Blastocyst score, a blastocyst quality ranking tool, is a predictor of blastocyst ploidy and implantation potential

Qiansheng Zhan, M.D., Ph.D.,[a] E. T. Sierra, Ph.D.,[b] Jonas Malmsten, Ph.D.,[a] Zhen Ye, M.S.,[a] Zev Rosenwaks, M.D.,[a] and Nikica Zaninovic, Ph.D.[a]

[a] Ronald O. Perelman and Claudia Cohen Center for Reproductive Medicine, Weill Cornell Medicine, New York, New York; and [b] QED Analytics LLC, Princeton, New Jersey

Objective: To convert blastocyst (BL) morphological grade and BL day into a numeric blastocyst score (BS).
Design: Retrospective cohort study.
Setting: Academic center.
Patient(s): A total of 5,653 BL of known implantation (fetal heart, FH) and 11,348 biopsied BL.
Intervention(s): Based on their FH rates and/or significance, a score (1–4) was assigned to each BL grade component. The BL morphological score (BMS) is the sum (BS = BMS on day 5; BS = BMS + 2 on day 6).
Main Outcome Measure(s): Statistics characterized the FH and euploidy odds with BS.
Result(s): All three morphology grade components and BL day were associated with implantation and euploidy probability. The FH rate and euploidy odds decrease with larger BS. The BS was the most important factor (odds ratio [OR] per unit change = 0.807, 95% confidence interval [CI] 0.784, 0.831) for untested and euploid BL implantation, and the sole one for euploid BL (OR/unit change = 0.845, 95% CI 0.803, 0.889). The BS is the second most significant factor after maternal age for euploidy probability (OR/unit change = 0.808, 95% CI 0.795, 0.822). In training and validation sets (75:25), the BS can predict implantation with similar area under the curve [AUC] (training = 0.628, 95% CI 0.613, 0.643; validation = 0.606, 95% CI 0.581, 0.631). The BS has better euploidy prediction ability (training AUC = 0.683, 95% CI 0.673, 0.693; validation AUC = 0.698, 95% CI 0.681, 0.715). The BS can stratify BL into good (3–5), fair (6–9), and poor (10–14) groups, reflecting their FH, live birth rates, and ploidy status. Advanced maternal age was associated with lower untested BL implantation and lower euploidy odds across all groups.
Conclusion(s): The BS is a predictor of BL ploidy and FH implantation. (Fertil Steril Rep® 2020;1:133–41. ©2020 by American Society for Reproductive Medicine.)
Key Words: Embryo implantation, blastocyst, morphology, grade, preimplantation genetic testing

Although the use of assisted reproductive technology has helped many couples conceive, it has also increased multiple gestation rates (1, 2). This has led to a substantial increase in maternal and perinatal complications, along with an increase in health care costs (1, 2). Thus, a major goal in reproductive medicine is to improve embryo selection and maximize success rates after single embryo transfers.

The development of sophisticated embryo culture systems has led to the increasing utility of blastocyst (BL) transfers as an evolving tool to aid in embryo selection. As aneuploidy is the major cause of implantation failure and pregnancy loss, preimplantation genetic testing for aneuploidy (PGT-A) has also been shown to improve implantation rates and lower miscarriage rates in patients of advanced maternal age (3–6). With advancing molecular genetic techniques and improved embryo culture methods, current assisted reproductive technology practice is increasingly shifting to the use of BL culture and PGT-A, leading the way toward PGT-A freeze-all cycles

Received March 10, 2020; revised May 11, 2020; accepted May 12, 2020.
Q.Z. has nothing to disclose. E.T.S. has nothing to disclose. J.M. has nothing to disclose. Z.Y. has nothing to disclose. Z.R. has nothing to disclose. N.Z. has nothing to disclose.
Reprint requests: Qiansheng Zhan, M.D., Ph.D., Ronald O. Perelman and Claudia Cohen Center for Reproductive Medicine, Weill Cornell Medicine, 530 East 70th Street, M903, New York, New York, 10065 (E-mail: qiz2006@med.cornell.edu).

