



Which embryo selection method should be offered to the patients?

Peter Kovacs¹ · Harry J. Lieman²

Received: 21 March 2019 / Accepted: 22 March 2019 / Published online: 5 April 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Despite the vast improvements of assisted reproduction, IVF live birth rates have only slowly improved. Delayed childbearing and the well-known negative effect of age on reproduction is one likely explanation. Another often cited explanation is that we still cannot choose with high accuracy the embryo with the highest implantation potential for transfer.

Various technologies (metabolomics, proteomics, time-lapse algorithm-based selection (TLM), preimplantation genetic testing for aneuploidy (PGT-A)) have been evaluated as tools to improve embryo selection [1]. TLM has received much attention in recent years. While only a few randomized controlled trials have been published assessing TLM's benefits (11 clinical trials listed on Pubmed from the past 8 years), far more reviews, meta-analyses and commentaries (4 times as many) have been published on it. Reviews, expert opinions, and meta-analyses often come up with conflicting conclusions leaving clinicians in a difficult position when couples need to be counseled [2, 3].

There are multiple reasons for conflicting opinions. It is always said that a review or meta-analysis is only as good as the studies it is based on. The TLM studies in the literature are the perfect example why it is almost impossible to come up with a good meta-analysis. The various RCTs (randomized controlled trials) enrolled rather heterogeneous patient populations. Culture conditions have not been standardized (culture medium, O₂ concentration) and even within an RCT, the conditions under which experimental and control embryos are cultured could differ. Due to technical differences in the equipment, it is also hard to compare the various time-lapse systems. The day of embryo transfer and the number of embryos

transferred in many of the studies are not similar either. Zaninovic et al. in their retrospective analysis tried to control as many confounding variables as possible (culture conditions similar though different culture media used; Embryoscope was used by both centers; two sets of analysis, one with day 3 development and single embryo transfer, one with day 5 development) but the oocyte source was still heterogeneous (autologous vs. donated oocytes) that may limit the generalizability of the results [4].

Time-lapse monitoring is claimed to have at least two advantages over “standard” embryo culture. It provides undisturbed culture conditions and significantly more embryo observation points. These multiple observations allow us to build algorithms that may be predictive of clinical outcome after the transfer of the selected embryo. However, the different studies use different algorithms. Some use only early kinetic markers. Others use both early and late kinetic markers and yet others add morphology to the kinetic markers as well. Finally, the different studies use algorithms to predict various clinical outcomes. Therefore, it is not surprising that meta-analyses published within one year of each other, which base their findings on only slightly different RCTs after including-excluding certain studies for various issues, draw opposing conclusions [2, 3]. Zaninovic et al. studied the correlation between a few early kinetic markers and implantation/blastocyst formation using two datasets. They measured receiver operator characteristic curve (ROC) and area under the curve (AUC) to measure how good the correlation was between these markers and endpoints. They obtained similar distribution of the data in the two clinics and the kinetics of early events tended to fall within similar time ranges. However, the predictive ability of the studied parameters was relatively poor as the AUCs fell in the 0.5–0.6 range. The combination of various parameters to provide an algorithm to improve the predictive value was not studied.

As mentioned above, multiple algorithms have been proposed by the different research groups. However, external validation has not been able to reproduce their predictive power [5]. Petersen et al. have proposed a universally applicable algorithm, but their results have not been reproduced by others [6]. It has been suggested that local, clinic-specific algorithms

✉ Peter Kovacs
peterkovacs1970@hotmail.com

¹ IVF Center, Kaali Institute, Istenhegy ut 54/a, Budapest 1125, Hungary

² Division of Reproductive Endocrinology And Infertility, Department of Obstetrics & Gynecology and Women's Health, Albert Einstein College of Medicine/Montefiore Medical Center, Bronx, New York, NY, USA

should be prepared. As a result, this technology becomes less attractive to smaller to medium size clinics with fewer cycles as they will need significantly more time to generate enough data to develop proper algorithms. Furthermore, it is not known how many algorithms we need to build. Can one algorithm fit all patients or should various subgroups (advanced vs. younger age, high vs. low BMI, male factor vs. non-male factor, etc.) be studied separately? Zaninovic et al. speculate that the timing differences obtained in the two clinics may be due to age differences of the patient populations. In one set of analysis, however, implantation data is not known as blastocyst development was studied as endpoint. In the second set of analysis, implantation was selected as endpoint and miscarriage rates are not known. Therefore, we still do not know whether embryos identified by the selected markers will result in a healthy live birth or whether the embryos identified are indeed healthy or not.

