolic activity of the mammalian embryo. In a series of studies, Wale and Gardner (2012) revealed that atmospheric oxygen exhibited a bi-phasic effect on amino acid utilization by the mouse embryo, enhancing turnover during the cleavage stages but reducing their turnover post-compaction. Furthermore, atmospheric oxygen was found to inhibit glucose uptake by the blastocyst (Wale and Gardner, 2012). Strikingly, when atmospheric oxygen was used in the presence of ammonium in the medium, the pattern of amino acid metabolism was dramatically altered (Wale and Gardner, 2013). Consequently, biomarker levels are affected under stress, a finding that warrants further investigation prior to the routine application of such technologies, and certainly no single overarching hypothesis of embryo metabolism can be applied to all stages and all culture conditions (Gardner and Wale, 2013).

To date, only a handful of studies have been performed on carbohydrate utilization and subsequent viability of the human embryo. In a retrospective analysis Conaghan et al. (1993) observed an inverse relationship between pyruvate uptake by 2- to 8-cell embryos cultured in the presence of serum and 20% oxygen and subsequent pregnancy. In a study on human morulae and blastocysts with different degrees of expansion, no conclusive data were generated on the ability of nutrient consumption or utilization to predict pregnancy outcome (Jones et al., 2001). However, in both studies the medium used to assess embryo metabolism was a simple one, lacking lactate, amino acids and vitamins. Consequently, embryos analyzed under such conditions could be expected to be experiencing considerable metabolic stress.

**Glucose utilization**

A relationship between metabolic activity and embryo development and subsequent viability has been established over several decades. As early as 1970, Menke and McLaren (1970) observed that mouse blastocysts, cultured in a basic medium (lacking amino acids) lost their ability to oxidize glucose. This initial observation was followed by several studies that elucidated that changes in embryo metabolism were associated with loss of developmental capacity in vitro. In 1980, Renard et al. (1980) determined that D10 cattle blastocysts which consumed glucose at a rate greater than 5 µg/h developed better in culture and in vivo after transfer than those blastocysts with a glucose uptake below this value. In 1987, using noninvasive microfluorescence, Gardner and Leese (1987) went on to measure glucose uptake by individual D4 mouse blastocysts prior to transfer to recipient females. Those embryos that went to term had a significantly higher glucose uptake in culture than those embryos that failed to develop after transfer. Interestingly, if glucose uptake is linked with lactate production, to create an indirect measure of glucose flux through the glycolytic pathway, it creates a second, and potentially greater reflective measure of normalcy. In the mouse model, this approach has been used to show that even a transient (6 h) exposure to environmental stress at the blastocyst stage, in this case provided by a medium lacking amino acids and vitamins, results in significant elevations in glycolytic activity culminating in a significant reduction in implantation and fetal development (weight) post transfer (Lane and Gardner, 1998). Encouragingly, the levels of glycolysis at the blastocyst stage have been used successfully to select prospectively viable mouse blastocysts for transfer by Lane and Gardner (1996). Morphologically identical mouse blastocysts with equivalent diameters were identified using metabolic criteria, as ‘viable’ prior to transfer and had a fetal development of 80%. In contrast, those embryos that exhibited an abnormal metabolic profile (compared with in vivo developed controls), developed at a rate of only 6%. Analysis of the relationship between human embryo nutrition and development in vitro was undertaken by Gardner et al. (2001), who determined that glucose consumption on D4 by human embryos was twice as high in those embryos that went on to form blastocysts. Subsequently, Gardner et al. (2011) went on to confirm a positive relationship between glucose uptake and human embryo viability on D4 and D5 of development (Fig. 4). Furthermore, the data generated indicate differences in nutrient utilization between human male and female embryos, a phenomenon previously documented in other mammalian species (Tiffin et al., 1991; Gardner et al., 2010). Of further interest was the observation by Gardner et al. (2011) that the relationship between glucose uptake by the blastocyst and pregnancy rate was not tied to the alphanumeric score of the embryo (Gardner et al., 2011). In other words, pregnancy was associated with the biomarker more than with morphology, a factor considered in the above sections on morphokinetics and proteomics.

Given the evidence of a relationship between glucose metabolism and blastocyst viability, the question is raised—why has glucose uptake not been adopted? At first, it would appear straightforward to quantify the levels of glucose in embryo culture media, given that glucose is measured routinely in biological fluids in so many areas of medicine. Whereas it is feasible to measure glucose concentration in blood and urine, the issue with regards to culture medium is one of accuracy and sensitivity. Standard approaches lack the ability to accurately measure the small differences in nutrient concentration that can be attributed to an individual
embryo. Hence, it has been those few laboratories with access to quantitative microfluorimetry, capable of the accurate quantification of nutrients in sub-microlitre volumes which have produced the majority of studies in this field. However, this could change rapidly with advances in microfluidic devices capable of performing single-cell analyses (Urbanski et al., 2008).