Fertil Steril Rep® Vol. 1, No. 2, September 2020 2666-3341

https://doi.org/10.1016/j.xfre.2020.05.004

(7). Despite the PGT-A randomized controlled trials in the past decades that reported improved ongoing pregnancy outcome, a recently published large Single Embryo TrAnsfeR of Euploid Embryo (STAR) trial (8) failed to find the outcome benefit from PGT-A versus morphology selection in good prognosis patients. The technique is costly, invasive (9, 10), and imperfect (11). Inaccurate PGT-A testing may be due to multiple factors, including the biological nature of embryos (i.e., mosaicism) and technical errors that result from imperfections in screening platforms (12–14). The PGT-A was actively marketed by overstating benefits and understating the losses of potential implantations, and should be applied selectively (15).

The two standard methods of noninvasive embryo selection are conventional embryo morphological assessment, which is performed mainly at the BL stage, and morphokinetic algorithm, which required expensive time-lapse microscopy (TLM) (16–20). Most BL morphological assessments are the Gardner system based, which grades expansion status, inner cell mass (ICM), and trophectoderm (TE) separately (21). Although BL morphological grade is widely used and has been shown to be associated with implantation potential (22–25), it has several limitations. First, embryo grading is subjective, with significant intrauser and interuser variability. Second, the combination of expansion, ICM, and TE grades introduces complexity when ranking blastocyst quality and assessing its prognostic value. Different approaches have been taken to simplify the system by reducing the number of expansion grades, rating the ICM or TE as good, fair, or poor (26), or by classifying BL into top, intermediate, and low quality groups (27, 28).

Because embryo morphology and morphokinetics are associated with implantation potential, they may also correlate with ploidy, as ploidy is the principal determinant of implantation. Although a few studies (19, 29–35) have shown some association between ploidy and embryo morphology and morphokinetic assessment, these techniques cannot yet replace PGT-A. Nonetheless, we should continue to improve our understanding of their relationship to ploidy.

In the present study we introduce a way to convert morphological grades and the rate of BL development (day 5/day 6, D5/D6) into a numeric BL score (BS), which can serve as a predictor of BL implantation potential and ploidy status. It can also stratify BL into three simple groups (good, fair, and poor) reflecting their implantation potential and ploidy status. The BS provides a quantitating way for the modified Gardner grading.

MATERIALS AND METHODS

Study design and participants

A retrospective cohort study at an academic institution from November 2011 through June 2017, included 5,653 BL of known implantation data (KID) (100% implantation rate/transfer and 0% implantation rate/transfer) for clinical implantation (fetal heart, FH) analysis and 11,348 biopsied BL for ploidy analysis. All BL were transferred fresh on D5, or frozen/biopsied-frozen on D5, D6, or D7. The D7 BL and unknown implantation (non-KID) BL were excluded from this study.

Ethical approval and clinical protocol

This retrospective cohort analysis was conducted following the research protocol approved by the Institutional Review Board of Weill Cornell Medicine (IRB #1304013779). Controlled ovarian hyperstimulation, the final oocyte maturation trigger, oocyte retrieval, and embryo transfer were performed per our standard protocols (36). Cryopreserved BLs were transferred in natural or programmed frozen-thawed embryo transfer cycles.

Fertilization, embryo culture, and cryopreservation

After retrieval, oocytes were fertilized by standard in vitro fertilization (IVF) or intracytoplasmic sperm injection (ICSI) depending on patient indications. Oocytes were individually loaded immediately after ICSI on D0 or after the fertilization check on D1. Embryos were cultured in a TLM system (EmbryoScope, Vitrolife) at 37°C, 5.8% CO_2, and 5% O_2 until reaching the BL stage (BL day) on D5 or D6. The culture slides (EmbryoSlide, Vitrolife) were filled with 25 μL of pre-equilibrated in-house sequential culture medium (cleavage medium, C1), which was replaced completely by BL medium (C2) on D3 and covered with tissue culture oil. Each batch of medium was vigorously tested by endotoxin, sperm survival test, mouse embryo test, and sibling oocyte test before clinical use. Embryos were vitrified and warmed using the Kitazato protocol (37).

