Review Article

Morphological assessment of embryo quality during assisted reproduction: A systematic review

Athanasios Papathanasiou, Bayan Osmani[1], Pek Joo Teoh[2], Abha Maheshwari[1]

Department of Gynaecology, Bourn-Hall Clinic, Cambridge/Norwich, England, [1]Department of Embryology, Aberdeen Centre of Reproductive Medicine, Aberdeen, Scotland, UK, [2]Department of Gynaecology, Concept Fertility Centre, Perth, Australia

ABSTRACT

Background: Various parameters of embryo morphology have been routinely used to select the embryo/s with maximum implantation potential during *in vitro* fertilization (IVF). Hence, there is a dilemma in clinical practice as to which morphological scoring system/test to use. We performed a systemic review to determine the predictive power as well as the clinical and cost-effectiveness of existing morphological tests of embryo quality described in an IVF setting. **Materials and Methods:** The preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines for systematic review were followed. A mixed-method analysis was performed. Qualitative and quantitative techniques were used to synthesize the final results. A narrative summary approach was used for initial data exploration and description, followed by the pooling of data, where appropriate, using Meta-DiSc software. Receiver operating characteristic (ROC) curves were plotted wherever appropriate, and the area under the curve (AUROC) was determined. **Results:** Day 3, day 5, and early cleavage (EC) all had similar discriminatory value for predicting implantation (AUC 0.66, 067, and 0.63 respectively). There was no evidence of improvement in pregnancy rates due to routinely doing EC. No studies were identified that determined the cost-effectiveness of any of the tests. **Conclusions:** All tests have low accuracy. They lack the discriminatory power to identify an embryo that will/will not lead to implantation. Appropriately designed studies are required to assess the predictive value and the clinical and cost-effectiveness of novel embryo scoring technologies.

Keywords: Embryo, implantation, pregnancy, quality, test

BACKGROUND

Multiple pregnancies are the single biggest risk of assisted reproduction. Single embryo transfer (SET) has the potential to virtually eliminate multiple pregnancies. However, despite widespread promotion of SET, only 16.8% of the embryo transfers in the United Kingdom (UK) in 2011 were elective SETs (http://www.hfea.gov.uk/docs/HFEA_Fertility_Trends_and_Figures_2011_-_Annual_Register_Report.pdf). As a result, multiple pregnancy rates were still over 20%. One of the stated barriers for SET is our inability to select the optimal embryo for implantation.[1] By using standard morphological criteria, it may not be possible to select the best embryo at the cleavage stage. Extended culture has been suggested as a preferential method to select the best embryo. However, this has not eliminated multiple embryo transfers, and over 25% of the double embryo transfers (DETs) in the UK in 2011 were at the blastocyst stage. In addition, concerns have been recently raised about preterm labor in pregnancies subsequent to blastocyst transfer.[2] Moreover, the

Address for correspondence:
Dr. Abha Maheshwari,
Consultant Reproductive Medicine and Surgery, Aberdeen Centre of Reproductive Medicine, Aberdeen, Scotland, UK.
E-mail: abha.maheshwari@abdn.ac.uk

Access this article online

Quick Response Code:

Website: www.fertilityscienceresearch.org

DOI: 10.4103/2394-4285.162776

Cite this article as: Papathanasiou A, Osmani B, Teoh PJ, Maheshwari A. Morphological assessment of embryo quality during assisted reproduction: A systematic review. Fertil Sci Res 2014;1:67-80.

cumulative pregnancy rate per woman, after combined fresh and subsequent frozen transfers, is lower for blastocyst transfers compared to transfers at the cleavage stage.[3] Ideally, one would like to be able to determine the embryo with the best implantation potential by day 3, followed by transfer and freezing, in order to maximize cumulative pregnancy rates and minimize multiple pregnancy rates.

Numerous morphological parameters and scoring systems have been advocated to determine the embryo with implantation potential, a testament to the fact that there is no single best test. Theoretically, a combination of multiple scoring methods should improve a test's predictive value. However, a considerable amount of time and money may be spent on doing such tests. Moreover, there are concerns regarding the repeated handling of embryos that may be required when performing such tests: This may adversely affect the incubation and culture process and, subsequently, the outcome of *in vitro* fertilization (IVF). Hence, uncertainty still exists in clinical practice as to which scoring system to use and how effective these tests are.

We performed a systematic review to determine the predictive value, clinical effectiveness, and cost-effectiveness of the various embryo scoring tests based on morphology described in the literature. The purpose of this exercise was to provide evidence-based guidance on the predictive properties of individual tests or combinations of tests, so as to enable IVF practitioners to select the best embryos for transfer to uterus or freezing, with minimal disruption.

MATERIALS AND METHODS

The preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines for systematic reviews were followed.[4]

Data sources and literature search

The searches were performed in two steps. An initial literature search was performed (1988-February 2015) on Medline, Excerpta Medica dataBASE (EMBASE), Cochrane Central Register of Clinical Trials, Cumulative Index to Nursing and Allied Health Literature (CINAHL), and Database of Abstracts of Reviews of Effects (DARE) for published studies (key words: "embryo quality," "embryo", "scoring", "zygote scoring", "cleavage scoring", "early cleavage scoring", "cumulative scoring", "implantation", "pregnancy", "ART"). This initial exercise helped in scoping which tests have been described in the literature. Once the tests were listed, the searches were repeated using key words specific for each test. There were no language restrictions. Relevant journals in the specialty (Human Reproduction, Human Reproduction Update, RBM online, and Fertility and Sterility) were also searched for advance access publications. Cross-references from the included studies were handsearched. Two review authors (AP, PT) independently conducted the searches and selected the studies to be included, while a third author conducted the searches in advance access publications (BO). Repeat searches for each test, as identified from the first step, were undertaken by two authors (AM and AP). Articles were included according to predetermined criteria. Differences of opinion were resolved after team discussion. Data were extracted using predesigned tables. Care was taken to avoid duplication of data in case of two studies from the same authors using the same population.

Study selection

The following inclusion and exclusion criteria were applied.

Inclusion criteria

To determine predictive power

All published studies in which the predictive value of any morphology test of embryo quality was calculated were included if it was feasible to create a 2 × 2 table from the published data, i.e., a normal test and an abnormal test were defined, and the cases with positive and negative tests were compared with a reference standard. In studies where two tests were compared with a reference standard, data for each test were separately extracted.

To determine clinical and cost-effectiveness

All studies that compared outcomes in two groups (those who either had or did not have the test) were included.

Exclusion criteria

Studies where blastocyst formation was used as the reference standard were excluded. We excluded studies evaluating invasive tests and those reporting on tests of oocyte and sperm quality. Conference abstracts and animal studies were also excluded.

Definition of reference standards

Implantation rate

This is defined as number of gestation sacs on ultrasound per embryo transferred.

Clinical pregnancy rate

This is defined as the presence of a fetal heart beat on 7-week ultrasound per embryo transfer.

Live birth rate

This is defined as live birth per embryo transfer.

As this review addresses the predictive power of embryo grading systems, a "per transfer" denominator was considered to be appropriate.

Statistical analysis

To determine predictive power

For each test, data were extracted in 2 × 2 tables. Data were pooled if there were at least two studies that defined the positive and negative test in the same way and compared the test with the same reference standard. When implantation rates acted as the reference standard, pooling of studies was restricted to studies with SETs or where per-embryo data could be extracted. Meta-analysis was attempted wherever appropriate.

The results were organized by entering all data reported on each test from several studies together. Studies were tested for heterogeneity, I^2 index calculated. Summary receiver operating characteristic (SROC) curves were produced wherever inverse correlation was evident, and based on a Spearman correlation coefficient between sensitivity and specificity of 0.6 or more. The Moses-Littenberg linear regression model[5] was used. The area

under the curve (AUROC) with standard error (SE) was calculated. When no SROC could be produced, positive likelihood ratios (LR+) were calculated and reported. The Meta-DiSc software was used.[6] Subgroup analysis was performed using specific features of the test.

To determine clinical effectiveness

For each test, data were extracted in 2 × 2 tables, and pooled if at least two studies had compared the same test. The data was pooled using Rev Man 5.2 (Review Manager 2012, Cochrane Collaboration) to calculate the odds ratio (OR), with 95% confidence interval, of pregnancy. The intervention group received the embryo scoring test of interest, while the control group did not.