Are we expected to see a difference in the kinetics of the development of a healthy embryo obtained from younger women as compared to a healthy embryo obtained from older women when culture conditions are standardized? It is known that aneuploid embryos follow a less strict division pattern, but do healthy embryos follow a more predictable pattern regardless of patient characteristics? Is this not what we do with measurements in medicine? As an example, we establish normal ranges for lab values and we apply those to men, women regardless of age. Can we establish normal and abnormal kinetic ranges based on euploidy using standardized culture conditions (at least temperature, pH, gas concentrations should be standard) that are independent of patient characteristics? Accordingly, it would be of interest to evaluate findings of a study in which time-lapse kinetic parameters of confirmed euploid and aneuploid embryos are compared in different patient groups. If under standardized culture conditions universal time-lapse kinetic differentiation could be established, then clinic specific time-lapse parameters would not need to be built and the system would be ready to use.

PGT-A is a competing technology with a similar aim, i.e., improving embryo selection for transfer. Unlike TLM, PGT-A is an invasive technique, often requires elective cryopreservation, off-site testing, and delayed transfer due to the pending biopsy results. The technology is not error free and may result in the loss of embryos that otherwise may have been able to implant and develop into a healthy fetus [7]. Not surprisingly, results in support of PGT-A mostly come from centers with high level expertise in the technological aspects and with an outstanding cryopreservation program. Their results cannot be applied to centers with less expertise in embryo biopsy and processing and less successful cryopreservation programs. The application of PGT-A is further complicated by the often-challenging interpretation of mosaic results.

Both TLM and PGT-A are now said to be tools to rank embryos and determine which one to transfer in order to

achieve success. However, success can also be defined in many ways. The most widely accepted definition of successful IVF treatment is the live birth of a healthy, full-term singleton after the transfer of all embryos created (fresh and frozen, preferably one at a time). If we accept this definition, then none of the currently available technologies can improve the outcome, i.e., live birth. This is true for morphology-based selection too, as it also prioritizes certain embryos for transfer and excludes others from being transferred. Selection based on morphology may also erroneously exclude embryos from transfer or cryopreservation. Loss of “viable” embryos may occur with any of these techniques. Embryo selection tools may influence the order in which the embryos are chosen for the transfer and hence could have significant clinical consequences. They may reduce the time to pregnancy and therefore decrease drop-out from treatment as well as diminish pregnancy loss rates.

Cost-effectiveness with TLM or PGT-A is yet another factor that has been poorly studied.

In the end, the question remains what should clinicians tell their patients? We need to explain that we do not have perfect tools to identify the embryo with the highest implantation potential yet. We have tools that may change the order in which the embryos are transferred but these technologies will not improve the outcome per treatment started. However, they are likely to lead to live birth sooner and may lower the risk of spontaneous abortion. This comes at the risk of potentially losing a small proportion of embryos along with the possibility of additional financial expenses. We should only offer these add on technologies after proper counseling and with the appropriate indication.

When it comes to TLM, further RCTs are needed that compare TLM to other embryo selection/ranking tools (standard morphology, PGT-A) using similar culture conditions and proper clinical aims (time-to pregnancy, pregnancy loss rate, drop-out rate from treatment, cost-effectiveness). It also should be evaluated whether truly universally applicable algorithms that are independent of patient factors (differentiate healthy from non-healthy embryos) could be built when culture conditions are standardized with the aim to further increase the proportion of single embryo transfer cycles.

Finally, once we have mastered embryo selection, we must address further issues as implantation is not all about the embryo. The significance of transfer technique, endometrial receptivity and embryo-endometrium interaction need to be better understood to maximize the chance of success during IVF.

References

1. Montag M, Toth B, Strowitzki T. New approaches to embryo selection. *Reprod BioMed Online*. 2013;27(5):539–46.

2. Armstrong S, Bhide P, Jordan V, Pacey A, Farquhar C. Time-lapse systems for embryo incubation and assessment in assisted reproduction. *Cochrane Database Syst Rev.* 2018;(25):5:CD011320.
3. Pribenszky C, Nilsseld AM, Montag M. Time-lapse culture with morphokinetic embryo selection improves pregnancy and live birth chances and reduces early pregnancy loss: a meta-analysis. *Reprod BioMed Online.* 2017;35(5):511–20.
4. Kovacs P. Time-lapse embryology: do we have an efficacious algorithm for embryo selection? *J Reprod Biotechnol Fertil.* 2016;5:1–12.
5. Zaninovic N, Nohales M, Zhan Q, de los Santos ZMJ, Sierra J, Rosenwaks Z, et al. A comparison of morphokinetic markers predicting blastocyst formation and implantation potential from two large clinical data sets. *J Assist Reprod Gen.* 2019. <https://doi.org/10.1007/s10815-018-1396>.
6. Petersen BM, Boel M, Montag M, Gardner DK. Development of a generally applicable morphokinetic algorithm capable of predicting the implantation potential of embryos transferred on Day 3. *Hum Reprod.* 2016;31(10):2231–44.
7. Paulson RJ. Preimplantation genetic screening: what is the clinical efficiency? *Fertil Steril.* 2017;108(2):228–30.

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