Preimplantation genetic testing

The TE biopsy was performed after BL was graded on D5 (108–115 hours after insemination) or D6 (130–138 hours after insemination), and up to 10 cells were collected after zona pellucida dissection and hatching by laser (ZILOS-tk, Hamilton Thorne). Biopsied samples were amplified and analyzed by either array comparative genomic hybridization (BlueGnome 24sure; Illumina, single nucleotide polymorphism (Natera), or next-generation sequencing (NGS-low, Illumina).

BL grading

The grading system used is a modified Gardner system (38) in which there was one grade difference in expansion grade, and intermediate ICM and TE grades added (detailed in Supplemental Table 1, available online). Most BL were assessed by three embryologists, each of whom with >20 years' experience of the same grading system. The grades can be reviewed using the archived TLM images.

Numeric conversion of BL morphology grade

The grades of BL expansion, ICM, and TE were classified into four groups based on their KID FH rate similarity and/or statistical significance. A score from 1 (highest FH rate) to 4 (lowest FH rate) was assigned to each group (Table 1). For non-BL (cavitating morula and morula), a number 5 was assigned to each component. The BL morphology score (BMS) is the sum of the three scores for expansion, ICM, and TE. The BL score (BS) is the BMS integrated with the BL day (BS

TABLE 1

Numeric conversion of BL grade components based on their fetal heart rates.

Score	Expansion	FH rate	ICM	FH rate	TE	FH rate
1	6	54.4 (62/114)[a]	A	60.6 (406/670)[a]	A	65.3 (179/274)[a]
1	4	54.5 (24/44)[a]	A-	54.9 (625/1,138)[a]		
1	3	50.8 (1,882/3,708)[a]				
2	2-3	42.7 (167/391)[a]	B	45.5 (1,105/2,430)[b]	A-	55.4 (331/597)[a]
2	5	44.0 (11/25)[a]			B	50.6 (1,485/2,936)[a]
3	2	37.5 (276/736)[a]	B-	36.8 (429/1,166)[b]	B-	36.3 (525/1,448)[b]
4	1-2	35.3 (55/156)[a]	B-/C	23.8 (15/63)[c]	B-/C	26.2 (50/191)[c]
4	1	33.4 (109/326)[a]	C	18.2 (6/33)[c]	C	29.6 (16/54)[c]
5	Non-BL	13.7 (21/153)[b]	Non-BL	13.7 (21/153)[d]	Non-BL	13.7 (21/153) [c]

Note: Values with different superscripts[(a-d)] in the same column differ significantly ($P<.05$). n = 5,653.
BL = blastocyst; FH = fetal heart; ICM = inner cell mass; Non-BL = include morula and cavitating morula; TE = trophectoderm.

Zhan. Blastocyst quality ranking score. Fertil Steril Rep 2020.

= BMS in D5 BL; BS = BMS + 2 in D6 BL). The BS can further classify BL into three simple groups based on their similar FH rates of euploid BL: good ($3 \leq BS \leq 5$), fair ($6 \leq BS \leq 9$), and poor ($10 \leq BS \leq 14$) (39).

Clinical outcome measures

Clinical pregnancy was defined as the presence of FHs within 6–8 weeks after transfer. Live birth was defined as the proportion of transfers resulting in delivery beyond 24 weeks of gestation. The BS system was based on KID FH data.

Statistical analysis

Statistical analysis was performed using JMP Pro 14 software (SAS Institute Inc.). The χ^2 test or logistic regression analyses were performed with $P<.05$ considered statistically significant. Odds ratios (ORs) with 95% confidence intervals (CIs) were calculated. Overall implantation and ploidy data were randomly divided into training and validation sets by 75:25 using JMP formula random method. Receiver operating characteristic curve was calculated, and the area under the curve (AUC) with 95% CI was used to determine the predictive value.