Quality of studies

Quality assessment of the included studies was performed by three authors (AP, PT, and BO) using the quality assessment of diagnostic accuracy studies (QUADAS) tool. Any disagreement regarding the type and quality of the studies was resolved after discussion.

RESULTS

Literature search

Table 1 lists the parameters and time of morphological assessment as described in the literature. For all morphology assessments, searches were simultaneously performed. Out of 56 articles, 28 studies were excluded with reasons; two studies had duplicate data. Most studies on morphological assessments were not necessarily designed to determine the predictive value of morphology, as morphology assessment is routine clinical practice in every embryology laboratory. However, we were able to extract data from these articles for prediction of implantation and/or pregnancy. We felt that it was important to assess the predictive value of morphological assessments, as this will put the newer tests into perspective. Morphology was assessed at various stages as follows:

Table 1: Studies assessing the predictive value of zygote scoring

Study	Design & Duration	Population	Index test with definition	SET	Day of ET	QUADAS score
Montag *et al.*, 2001	Prospective study Nov 1999-Oct 2000	512 patients Fresh treatments with ET	≤7 nucleoli polarised vs any other	No	2 or 3 or 5	11
Chen *et al.*, 2006	Prospective randomized study Jun 2002-Jun 2004	165 patients with ET	A: Nucleoli large or medium in size and aligned between the two nucleoli B: Nucleoli large or medium without any particular alignment C: Nucleoli small or pinpoint without and alignment	No	3	11
Balaban *et al.*, 2001	Retrospective study, unknown time period	86 fresh cycles with ET	Number of nucleolar precursor bodies (NPB) in both pronuclei	No	5	10
Liu *et al.*, 2008	Uncertain design Jul 2006-Oct 2006	409 fresh cycles with ET	Z-1: Equal numbers of nucleoli aligned at the pronuclear junction Z-2: Equal numbers and sizes, equally scattered Z-3: Unequal numbers (difference more than one) and/or sizes, or ones with equal numbers of equal sizes but with one pronucleus at the pronuclear junction and the other with scattered nucleoli Z-4: Pronuclei not aligned, of different sizes or not located in the central part of the zygote	No	3	11
Brezinova *et al.*, 2009	Retrospective study, 2004-2006	364 fresh cycles with ET	Number of small evenly distributed nucleolar precursor bodies (NPB) or large NPB with polarized distribution	No	3	11
Nicoli *et al.*, 2013	Retrospective clinical analysis Apr 2008-Nov 2010	755 fresh cycles with ET	Z1: Simultaneously juxtaposed and centralized PN, nucleoli of large size and orientated, and orientation of polar bodies in the longitudinal axis of PN Z2: All other configurations	Yes	2	10
Ludwig *et al.*, 2006	Case-control study Oct 2002-Sep 2003	338 fresh cycles with ET matched with 338 controls	Z1: Equal numbers of nuclear precursor bodies (NPBs) aligned at the furrow between the nucleoli Z2: Equal numbers but not aligned Z3: Not aligned, not polarised Z4: Unequally sized and not aligned	No	2	11
Payne *et al.*, 2005	Prospective clinical study, Unknown time period	100 fresh cycles with ET	(Scott's revised Z-score) *One Z1 *>1 Z1 *One Z3, no Z1 *>1 Z3 *Else, Z2 or Z4	No	2 or 3	11

Zygote scoring (pronuclear morphology)

A total of eight studies assessed the prediction of zygote scoring for embryo quality. Most were retrospective studies. The precise definitions of the index test varied among the included studies [Table 1]. Embryo transfers were performed on either day 2,[7,8] day 3,[9-11] day 5,[12] day 2, 3, or 5,[13] or day 2 or 3.[14]

Prediction of implantation

Six studies assessed the impact of zygote scoring systems on implantation rates.[7,9-13] Except for one study,[7] DETs were performed. Data from these studies were pooled. No heterogeneity (I^2 = 0%) was detected. The Spearman correlation coefficient was 0.829, so a SROC curve was constructed [AUROC 0.57 (SE 0.017)].

Prediction of pregnancy

Seven studies assessed the prediction of clinical pregnancy using a zygote scoring system.[7,8,10-14] Data from these studies could be pooled. No heterogeneity (I^2 = 0%) was detected among the pooled studies. The Spearman correlation coefficient was 1, so a SROC curve was constructed (AUROC 0.58 (SE 0.023)) [Figure 1].

No studies that assessed the clinical or cost-effectiveness of performing zygote scoring were identified. Chen *et al*.[9] compared zygote scoring to early cleavage (EC) in a prospective randomized control trial (RCT) and reported no significant differences with regard to pregnancy rates.

Value in clinical practice

As is evident from Table 1, there is a lack of consensus among the included studies on the exact method necessary to evaluate pronuclear morphology. Even though pronuclear scoring is statistically better than chance in order to predict pregnancy or implantation, it possesses limited accuracy (based on the low LR+) and discrimination (based on the low AUROC). Currently, there is no strong evidence for its routine use in clinical practice.

Day 2 morphology

Six studies were identified where day 2 morphology was assessed with regard to pregnancy or implantation [Table 2]. In all studies, embryo transfers were performed on day 2, except for one,[15] where transfers were performed on day 3. The studies assessed morphology mainly by means of blastomere numbers, fragmentation, or multinucleation. Sjoblom *et al*.[16] assessed morphology by an elaborate weighted score, which also examined other features, such as the zona pellucida thickness and the appearance of the cell cytoplasm, membrane, and perivitelline space. Holte *et al*.[17] included symmetry of cleavage in their scoring criteria. The cutoff values for blastomere cell numbers, fragmentation, and multinucleation differed between studies.

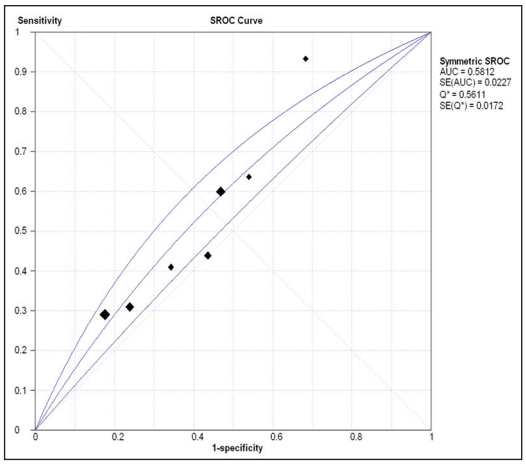

Figure 1: SROC curve for prediction of clinical pregnancy by zygote scoring

Prediction of implantation

Three studies reported data on prediction of implantation.[15-17] Significant heterogeneity was detected (I^2 = 91.5%), and the reported LR+ was 1.56 (95% CI 1.13-2.14). No SROC curve was constructed (Spearman correlation coefficient 0.5).

Prediction of pregnancy

Four studies reported on prediction of pregnancy.[18-20] Significant statistical heterogeneity (I^2 = 83.7%) was detected. A SROC curve could be generated (Spearman correlation coefficient 0.8), and the AUROC was 0.61 (SE 0.05).

Prediction of live birth

Three studies reported on live birth.[18-20] Significant statistical heterogeneity (I^2 = 77.2%) was detected. A SROC curve could be generated (Spearman correlation coefficient 1), and the AUROC was 0.66 (SE 0.08).

No studies that assessed the clinical effectiveness or cost-effectiveness of performing day 2 embryo morphology scoring were identified.

Value in clinical practice

Although there are no separate studies on clinical effectiveness, day 2 morphology assessment prior to embryo transfer is routine practice. However, performing day 2 morphology scoring as an extra test to select embryos for day 3 and beyond is not backed by current evidence as it possesses limited accuracy (based on the low LR+) and discrimination (based on the low AUROC).

Day 3 morphology

Data on the predictive power of day 3 morphology could be obtained from seven studies [Table 3]. Day 3 assessment was based on the number of blastomeres and the degree of fragmentation in all studies but one.[21] As with day 2 morphology, cutoff points varied among studies. In one study,[22] terminology of a good, fair and a poor embryo was used to describe embryo quality. However, various clinics used their own criteria to classify embryos into the three grades mentioned above. Embryo transfers were performed on day 3.

Prediction of implantation

Data could be extracted from five studies for prediction of implantation.[15,21,23-25] There was significant statistical heterogeneity (I^2 = 79%). The SROC was plotted (Spearman correlation coefficient 0.9) and the AUROC was 0.66 (SE 0.05).