Amino acid utilization
Houghton et al. (2002) quantified amino acid turnover by individual human embryos cultured to the blastocyst stage. A different pattern of amino acid utilization was determined between embryos that went on to form a blastocyst and those embryos that failed to develop. It was found that leucine was taken up from the culture medium more by embryos that went on to develop. The profiles of the amino acids alanine, arginine, glutamine, methionine and asparagine flux also predicted blastocyst potentiality at 95%. Subsequently, Brison et al. (2004) went on to report changes in concentration of amino acids in the spent medium of human zygotes cultured for 24 h to the 2-cell stage in an embryo culture medium containing a mixture of amino acids. It was found that aspargine, glycine and leucine were all significantly associated with clinical pregnancy and live birth. An interesting observation from these studies was that overall embryos that developed displayed a lower turnover of amino acids than ‘nonviable’ embryos, leading to the hypothesis that a ‘quiet metabolism’ is reflective of viability (Leese, 2002; Baumann et al., 2007). It was suggested that embryos with a low metabolic activity reflected a less stressed physiology, and consequently embryos classified as viable would be those that had low nutrient uptake and turnover (Leese et al., 2008). However, a number of studies provide data that do not support this hypothesis (Lane and Gardner, 1996; Gardner et al., 2011), rather they indicate that viability is associated with increased metabolic activity. Indeed, data on oxygen consumption by cleavage stage human embryos indicate that viability is associated with an increased oxygen consumption rate, reflecting an increase in respiration rate (Tejera et al., 2012). So how can this paradox be resolved? An examination of the studies upon which the ‘quiet hypothesis’ was built uncovers that a common factor among them is the use of 20% oxygen for either embryo culture and analysis, or during the actual analysis of metabolism. Given the documented negative impact of 20% oxygen on gene expression, embryonic proteome and metabolism described above, the significance of the ‘quiet hypothesis’ for embryos cultured under physiological oxygen conditions must be carefully reviewed and studies performed to determine what is the optimal nutrient utilization profile under physiological conditions. It is most plausible that there is an upper and lower value of metabolic normality, outside which the embryo will show a decline in viability. Animal data to date indicate that a metabolism similar to that of the embryo developed in vivo would reflect the highest viability, and that metabolic adaptations to inappropriate conditions in vitro carries a cost, culminating in compromised development post transfer (Lane and Gardner, 1998). Furthermore, as glucose uptake (Gardner et al., 2010, 2011) and amino acid utilization (Picton et al., 2010) are related to the sex of embryo, such data need to be factored in to further hypotheses on embryo metabolism and in identifying the optimal rates of metabolic activity for embryo selection.

Figure 4 Relation between glucose consumption on D4 of development and human embryo viability and embryo sex. (A) Glucose uptake on D4 of embryonic development and pregnancy outcome (positive fetal heart beat). Notches represent the confidence interval of the median, and the depth of the box represents the interquartile range (50% of the data), whiskers represent the 5% and 95% quartiles. The line across the box is the median glucose consumption. **significantly different from pregnant ($P < 0.01$). (B) Glucose uptake by male and female embryos on D4 of development. Notches represent the confidence interval of the median, and the depth of the box represents the interquartile range (50% of the data), whiskers represent the 5% and 95% quartiles. The line across the box is the median glucose consumption. * significantly different from male embryos ($P < 0.05$). Data from Gardner et al. 2011).
Figure 5 Concept microfluidic device for biomechanical analysis of a preimplantation embryo. A blastomere is first biopsied from a cleavage stage embryo (A). The blastomere is then loaded into the microfluidic device and driven through a channel constriction (B). This causes morphological compression (C), from which some basic biomechanical properties can be calculated from curvatures of the rear (α), side (β) and front (δ) aspects of the plasma membrane. The same blastomere can then be moved to an actuation zone and lysed (D and E), with subcellular fragments divided by simple channel branching to other regions for parallel molecular analysis on-chip, or to other off-board platforms (F). From Thouas et al. (2013) (with permission from Springer).