RESULTS

The implantation analysis included 5,653 BL created from 3,247 cycles (2,734 ICSI and 513 standard IVF; patient age, 36.0 ± 4.9 [range, 20–47] years) and transferred in 4,477 fresh and frozen cycles. The PGT-A analysis included 11,348 biopsied BL from 2,306 cycles (2,258 ICSI and 48 standard IVF; patient age, 36.8 ± 4.7 [range, 21–48] years).

All grade components correlated with implantation and euploid probability

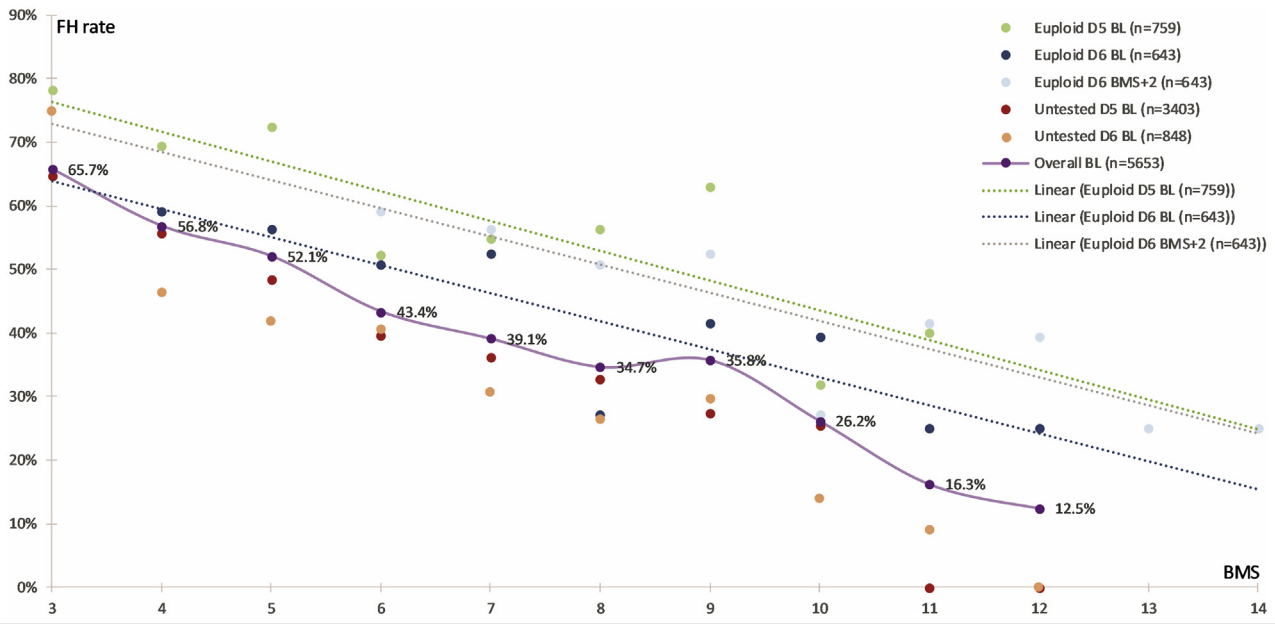
Table 1 summarizes the FH KID rate by each grade component. The FH rate increased significantly with better expansion and ICM or TE grades. Nominal logistic regression in overall (euploid and untested) BL (n = 5,653) demonstrated that the FH rate was affected by the following factors (in descending order of significance, all $P<.05$): euploid transfer, maternal age, TE, ICM, expansion, and BL day. The oocyte source (donor vs. autologous), fertilization method (ICSI vs. IVF), and episodes of direct uneven cleavage (DUC) were not significant factors. For euploid BL (n = 1,402) FH implantation, the BL day had the greatest impact, followed by ICM, expansion, and TE grades; however, maternal age became insignificant.