Table 2: Studies assessing the predictive value of day 2 morphology

Study	Design & Duration	Population	Index test with definition	SET	Day of ET	QUADAS score
Pelinck *et al.*, 2010	Retrospective cohort study Jan 2001-Aug 2006	449 modified natural cycles with ET	Blastomere cell number on Day 2 Fragmentation rate: Group 1: <10%, Group 2: >10% Cleavage rate (number of blastomeres on Day 3 divided by number of blastomeres on Day 2)	Yes	3	10
Holte *et al.*, 2007	Retrospective record analysis 1999-2001	2266 fresh IVF cycles & 928 ICSI cycles with ET	Integrated morphology cleavage (IMC) embryo score *Number of blastomeres *Degree of fragmentation (0: no fragmentation, 1: <10%, 2: 10-25%, 3: 25-50%, 4: >50%) *Variation in blastomere sizes (0: uniform, 1: <50% variation, 2: >50% variation) *Symmetry of cleavage (0: full symmetry, 1: some asymmetry, 2: pronounced asymmetry) *Presence of single nucleus within the blastomere (ratio of mononucleated blastomeres/total blastomeres) (0: 0-0.25, 1: 0.25-0.5, 2: 0.5-0.75, 3: >0.75) Give-1 for at least one multinucleated blastomere	No	2	8
Sjoblom *et al.*, 2006	Retrospective audit of data Jan 2001-Dec 2001	268 couples 357 fresh cycles with ET	Day 2 weighted scoring-42 hours after insemination. Combined score was generated (maximum 50) based on appearance of zona pellucida thickness, cytoplasm, membrane, cell size, cell shape, perivitelline space, fragmentation and development rate	No	2	9
Lewin *et al.*, 1994	Retrospective analysis	197 patients undergoing frech treatment with ET	Blastomere cell number on Day 2	No	2	10
Jackson *et al.*, 1998	Retrospective review Jan 1991-Jul 1996	483 IVF-ET cycles	Multinucleation Group 1: No multinucleated embryos Group 2: >50% transferred multinucleated embryos Group 3: At least one multinucleated embryo produced, but no multinucleated embryos were transferred	No	2	10
Visser *et al.*, 1993	Uknown design, unknown time period	602 fresh cycles with ET	Day 2 morphology Cumulative Embryo Score (CES): multiply number of blastomeres of each embryo with its morphological grading and add the score of all transferred embryos Morphological grading: Grade 4: >2 cells with equal blastomeres, smooth membranes, translucent cytoplasm and no fragments Grade 3: similar to grade 4 but one morphological characteristic not ideal or <3 blastomeres Grade 2: moderate fragmentation and/or more than one characteristic less than ideal Grade 1: extensive fragmentation or aberrations in most morphological characteristics	No	2	10

Prediction of pregnancy

Data could be extracted from three studies for prediction of clinical pregnancy.[15,23,26] There was significant statistical heterogeneity among studies (I^2 = 90.3%). An SROC was plotted (Spearman correlation coefficient 0.8) and the AUROC was 0.68 (SE 0.08) [Figure 2].

Prediction of live birth

Two studies reported on live birth.[22,26] Significant statistical heterogeneity (I^2 = 92.7%) was detected. The LR+ was 1.29 (95% CI 0.87-1.92).

No studies that assessed the clinical effectiveness or cost-effectiveness of performing day 3 embryo morphology scoring were identified.

Value in clinical practice

Although there are no separate studies on clinical effectiveness, day 3 morphology assessment prior to embryo transfer is routine

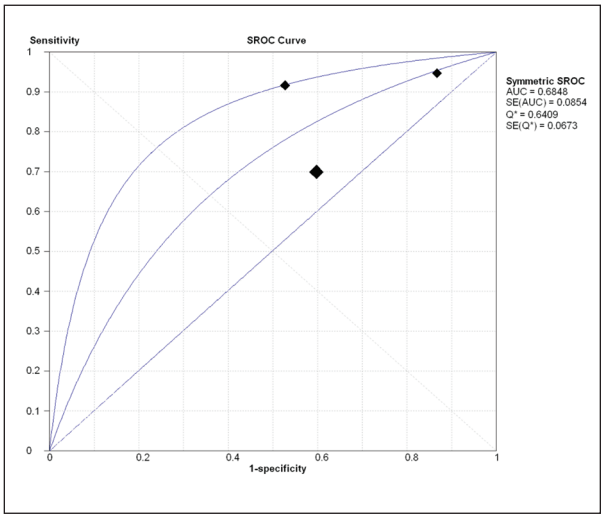

Figure 2: SROC curve for prediction of pregnancy by day 3 morphology scoring

Table 3: Studies assessing the predictive value of day 3 morphology

Study	Design & Duration	Population	Index test with definition	SET	Day of ET	QUADAS score
Qian *et al.*, 2008	Prospective randomized study	117 fresh IVF or ICSI cycles (study group) 420 fresh IVF or ICSI cycles (control group) Jun 2003-Apr 2004	Cumulative embryo score system (CES): Score 0-4 4: at least 5 cells, equal blastomeres, no fragmentation 3: at least 5 cells, equal blastomeres, <30% fragmentation 2: at least 5 cells, distinctly unequal blastomeres,no fragmentation 1: at least 5 cells, equal or unequal blastomeres, 30-50% fragmentation 0: less than 5 cells or >50% fragmentation Multiply the blastomere number with CES	No	3	9
Sajko *et al.*, 2010	Retrospective analysis	115 consecutive unstimulated cycles (73 IVF and 42 ICSI) with ET	Fragmentation: Grade 1: <20%, Grade 2: 20-50%, Grade 3: >50% Blastomere cell number: Group 1: 6 or more, Group 2: less than 6 Blastomere cleavage dynamics : Group 1: number of blastomeres increased by 2 or more from Day 2 to Day 3; Group 2: any other change. Good quality embryo: Belonging to group 1 and grade 1 depending on the combined grouping	Yes	3	10
Fisch *et al.*, 2003	Prospective cohort analysis	106 patients undergoing fresh cycles	Grade A: 7 or more cells with <20% fragmentation	No	3 or 5	11
Pelinck *et al.*, 2010	Retrospective cohort study Jan 2001-Aug 2006	449 modified natural cycles with ET	Blastomere cell number on Day 3 Fragmentation rate: Group 1: <10%, Group 2: >10% Cleavage rate (number of blastomeres on Day 3 divided by number of blastomeres on Day 2)	Yes	3	10
Van Royen *et al.*, 2001	Retrospective analysis May 1997-Nov 1999	745 fresh IVF or ICSI cycles with ET	Fragmentation; <10%, 10-20%, >20%	No	3	9
Check *et al.*, 2007	Retrospective cohort analysis, Jan 1997-Nov 2005	129 IVF-ET cycles	Day 3 morphology *Blastomere cell number *Fragmentation A: none B: 1-25% C: >25%	Yes	3	9
Vernon *et al.*, 2011	Retrospective analysis of the data from multiple clinics	3719 fresh cycles	Embryo quality was labelled as good, fair and poor, depending on individual unit`s criteria	no	2,3,5	11

practice in all embryology laboratories. As a test it possesses limited accuracy (based on the low LR+) and discrimination (based on the low AUROC).

Day 5 morphology

Three studies[27-29] were identified where the predictive value of blastocyst grading on implantation rates was assessed [Table 4]. Blastocyst grading was performed using similar parameters. Two studies were retrospective and one was prospective. All embryo transfers were performed on day 5 or day 6. SETs were exclusively done in one study.[27] All studies assessed the same parameters to estimate blastocyst quality: Blastocyst expansion, inner cell mass appearance, and trophectoderm appearance.

Prediction of implantation

There was significant statistical heterogeneity among the studies (I^2 = 98.6%). A SROC was plotted (Spearman correlation coefficient 1), and AUROC was 0.67 (SE 0.028) [Figure 3]. The pooled LR+ was 1.30 (0.76-2.24). There was significant heterogeneity among the studies (98.6%).

Prediction of clinical pregnancy

Data for prediction of clinical pregnancy could only be extracted from one study.[28] Data from this study showed a clinical pregnancy rate of 52.5% when the blastocyst or early blastocyst was transferred (considered as an embryo with implantation potential).