Figure 6 Schematic of an embryo culture system for perfusion culture and analysis of biomarkers. Culture media are continuously passed over the embryo(s). The composition of the culture media can be changed according to the specific requirements of each stage of embryonic development. Toxins, such as ammonium, are not able to build up and impair embryo development, while more labile components of the culture system are not denatured. Samples of culture media can be removed for biomarker analysis. Adapted from Gardner (1994).
When morphokinetics meets metabolism
A recent study has considered the relationship between embryo morphokinetics and metabolic activity. In the first study of its kind, mouse embryos underwent both time-lapse and metabolic analyses. Embryos had their rate of development quantified together with the utilization of glucose and amino acids of the resultant blastocysts. It was determined that embryos assigned to the fastest quartile gave rise to blastocysts with a higher glucose uptake and lower glycolytic activity (reflective of increased viability) than their counterparts which developed from the slowest quartile (Lee et al., 2015). Furthermore, differences in amino acid utilization were also detectable, with those blastocysts developing from fast embryos exhibiting higher aspartate uptake, which is of significance for the regulation of the malate-aspartate shuttle in the blastocyst (Mitchell et al., 2009; Gardner and Wale, 2013). Of interest, within each quartile embryos exhibited a range of metabolic activities, supporting the move to use multiple parameters in diagnosing embryo viability. These data indicate that morphokinetics and targeted metabolomics are complementary approaches for the assessment of embryo development and viability.

Laboratory on a chip
The human embryo represents limited biological material for quantitative analysis, which makes it very attractive for analysis through microfluidic devices, which are capable of moving, mixing and accurate analysis of sub-microlitre volumes. The potential of such 'Laboratory-on-a-Chip' devices in reproductive biology has been considered for over a decade, and to date microfluidic devices exist, capable of performing several of the tasks associated with human ART including oocyte denudation (Beebe et al., 2002), sperm sorting (Schuster et al., 2003), IVF and embryo culture (Swain et al., 2013), and analysis of biomarkers (Urbanski et al., 2008; Heo et al., 2012). Although currently a single platform to undertake all tasks does not exist, it is envisaged that in the near future such devices will be employed clinically for specific tasks. Furthermore, such devices could be used for both genetic and biomarker analysis (see Fig. 5), given that microfluidic devices are commercially available for qPCR and gene arrays. As it is possible to visualize the embryo within such devices, it is envisaged that morphokinetic data will be collected concomitantly, and that the ‘IVF-Lab-on-a-Chip’ will reside within a time-lapse incubation system (Fig. 6).

Conclusions
Edwards et al. (1984) observed that faster cleaving embryos gave rise to more viable pregnancies, and in an animal model, it was determined that there existed a relationship between glucose uptake by blastocysts and subsequent development post-transfer (Renard et al., 1980). With developments in time-lapse technologies, together with advances in culture systems, we are now confirming, and advancing, such observations made three decades earlier. Furthermore, together with rapid developments and implementation of new molecular means of CCS, we are creating new paradigms for the way human IVF and embryo diagnosis are being performed. Additionally, we are on the verge of understanding what relationships exist between embryo morphokinetics, karyotype, gene expression and physiology, which will help to formulate comprehensive and robust algorithms for embryo selection (as well as the development of more refined culture systems). Microfluidic devices for embryo culture and analysis may yet prove to be a practical means for the integration and application of these new technologies. Combined with developments in vitrification technologies for embryo cryopreservation (Vajta, 2013), there is now no longer a need to transfer the embryo back to the uterus in same cycle that the oocytes were collected, thereby avoiding a uterus whose receptivity has been compromised by exposure to exogenous gonadotrophins during ovarian stimulation (Gardner, 2008; Evans et al., 2012, 2014). Rather, following time-lapse monitoring, biopsy for karyotyping and analysis of spent culture media for biomarkers, one could rank each euploid embryo for transfer in a natural cycle on the appropriate day of receptivity for the patient (Ruiz-Alonso et al., 2013), thereby significantly reducing the time to conception.

Authors’ roles
All four authors contributed equally to the inception and writing of this manuscript.

Funding
No funding was received for the preparation of this work.

Conflict of interest
None of the authors have any conflict of interest.

References


Diagnosis of human embryo viability


Lee YSL, Thousa GA, Gardner DK. Developmental kinetics of cleavage stage mouse embryos are related to their subsequent carbohydrate and amino acid utilisation at the blastocyst stage. Hum Reprod 2013; 28:543 –552.


Swain JE. Thinking big by thinking small: application of technology to reduce the need for biopsy and PGS. Semin Reprod Med 2010;28:305–315.


Wale PL, Gardner DK. Oxygen affects the ability of mouse blastocysts to regulate ammnonium. Biol Reprod 2013;89:75.


Wells D, Frugoli E, Alfarawati S et al. Increased embryo implantation and high birth rates following comprehensive chromosomal screening of in vitro fertilized embryos by...