The ploidy status distributions by expansion, ICM, and TE grades were interesting (Supplemental Fig. 1, available online). The proportions of single chromosome aneuploidy and segmental abnormalities were relatively constant across all grades; only the ratio of euploid versus complex abnormalities significantly decreased reciprocally with lower grades, especially the TE grade. Logistic regression analysis of 11,348 BL showed that euploid probability was affected drastically by the following factors (in descending order of significance): maternal age at oocyte retrieval, expansion, TE grade, BL day, DUC, and ICM grade.

Development of the BMS, which is associated with FH implantation and euploidy probability

Based on the FH rates of each grade and/or statistical significance, the grades were divided into four groups. A score from 1–4 was assigned to each expansion, ICM, and TE groups separately. A score 5 was assigned to each grade component of morula and cavitating morula (Table 1). The BMS is the sum of three components scores. Figure 1 shows that the FH rate of overall BL gradually decreases from 65.7%–12.5% when the BMS increases from 3–12. Logistic regression of overall BL (n = 5,653) was affected by the following factors (in descending order of significance): BMS, euploidy transfer, maternal age, and BL day. For euploid BL transfer (n = 1,402), only the BMS and the BL day were significant factors. For euploid BL prediction (n = 11,348), the BMS was the second most significant factor after maternal age, followed by the BL day and DUC. Comparing to D6, D5 BL had higher FH rate in overall BL (OR = 1.343, 95% CI 1.17, 1.564) and in euploid BL (OR = 1.612, 95% CI 1.289, 2.016; all $P<.0001$). The D5 BL had higher euploidy odds (OR = 1.403, 95% CI 1.283, 1.534; $P<.0001$).

FIGURE 1



Fetal heart (FH) implantation rate by blastocyst (BL) morphology score (BMS). The FH rate in overall blastocysts gradually decreases from 65.7% to 12.5% with increasing BMS. The same trends were shown in day 5 (D5)/day 6 (D6) euploid/untested blastocysts. The trendline of D5 euploid blastocysts is near parallel to that of D6 euploid blastocysts. The linear trendline of euploid D6 BL with a 2 unit offset (euploid D6 BMS + 2) almost overlapped with the trendline of euploid D5 BL (euploid D5 BL).

Zhan. Blastocyst quality ranking score. Fertil Steril Rep 2020.

The BS, integration of the BMS and the BL day, is associated with FH implantation and euploidy probability

The BMS and the BL day were the only significant factors of euploid BL implantation, indicating that both are important measures of BL quality. Figure 1 shows the impact that the BL day has on the implantation of euploid and untested BL. We focused on euploid BL only as maternal age was not a significant factor. The near-parallel between the FH rate linear trend lines of D5 and D6 euploid BL (Fig. 1) suggests the possible integration of the BL day (D5/D6) with the BMS.

To find out the optimal score for BL day (D6), we tried different methods to modify the D6 score to merge both groups (D6/D6) into a single score, and then repeat the logistic regression to determine the significance of the BL day. After testing from 1–5, the BL day became insignificant after applying the number 2 (Table 2). In addition, the linear trendline of euploid D6 BL with a 2-unit offset almost overlapped with the trendline of euploid D5 BL (Fig. 1), indicating that number 2 is a reasonable correction on D6 BL to unify the scoring. The BS is defined as: BS = BMS for D5 BL; BS = BMS + 2 for D6 BL.

For overall BL FH implantation, the BS is the most significant factor (OR/unit change = 0.807, 95% CI 0.784, 0.831), followed by euploidy and maternal age (Table 2A). For euploid BL FH implantation, it is the only significant factor (OR/unit change = 0.845, 95% CI 0.803, 0.889; Table 2B). For BL euploidy odds, the BS is the second significant factor (OR/unit change = 0.808, 95% CI 0.795, 0.822) after maternal age (Table 2C). In the previous analyses, the impact of BL day became insignificant, demonstrating the successful integration of the BL day and BL morphology grade in BS. Figure 2 shows that the larger the BS, the lower the FH rate of euploid and untested BL, as well as the lower euploid probability.