Prediction of live birth

Only one study[27] assessed live birth rates in association with blastocyst morphological grading. They found that the appearance of the trophectoderm correlates strongly with live birth rate.

No studies have separately assessed the clinical and cost-effectiveness of this test.

Value in clinical practice

Day 5 morphology assessment is performed routinely prior to embryo transfer at day 5. As a test it possesses limited accuracy (based on the low LR+) and discrimination (based on the low AUROC).

Table 4: Studies assessing the predictive value of day 5 morphology

Study	Design & Duration	Population	Index test with definition	SET	Day of ET	QUADAS score
Rehman *et al.*, 2006	Retrospective observational study Apr 1998 – Nov 2004	1292 ICSI and 842 IVF cycles	Blastocyst Quality Score (BQS)= degree of expansion (1-6) + ICM (A=3, B=2, C=1)+ TE (A=3, B=2, C=1)	No	5 or 6	10
Rijnders PM *et al.*, 1998	Prospective study May 1995-Dec 1995	48 fresh cycles	Blastocyst grading E/EB Blastocyst or/and expanded blastocyst M/C Morula or/and compaction	No	5	10
Hill *et al.*, 2013	Retrospective cohort study 2010	694 fresh cycles	Blastocyst expansion: *Morula *Early blastocyst (blastocele <50% of blastocyst) *Expanded blastocyst (blastocele >50% of blastocyst) *Hatched blastocyst Inner cell mass (ICM) score: A: Numerous tightly packed cells B: Several and loosely packed cells C: Very few cells Trophoectoderm (TE) score: A: Many cells organized in the epithelium B: Several cells organized in loose epithelium C: Few large cells	Yes	5	11

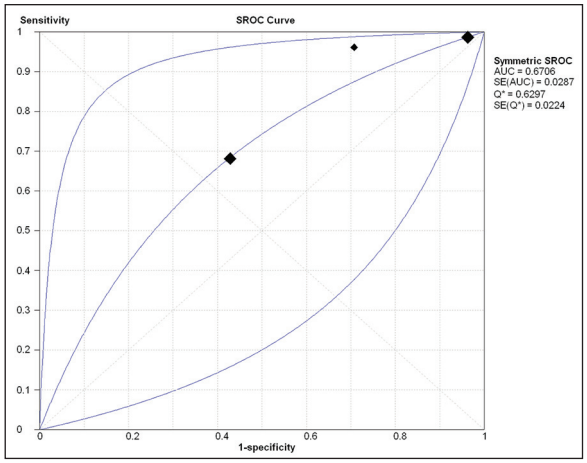

Figure 3: SROC curve for prediction of pregnancy by day 5 assessment

Cumulative embryo score

Six studies[14,16,23,25,30,31] described the predictive value of performing cumulative embryo scoring (CES) [Table 5]. In only one study, the transfers were SETs.[30] Embryo transfers were performed on day 2, day 3, day 4, or day 5. A combination of zygote scoring, EC scoring, and day 2 and/or day 3 scoring was used. The exact methodology for calculating CES scores in each study is described [Table 5]. Qian *et al*. (2008) compared two systems of cumulative scoring. A different scoring system was developed by each study, with different weighting given to various components. Hence, pooling of data was not deemed appropriate.

No studies have addressed the clinical effectiveness or cost-effectiveness of performing CES. It is, therefore, not possible to determine its value in routine clinical practice.

Embryo development rate

Two studies[15,24] assessed the predictive value of embryo development rate assessment on implantation [Table 6]. In both studies, patients were treated in a natural cycle and day 3 SET was performed. However, they used entirely different criteria for identifying a good quality embryo. Hence, pooling of data was not deemed appropriate. However, data for individual study are provided in Table 6.

There are no relevant studies on clinical or cost-effectiveness. Currently, there is no clear evidence to justify the routine use of CES in clinical practice.

EC

Shoukir *et al*.[32] were first to demonstrate that human embryos that had undergone their first cleavage cycle by 25 h *post* insemination achieved higher pregnancy rates during IVF. It is not clear why the time to first cell division varies among embryos; it could be related to the culture conditions as well as intrinsic factors of the oocyte and sperm, maturity issues, genetic competence, and metabolic activity. It has been suggested that metabolically fit embryos cleave earlier due to the availability of energy molecules, such as adenosine triphosphate (ATP), and their highly active mitochondria.[33]

Literature search

Using the search terms "early cleavage"; "IVF" or "ICSI" or "Assisted Conception", and "embryo", 195 articles were identified, of which 65 abstracts were considered relevant. Full texts were obtained and, subsequently, 27 appropriate articles were identified. Twenty articles were included and seven articles were excluded. The full text article of one article could not be accessed, and it was not included. Three more studies were found by cross-searching, and 22 articles were thus included in total [Table 7].[11,32-51]

The characteristics of the included studies are detailed in Table 8. All were observational studies. Recruitment was consecutive in some studies. Blinding was not used in any of these studies. The presence of EC was defined as first cleavage by 25-27 h *post* insemination by most but not all. Some authors considered the presence of two cells, blastomeres, as EC, whereas others included

Table 5: Studies assessing the predictive value of cumulative embryo scoring

Study	Design & Duration	Population	Index test with definition	SET	Day of ET	QUADAS score
Qian *et al.*, 2008	Prospective randomized study	117 fresh IVF or ICSI cycles (study group) 420 fresh IVF or ICSI cycles (control group) Jun 2003-Apr 2004	Cumulative embryo score system (CES) CSS: Add Zygote score (at 16-18h post insemination) + early cleavage (25-27 hours) + embryo morphology (64-67 hours) *Give 20 for Z-1 Zygote (equal number of nucleoli at the pronuclear junction) *Give 10 for Z-2 Zygote (equal numbers and sizes of nucleoli , equally scattered) *Give 20 for regular and symmetrical cleavage, 30 for no fragmentation, 25 for <20% and 0 for >20% *Give 30 for 7c1, 8c1, 9c1, 8c2, compacting 1 and 15 for 7c2, 9c2, 10c3, 11c3 (c1: symmetrical blastomeres with no fragmentation, c2: slightly uneven blastomeres with <20% fragmentation, c3: uneven blastomeres with >20% fragmentation)	No	3	9
Lan *et al.*, 2003	Retrospective analysis Jan 2001-May 2002	fresh IVF or ICSI cycles	Top-quality embryo is: Originates from Z1 zygote (equal number of nucleoli aligned at the pronuclear junction) and has 8 cells, equal blastomeres and no cytoplasmic fragments	Yes and No	3 or 4 or 5	12
Fisch *et al.*, 2003	Prospective cohort analysis	106	Graduated embryo scoring (GES) : = Add Zygote score (at 16-18h post insemination) + early cleavage (25-27 hours) + embryo morphology (64-67 hours) *Give 20 for nucleoli aligned along pronuclear axis *Give 30 for regular and symmetrical cleavage, 30 for no fragmentation, 25 for <20% and 0 for >20% *Give 20 for 7c1, 8c1, 9c1, 8c2 and 10 for 7c2, 9c2, 10c1, 11c1, compacting 1 (c1: symmetrical blastomeres with no fragmentation, c2: slightly uneven blastomeres with <20% fragmentation, c3: uneven blastomeres with >20% fragmentation)	No	3 or 5	11
Meseguer *et al.*, 2011	Retrospective analysis Sep 2009-Sep 2010	247 fresh ICSI own-eggs and oocyte donation cycles	Grade 1: 2PN embryo has 2 cells at 27h, 4 cells at Day 2 and 8 cells at Day 3 Grade 2: 2PN embryo has 1-2 cells at 27h, 3-4 cells at Day 2 and 6-8 cells at Day 3 (only one mismatch is allowed) Grade 3: 2PN embryo has 1-2 cells at 27h, 2-4 cells at Day 2 and 6-8 cells at Day 3 (may have asymetric blastomeres, and multinucleation in up to one blastomere, fragmentation <20%) Grade 4: 1-2 PN embryo has 1-2 cells at 27h, 2-6 cells at Day 2 and 4 to more than 8 cells or morula at Day 3 (asymetric blastomeres and Multinucleation allowed, fragmentation <50%) Grade 5: Any other	No	3	11
Payne *et al.*, 2005	Prospective clinical study, Unknown time period	46 IVF and 54 ICSI cycles	Pronuclear morphology (Scott's revised Z-score) Groups: *One Z1 *>1 Z1 *One Z3, no Z1 *>1 Z3 *Else, Z2 or Z4 Day 2 and Day 3 morphology grading (included uniformity of blastomeres, percentage of fragmentation, rate of cleavage and blastomere multinucleation, no details given): good embryo Two or more good embryos One best embryo Two best embryos 3 best embryos	No	2 or 3	11
Sjoblom *et al.*, 2006	Retrospective audit of data	268 couples, 357 fresh cycles	Day 2 weighted scoring-42 hours after insemination. Combined score was generated (maximum 50) based on appearance of zona pellucida thickness, cytoplasm, membrane, cell size, cell shape, perivitelline space, fragmentation and development rate For embryo selectionsum of the scores from all individuals characteristics D0+D1+D2, making a corrected day 2 score-for ICSI (maximum-20) for IVF(D1+D2 score-maximum of 16)	YES	2	9

Table 6: Studies assessing the predictive value of assessing embryo development rate

Study	Index test with definition	Sensitivity	Specificity	LR+	LR–	QUADAS scoring
Sajko *et al.*, 2010[24] Retrospective analysis N=115 cycles	Good quality embryo: Belonging to group 1 and grade 1 depending on the combined grouping as follows Day 3 embryo morphology Fragmentation: Grade 1: <20% Grade 2: 20-50% Grade 3: >50% Blastomere cleavage dynamics Group 1: number of blastomeres increased by 2 or more from Day 2 to Day 3 Group 2: any other change	0.91	0.24	1.20	0.34	10
Pelinck *et al.*, 2010[15] Retrospective cohort study N=449 cycles	Cleavage rate (number of blastomeres on Day 3 divided by number of blastomeres on Day 2)	0.68	0.48	1.33	0.64	10

Table 7: Studies assessing the predictive value of early cleavage

Study	Time of EC	Type of insemination	Assessment of Symmetry of EC	Definition of EC	SET	Difference in two groups	Outcome	QUADAS scoring
Giorgetti *et al.*, 2007[34]	26 hours	IVF & ICSI	Yes	2cell	Yes	No	Ongoing pregnancy	12
Shoukir *et al.*, 1997[32]	25 hours	IVF	No	2 cell	No	No	Clinical pregnancy	10
Sakkas *et al.*, 1998[35]	27 hours	ICSI	No	2 cell	No	Not known	Clinical pregnancy	12
Bos-Mikich *et al.*, 2001[36]	25-29 hours	IVF & ICSI	No	2 cell	No		Clinical pregnancy	10
Lundin *et al.*, 2001[33]	25-27 hours	IVF & ICSI	No	2 cell 1 cell	No		Clinical pregnancy	
Fenwick *et al.*, 2002[37]	24.5-25.5 hours	IVF	No	2 cell	No	No	Clinical pregnancy	10
Salumets *et al.*, 2003[38]	25-27 hours	IVF & ICSI	Yes	2 cell	Yes	No	Clinical pregnancy	11
Van Montfoort *et al.*, 2004[39]	25-28 hour 23-26 hours	IVF ICSI	No	2 cell	Yes	No	Ongoing pregnancy	11
Van Montfoort *et al.*, 2004[39]	23-26 hour	IVF and ICSI	No	2 cell	No*	No	Ongoing pregnancy	11
Emiliani *et al.*, 2006[40]	25	IVF & ICSI	No	2 cell	Yes	Yes	Clinical pregnancy	12
Yang *et al.*, 2009[41]	25-27	IVF & ICSI	No	2 cell	No	Not known	Ongoing pregnancy	10
Tsai *et al.*, 2002[42]	24-26	IVF & ICSI	No	2cell	No	Not Known		10
Lee *et al.*, 2012[43]	25-27 hours	IVF & ICSI	No	2 cell	No*	Not known	Clinical pregnancy	8
Hammound *et al.*, 2008[44]	25 hours	IVF & ICSI	No	2 cell 0 pn	No*	No	Clinical pregnancy	10
Sakkas *et al.*, 2001[45]	25-27 hours	IVF & ICSI	No	2cell	No	No	Clinical Pregnancy	10
Hesters *et al.*, 2008[46]	25-27 hours 25-26 hours	IVF ICSI	Yes	Even 2 cell	No	No	Clinical pregnancy	11
Lundi *et al.*, 2001[33]	25-27 hours	IVF & ICSI	No	1 cell 2 cell	No	Yes	Pregnancy	10
Fu *et al.*, 2009[47]	25-27 hours	IVF & ICSI	No	2 cell	No	Yes	Ongoing Pregnancy	10
Ciray *et al.*, 2006[48]	26 hours	ICSI	Yes	Even 2cell	No	Yes	Pregnancy	10
Fancsovitis *et al.*, 2005[49]	22-25 hours	IVF & ICSI	No	2 cell	No	No	Clinical pregnancy	10
Ciray *et al.*, 2004[50]	25-27 hour	ICSI	Yes	2 cell	No	No	Clinical pregnancy	10
Brezinova *et al.*, 2009[11]	23-27 hour	IVF & ICSI	No	2 cells, 0 cell	No*	No	Clinical pregnancy	10
Isiklar *et al.*, 2002[51]	27 hours	ICSI	No	2 cells	No	No	Clinical pregnancy	11

*only early cleaved embryos were transferred even if there was DET

any type of cleavage, such as the presence of one cell or the absence of two pronuclei. Some authors only evaluated the number of cells, while others also explored the symmetry of cell division [Table 7].

In eight studies, only EC embryos were transferred in the study group, while at least one EC embryo transfer was included in the study group in the remaining studies. The non-EC groups only included transfers embryos with late cleavage. The time interval from IVF or intracytoplasmic sperm injection (ICSI) to assessment for EC was the same for all studies except for one.[39]

Lundin *et al.*[33] included only one cycle per woman. Yang *et al.*[41] were the only ones that explored subgroups of agonist and antagonist treatment cycles. They found no difference in their antagonist treatment cycles. There was variation among the included studies in the stimulation regimens used, the starting dose of gonadotropins, and the media used for culture.

Van Montfoort *et al.*[39] has three entries in the table, as they provided data separately for IVF and ICSI, and also for DET, where both embryos either had EC or no EC.

Prediction of implantation

In six studies, only SET was performed, and these were used to assess the test's predictive value for implantation. The Spearman

correlation coefficient was considered satisfactory (0.89) for plotting a SROC curve. The AUROC was 0.63 (SE 0.02). There was significant statistical heterogeneity among the studies (I^2 = 88.5%).

Prediction of pregnancy

As all included studies made use of similar methodology to assess EC, they were appropriate for data pooling in order to determine the test's predictive value for pregnancy. The Spearman correlation coefficient was considered satisfactory (0.66) for plotting a SROC curve. The AUROC was 0.62 (SEM 0.02) [Figure 4]. There was significant statistical heterogeneity among the studies (I^2 = 85.7%).

Subgroup analysis after excluding studies that transferred both EC and non-EC embryos did not alter the results.

Clinical effectiveness

Four studies determined the clinical effectiveness of performing EC assessment [Table 9]. Their characteristics are summarised in Table 9. Two out of four studies secured prospective recruitment and random allocation for the two groups. In all four studies, baseline characteristics in both groups were similar. Pooling of data revealed no statistically significant heterogeneity. No statistical difference in the odds of achieving pregnancy was achieved, when comparing EC assessment with no EC assessment (OR 1.29 95% CI 0.98-1.70) [Figure 5].

No studies have been identified that have evaluated the cost-effectiveness of this test.

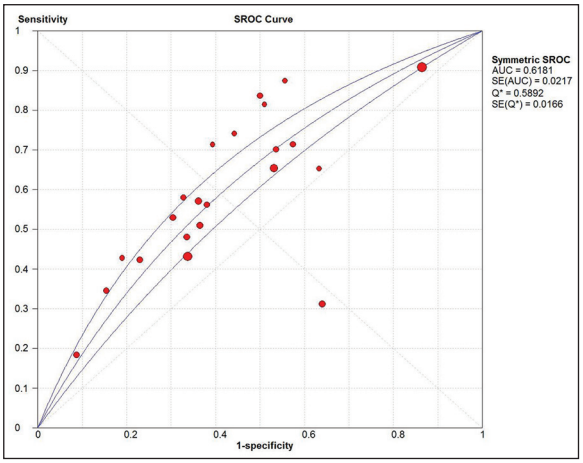

Figure 4: SROC curve for prediction of pregnancy in the presence of early cleavage

Table 8: Summary data for prediction of pregnancy and implantation

Summary prediction of implantation					
Embryo scoring test	**Pooled sensitivity (95% CI)**	**Pooled Specificity (95% CI)**	**Pooled LR+ (95% CI)**	**Pooled LR– (95% CI)**	**SROC (95% CI)**
Zygote scoring	0.52 (0.48-0.57)	0.61 (0.59-0.63)	1.21 (1.12-1.31)	0.85 (0.78-0.93)	0.57 (0.54-0.60)
Day 2 morphology	0.72 (0.68-0.76)	0.59 (0.57-0.61)	1.56 (1.13-2.14)	0.56 (0.39-0.80)	
Day 3 morphology	0.91 (0.88-0.93)	0.20 (0.18-0.22)	1.10 (1.03-1.18)	0.45 (0.23-0.87)	0.66 (0.56-0.76)
Day 5 morphology	0.89 (0.86-0.91)	0.35 (0.31-0.38)	1.30 (0.76-2.24)	0.46 (0.27-0.78)	0.67 (0.61-0.73)
Early cleavage	0.70 (0.66-0.73)	0.34 (0.31-0.37)	1.4 (1.05-1.89)	0.69 (0.51-0.92)	0.63 (0.59-0.67)
Summary prediction of clinical pregnancy					
Embryo scoring test	**Pooled sensitivity (95% CI)**	**Pooled Specificity (95% CI)**	**Pooled LR+ (95% CI)**	**Pooled LR– (95% CI)**	**SROC (95% CI)**
Zygote scoring	0.50 (0.46-0.54)	0.70 (0.68-0.72)	1.26 (1.14-1.40)	0.86 (0.77-0.97)	0.58 (0.53-0.63)
Day 2 morphology	0.51 (0.45-0.57)	0.64 (0.61-0.67)	1.42 (1.08-1.86)	0.79 (0.66-0.95)	0.61 (0.51-0.71)
Day 3 morphology	0.82 (0.76-0.87)	1.29 (1.00-1.66)	1.29 (1.00-1.66)	0.41 (0.16-1.10)	0.68 (0.51-0.85)
Early cleavage	0.58 (0.56-0.60)	0.52 (0.50-0.54)	1.41 (1.23-1.61)	0.74 (0.66-0.83)	0.61 (0.57-0.65)

Table 9: Studies evaluating clinical effectiveness of doing early cleavage

Study	Type of study	Population	Sample size	How the groups were divided Assessed group	SET	Outcome measure	Difference in two groups in baseline characteristics
Giorgetti *et al.*, 2007	Prospective Jan 2003-March 2006	Day 2 ET Both IVF & ICSI No inclusion/exclusion criteria specified	N=193 & 84 cycles No power calculation	Weekday/Weekend	Yes	Ongoing pregnancy	No
Emiliani *et al.*, 2006	Prospective randomized	All couples with at least 2 embryo fertilized and had SET	N=93 and 94 patients No power calculation	Random number table	Yes	Clinical pregnancy	No
Sakkas *et al.*, 2001	Prospective 18 weeks (starting from April 2000)	Day 2-3 ET Both IVF & ICSI No inclusion/exclusion criteria specified	N=77 & 90 cycles No power calculation	Alternate week	No	Clinical Pregnancy	No
Ciray *et al.*, 2004	Prospective	All patients undergoing ICSI over 3 months period	N=138 and 153 patients No power calculation	Random number table	No	Clinical pregnancy	No

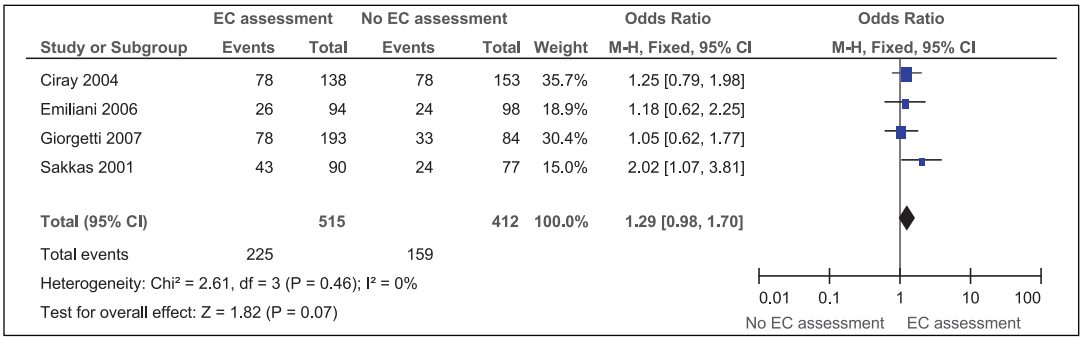

Figure 5: Impact of assessment of early cleavage on pregnancy rates

Value in clinical practice

As a test, EC assessment possesses limited accuracy (based on the low LR+) and discrimination (based on the low AUROC). In addition, the clinical effectiveness studies suggest that it is not an effective test. Based on the available evidence, routine assessment for EC is not recommended as a routine test in IVF practice.

According to the Alpha/European Society of Human Reproduction and Embryology (ESHRE) consensus,[52] checking for EC should be performed 25-27 h *post* ICSI and 27-29 h *post* IVF. The included studies have used a fixed time frame for assessing for EC regardless of the fertilization technique used (IVF or ICSI). Moreover, in all studies except one, assessment was performed no later than 27 h, which is not appropriate after IVF treatment. There was also variation in the definition of EC within the included studies, ranging 0-2 cells. The latest ESHRE consensus has agreed that the presence of 2 cells is required.

DISCUSSION

Main findings

Numerous tests for assessing embryo quality have been described in the literature. Our review has shown that none of the morphological assessments described have a high accuracy to identify the embryos that have good implantation potential. At no point was morphology discriminatory to exclude embryos from transfer or freezing. The predictive capacity for implantation and pregnancy are similar for day 3 and day 5 morphology assessments (AUROC of 0.66 and 0.67 for implantation, and 0.68 and 0.67 for pregnancy, respectively). There is currently no evidence of improvement in clinical pregnancy rates by routinely performing EC assessment during assisted conception treatment.

Strengths

This is the first systematic review of morphological assessment of embryo quality-predicting outcomes during IVF. Two-step searches have been performed to ensure that all tests described in the literature are included. We not only attempted to determine the predictive value of these tests but also explored clinical effectiveness and cost-effectiveness, as these aspects pertain to the application of any test in clinical practice.

Weaknesses

This systematic review is based on observational data. Individual methodological differences, variation in design, inclusion or exclusion criteria as well as differences in the definition of the index tests and reference standards are inherent in systematic reviews of observational studies. In addition, there were a number of limitations.

Exclusion of studies

A number of studies were excluded as they had used development to blastocyst as the reference standard. Although it is assumed that the embryos that reach the blastocyst stage have proved their potential, it is an accepted fact that not all blastocysts implant. Moreover, a meta-analysis of RCTs showed that the cumulative pregnancy rate, after fresh and subsequent frozen embryo transfers, is higher if the transfer takes place on day 3, indicating that those who do not proceed to the blastocyst stage may indeed have embryos with implantation potential.[3] For this reason, we did not consider blastocyst development as an appropriate reference standard for this review.

An ideal study for a predictive test and its comparison to currently available studies

An ideal study testing a predictor of implantation/pregnancy should have a well-defined population, prospective and consecutive recruitment, blinding of those involved in assessing the test results and outcomes, adequate test description, predetermined normal and abnormal test values, and comparison with a gold standard such as live birth. An ideal study to determine the predictive power of any test of embryo quality in this case would have predetermined definitions of a good embryo and an inferior embryo. The ideal outcome should be live birth rate, but implantation and pregnancy rate would also be appropriate. Women should not have a combination of good and lower-quality embryos transferred at the same time. However, within this review, most of the available studies were retrospective, without consecutive involvement. In a significant proportion of them, embryos of varying quality, as determined by the index test, were transferred.

An ideal test and its comparison to currently available morphological assessment as test of embryo quality

An ideal test should be valid both internally and externally, reliable, replicable, discriminatory, cheap, easily available, simple to perform, and noninvasive. In addition, there should be a clear definition of what a normal or an abnormal test is. For any predictive test, it is important to consider what exactly is being predicted. In the present context, it would be either implantation rate or pregnancy or live birth. Assuming that a positive test result indicates a favorable prognosis, sensitivity

reflects the ability of the test to identify all embryos that will result in implantation; specificity reflects its ability to exclude embryos that are not likely to implant; positive predictive value represents the probability of implantation when the index test is positive; and negative predictive value represents the probability of embryos not implanting if the index test is negative. The LR of a positive test quantifies how much more likely it is that a positive test will be found in an embryo that will implant than in an embryo that will not; the LR of a negative test indicates how much more likely it is that a negative test will be found in an embryo that will not implant than in an embryo that will. It is generally accepted that an LR+ of >10 represents a highly accurate test, an LR+ of 5-10 reflects a moderately accurate test, an LR+ of 2-5 indicates weak accuracy, an LR+ of 1-2 very weak accuracy, and a LR+ of 1 indicates no value in terms of predictive accuracy.[53] The LR ratios of the reviewed tests were all low (range 0-2), indicating that these tests perform poorly in terms of prediction of implantation or pregnancy. The AUROC represents the ability of a test to discriminate between a positive and a negative outcome. By definition, an AUROC of 0.5 is consistent with a test that completely lacks discrimination: No better than tossing a coin. None of the reviewed embryo scoring tests performed well in terms of discrimination, as shown by their respective AUROCs, all of which are less than 0.7 [Table 8].

Cost-effectiveness of embryo assessment

Although the cost-effectiveness of performing any embryo assessment was not addressed by the studies mentioned above, there may well be implications in terms of staff time. For example, when assessing for EC, laboratory time schedules are likely to be affected if this stage of examination is introduced into everyday practice. Oocyte collections are usually planned during morning hours, with insemination being performed in the afternoon. According to the ESHRE consensus,[52] the ideal time to determine EC would then be during the evening hours of the day, which may have implications for staff time and subsequent costs incurred. Assessment for EC should lead to significant improvement in pregnancy rates in order to justify the extra effort.

When considering the benefits of a test that involves additional examinations, alongside the standard visual assessments of the developing embryos it is recommended that the temporary interruption inflicted on the culture ecosystem is considered and that the detriment this may have toward the treatment outcome is taken into account.

Implication for clinical practice

Based on current available evidence, there is no justification for using extra morphological assessments that involve taking embryos out of incubators and interfere with the embryo culture system. The need for a test of embryo quality has been recently questioned as, with improved freezing, one could perform a fresh transfer and freeze the rest for subsequent transfers.[54] With successful freezing techniques, the only significant drawback of such an approach would be the potential time delay to achieve pregnancy.

Implication for future research

Like other health interventions, a new diagnostic test is ideally required to pass through various stages of critical assessments. It should be deemed biologically plausible. It is also necessary for the test to be clearly defined, including what constitutes a normal and an abnormal test. Appropriate reference standards should be used in the analysis of the test's sensitivity, specificity, and LRs. The clinical effectiveness and also cost-effectiveness should be demonstrated by high-quality prospective RCTs. Based on these criteria, clinical trial evidence is lacking for many tests of embryo quality. One good example is the time lapse systems, which have already been advocated in clinical practice without proven clinical or cost-effectiveness by appropriate studies. In a parallel example, preimplantation genetic screening had been shown to be of value by retrospective studies. However, when put to the test by an RCT, its usefulness was dismissed — in fact, this was found to be detrimental.[55] Therefore, further research in the form of appropriately designed RCTs is required before introducing such novel modalities into routine clinical practice.

CONCLUSIONS

A large number of morphological assessments of embryo quality have been described in the literature: Evidence that no ideal test exists. The accuracy of all these tests is low. Our review has also shown that none of these tests or combinations of tests has sufficient discriminatory power to exclude an embryo from embryo transfer. Newer techniques need to be further explored prior to their introduction in routine clinical practice.

Financial support and sponsorship

No external funding was sought for this systematic review.

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. Maheshwari A, Griffiths S, Bhattacharya S. Global variations in the uptake of single embryo transfer. Hum Reprod Update 2011;17:107-20.
2. Dar S, Librach CL, Gunby J, Bisonnette F, Cowan L; IVF Directors Group of Canadian Fertility and Andrology Society. Increased risk of preterm birth in singleton pregnancies after blastocyst versus Day 3 embryo transfer: Canadian ART Register (CARTR) analysis. Hum Reprod 2013;28:924-8.
3. Glujovsky D, Blake D, Farquhar C, Bardach A. Cleavage stage versus blastocyst stage embryo transfer in assisted reproductive technology. Cochrane Database Syst Rev 2012;11:CD002118.
4. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med 2009;6:e1000097.
5. Littenberg B, Moses LE. Estimating diagnostic accuracy from multiple conflicting reports: A new meta-analytic method. Med Decis Making 1993;13:313-21.
6. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: A software for meta-analysis of test accuracy data. BMC Med Res Methodol 2006;6:31.
7. Nicoli A, Capodanno F, Rondini I, Valli B, Villani MT, Morini D, *et al.* Pronuclear morphology evaluation in *in vitro* fertilization (IVF)/intracytoplasmic sperm injection (ICSI) cycles: A retrospective clinical review. J Ovarian Res 2013:6:1.
8. Ludwig AK, Werner S, Diedrich K, Nitz B, Ludwig M. The value of pronuclear scoring for the success of IVF and ICSI-cycles. Arch Gynecol Obstet 2006;273:346-54.
9. Chen C, Kattera S. Comparison of pronuclear zygote morphology and early cleavage status of zygotes as additional criteria in the selection of day 3 embryos: A randomized study. Fertil Steril 2006;85:347-52.

10. Liu Q, Zhu G, Hu J, Wei Y, Ren X, Zhang H, *et al*. Relationship between pronuclear scoring and embryo quality and implantation potential in IVF-ET. J Huazhong Univ Sci Technolog Med Sci 2008;28:204-6.
11. Brezinova J, Oborna I, Svobodova M, Fingerova H. Evaluation of day one embryo quality and IVF outcome — A comparison of two scoring systems. Reprod Biol Endocrinol 2009;7:9.
12. Balaban B, Urman B, Isiklar A, Alatas C, Aksoy S, Mercan R, Mumcu A, Nuhoglu A. The effect of pronuclear morphology on embryo quality parameters and blastocyst transfer outcome. Hum Reprod 2001;16:2357-61.
13. Montag M, van der Ven H; German Pronuclear Morphology Study Group. Evaluation of pronuclear morphology as the only selection criterion for further embryo culture and transfer: Results of a prospective multicentre study. Hum Reprod 2001;16:2384-9.
14. Payne JF, Raburn DJ, Couchman GM, Price TM, Jamison MG, Walmer DK. Relationship between pre-embryo pronuclear morphology (zygote score) and standard day 2 or 3 embryo morphology with regard to assisted reproductive technique outcomes. Fertil Steril 2005;84:900-9.
15. Pelinck MJ, Hoek A, Simons AH, Heineman MJ, van Echten-Arends J, Arts EG. Embryo quality and impact of specific embryo characteristics on ongoing implantation in unselected embryos derived from modified natural cycle *in vitro* fertilization. Fertil Steril 2010;94:527-34.
16. Sjöblom P, Menezes J, Cummins L, Mathiyalagan B, Costello MF. Prediction of embryo developmental potential and pregnancy based on early stage morphological characteristics. Fertil Steril 2006;86:848-61.
17. Holte J, Berglund L, Milton K, Garello C, Gennarelli G, Revelli A, *et al*. Construction of an evidence-based integrated morphology cleavage embryo score for implantation potential of embryos scored and transferred on day 2 after oocyte retrieval. Hum Reprod 2007;22:548-57.
18. Lewin A, Schenker JG, Safran A, Zigelman N, Avrech O, Abramov Y, *et al*. Embryo growth rate *in vitro* as an indicator of embryo quality in IVF cycles. J Assist Reprod Genet 1994;11:500-3.
19. Jackson KV, Ginsburg ES, Hornstein MD, Rein MS, Clarke RN. Multinucleation in normally fertilized embryos is associated with an accelerated ovulation induction response and lower implantation and pregnancy rates in *in vitro* fertilization-embryo transfer cycles. Fertil Steril 1998;70:60-6.
20. Visser DS, Fourie FR. The applicability of the cumulative embryo score system for embryo selection and quality control in an *in-vitro* fertilization/embryo transfer programme. Hum Reprod 1993;8:1719-22.
21. Van Royen E, Mangelschots K, De Neubourg D, Laureys I, Ryckaert G, Gerris J. Calculating the implantation potential of day 3 embryos in women younger than 38 years of age: A new model. Hum Reprod 2001;16:326-32.
22. Vernon M, Stern JE, Ball GD, Wininger D, Mayer J, Racowsky C. Utility of the national embryo morphology data collection by the Society for Assisted Reproductive Technologies (SART): Correlation between day-3 morphology grade and live-birth outcome. Fertil Steril 2011;95:2761-3.
23. Qian YL, Ye YH, Xu CM, Jin F, Huang HF. Accuracy of a combined score of zygote and embryo morphology for selecting the best embryos for IVF. J Zhejiang Univ Sci B 2008;9:649-55.
24. Sajko MC, Vlaisavljevic V. Determination of criteria for the assessment of embryo quality with good implantation predictability on a model of unstimulated *in vitro* fertilization cycles — Prediction of implantation in unstimulated cycles. Zdravniski Vestnik 2011;80(Suppl):I-39.
25. Fisch JD, Sher G, Adamowicz M, Keskintepe L. The graduated embryo score predicts the outcome of assisted reproductive technologies better than a single day 3 evaluation and achieves results associated with blastocyst transfer from day 3 embryo transfer. Fertil Steril 2003;80:1352-8.
26. Check JH, Summers-Chase D, Yuan W, Horwath D, Wilson C. Effect of embryo quality on pregnancy outcome following single embryo transfer in women with a diminished egg reserve. Fertil Steril 2007;87:749-56.
27. Hill MJ, Richter KS, Heitmann RJ, Graham JR, Tucker MJ, DeCherney AH, *et al*. Trophectoderm grade predicts outcomes of single-blastocyst transfers. Fertil Steril 2013;99:1283-9. e1.
28. Rijnders PM, Jansen CA. The predictive value of day 3 embryo morphology regarding blastocyst formation, pregnancy and implantation rate after day 5 transfer following *in-vitro* fertilization or intracytoplasmic sperm injection. Hum Reprod 1998;13:2869-73.
29. Rehman KS, Bukulmez O, Langley M, Carr BR, Nackley AC, Doody KM, *et al*. Late stages of embryo progression are a much better predictor of clinical pregnancy than early cleavage in intracytoplasmic sperm injection and *in vitro* fertilization cycles with blastocyst-stage transfer. Fertil Steril 2007;87:1041-52.
30. Lan KC, Huang FJ, Lin YC, Kung FT, Hsieh CH, Huang HW, Tan PH, Chang SY. The predictive value of using a combined Z-score and day 3 embryo morphology score in the assessment of embryo survival on day 5. Hum Reprod 2003;18:1299-306.
31. Meseguer M, Herrero J, Tejera A, Hilligsøe KM, Ramsing NB, Remohí J. The use of morphokinetics as a predictor of embryo implantation. Hum Reprod 2011;26:2658-71.
32. Shoukir Y, Campana A, Farley T, Sakkas D. Early cleavage of *in-vitro* fertilized human embryos to the 2-cell stage: A novel indicator of embryo quality and viability. Hum Reprod 1997;12:1531-6.
33. Lundin K, Bergh C, Hardarson T. Early embryo cleavage is a strong indicator of embryo quality in human IVF. Hum Reprod 2001;16: 2652-7.
34. Giorgetti C, Hans E, Terriou P, Salzmann J, Barry B, Chabert-Orsini V, *et al*. Early cleavage: An additional predictor of high implantation rate following elective single embryo transfer. Reprod Biomed Online 2007;14:85-91.
35. Sakkas D, Shoukir Y, Chardonnens D, Bianchi PG, Campana A. Early cleavage of human embryos to the two-cell stage after intracytoplasmic sperm injection as an indicator of embryo viability. Hum Reprod 1998;13:182-7.
36. Bos-Mikich A, Mattos AL, Ferrari AN. Early cleavage of human embryos: An effective method for predicting successful IVF/ICSI outcome. Hum Reprod 2001;16:2658-61.
37. Fenwick J, Platteau P, Murdoch AP, Herbert M. Time from insemination to first cleavage predicts developmental competence of human preimplantation embryos *in vitro*. Hum Reprod 2002;17:407-12.
38. Salumets A, Hydén-Granskog C, Mäkinen S, Suikkari AM, Tiitinen A, Tuuri T, Early cleavage predicts the viability of human embryos in elective single embryo transfer procedures. Hum Reprod 2003;18:821-5.
39. Van Montfoort AP, Dumoulin JC, Kester AD, Evers JL. Early cleavage is a valuable addition to existing embryo selection parameters: A study using single embryo transfers. Hum Reprod 2004 19:2103-8.
40. Emiliani S, Fasano G, Vandamme B, Vannin AS, Verdoodt M, Biramane J, *et al*. Impact of the assessment of early cleavage in a single embryo transfer policy. Reprod Biomed Online 2006;13:255-60.
41. Yang WJ, Hwu YM, Lee RK, Li SH, Fleming S. Early-cleavage is a reliable predictor for embryo implantation in the GnRH agonist protocols but not in the GnRH antagonist protocols. Reprod Biol Endocrinol 2009;7:20.
42. Tsai YC, Chung MT, Sung YH, Tsai TF, Tsai YT, Lin LY. Clinical value of early cleavage embryo. Int J Gynecol Obstet 2002;76:293-7.
43. Lee MJ, Lee RK, Lin MH, Hwu YM. Cleavage speed and implantation potential of early-cleavage embryos in IVF or ICSI cycles. J Assist Reprod Genet 2012;29:745-50.
44. Hammoud I, Vialard F, Casasnovas P, Lefebvre G, Vauthier-Brouzes D, Poirot C. How viable are zygotes in which the PN are still intact at 25 hours? Impact on the choice of embryo for transfer. Fertil Steril 2008;90:551-6.
45. Sakkas D, Percival G, D'Arcy Y, Sharif K, Afnan M. Assessment of early cleaving *in vitro* fertilized human embryos at the 2-cell stage before transfer improves embryo selection. Fertil Steril 2001;76:1150-6.
46. Hesters L, Prisant N, Fanchin R, Méndez Lozano DH, Feyereisen E, Frydman R, *et al*. Impact of early cleaved zygote morphology on embryo development and *in vitro* fertilization-embryo transfer outcome: A prospective study. Fertil Steril 2008;89:1677-84.
47. Fu J, Wang XJ, Wang YW, Sun J, Gemzell-Danielsson K, Sun XX. The influence of early cleavage on embryo developmental potential and IVF/ICSI outcome. J Assist Reprod Genet 2009;26:437-41.
48. Ciray HN, Karagenç L, Ulug U, Bener F, Bahçeci M. Early cleavage morphology affects the quality and implantation potential of day 3 embryos. Fertil Steril 2006;85:358-65.
49. Fancsovits P, Toth L, Takacs ZF, Murber A, Papp Z, Urbancsek J. Early pronuclear breakdown is a good indicator of embryo quality and viability. Fertil Steril 2005;84:881-7.

50. Ciray HN, Ulug U, Bahçeci M. Transfer of early-cleaved embryos increases implantation rate in patients undergoing ovarian stimulation and ICSI-embryo transfer. Reprod Biomed Online 2004;8:219-23.
51. Isiklar A, Mercan R, Balaban B, Alatas C, Aksoy S, Urman B. Early cleavage of human embryos to the two-cell stage. A simple, effective indicator of implantation and pregnancy in intracytoplasmic sperm injection. J Reprod Med 2002;47:540-4.
52. Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology. The Istanbul consensus workshop on embryo assessment: Proceedings of an expert meeting. Hum Reprod 2011;26:1270-83.
53. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. JAMA 1994:271;703-7.
54. Mastenbroek S, van der Veen F, Aflatoonian A, Shapiro B, Bossuyt P, Repping S. Embryo selection in IVF. Hum Reprod 2011;26: 964-6.
55. Mastenbroek S, Twisk M, van der Veen F, Repping S. Preimplantation genetic screening: A systematic review and meta-analysis of RCTs. Hum Reprod Update 2011;17:454-66.