BS is a predictor of BL implantation and BL euploidy probability

Implantation and ploidy data were sampled randomly into training and validation sets (75:25) by formula randomization. For the FH implantation prediction by BS, the AUC of 0.606 (95% CI 0.581, 0.631; n = 1,475) in validation is similar to that in training (0.628, 95% CI 0.613, 0.643; n = 4,178). For the euploidy prediction, the accuracy was similar in training (AUC = 0.683, 95% CI 0.673, 0.693; n = 8,539) and validation (AUC = 0.698, 95% CI 0.681, 0.715; n = 2,809). The BS has a better predictive ability for BL euploidy than for FH implantation.

Stratifying BL into simple groups by BS

Based on the similar FH incidence rates of euploid BL shown in Figure 2, the BS can stratify BL into three simple groups: good (3 ≤ BS ≤ 5, average FH rate = 71.2%), fair (6 ≤ BS ≤ 9, average FH rate = 54.8%), and poor (10 ≤ BS ≤ 14, average FH rate = 33.3%). Supplemental Figure 2, available online, shows the descending FH and live birth rates of

TABLE 2

Nominal logistic regression: fetal heart implantation, euploidy versus blastocyst score.

Analysis	Effect summary			ORs for odds of positive versus negative				
A	**Source**	**LogWorth**	***P* value**	**Level 1**	**/Level 2**	**OR**	**Lower 95% CI**	**Upper 95% CI**
Overall (euploid and untested)	BS[a]	52.828	.00000[a]	Per unit change		0.807	0.784	0.831
Blastocyst implantation (FH)				Entire range (3–15)		0.077	0.054	0.108
N = 5,653	Euploidy[a]	34.047	.00000[a]	Euploid	Untested	2.391	2.076	2.753
	Maternal age[a]	18.669	.00000[a]	Per year change		0.940	0.927	0.953
				Entire range (20–47)		0.186	0.129	0.270
	Blastocyst day	0.994	.10144	D5	D6	0.878	0.751	1.026
B								
Euploid blastocyst implantation (FH)	BS[a]	10.699	.00000[a]	Per unit change		0.845	0.803	0.889
				Entire range (3-15)		0.133	0.072	0.244
N = 1,402	Blastocyst day	0.602	.25005	D5	D6	1.163	0.899	1.505
C								
Blastocyst euploidy odds N = 11,348	Maternal age	242.788	.00000[a]	Per year change		0.856	0.848	0.865
				Entire range (21–48)		0.015	0.012	0.020
	BS[a]	152.700	.00000[a]	Per unit change		0.808	0.795	0.822
				Entire range (3–15)		0.078	0.064	0.095
	Direct uneven cleavage[a]	2.468	.00340[a]	Non-DUC	DUC	0.814	0.710	0.934
	Blastocyst day	0.737	.18312	D5	D6	0.933	0.842	1.033

Note: The same variables in analysis A, B, and C were BS, maternal age, oocyte source (donor/autologous), fertilization method (IVF/ICSI), and direct uneven cleavage. Euploidy (euploid/untested) was included in analysis A only.
Except for blastocyst day, only the variables listed were significant. BS = blastocyst score; CI = confidence interval; D5/D6 = day 5 or day 6; DUC = direct uneven cleavage; FH = fetal heart; ICSI = intracytoplasmic sperm injection; IVF = in vitro fertilization; OR = odds ratio.
[a] $P < .05$.

Zhan. Blastocyst quality ranking score. Fertil Steril Rep 2020.

euploid and untested BL from the good to non-BL groups. With advancing maternal age, the implantation rates were descending across all groups of untested BL (Supplemental Fig. 2, right). This declining trend disappeared in euploid BL (Supplemental Fig. 2, left). The ploidy ratios (euploid/complex abnormalities) followed the same trend among different maternal ages and BL groups (Supplemental Fig. 3, available online).

DISCUSSION

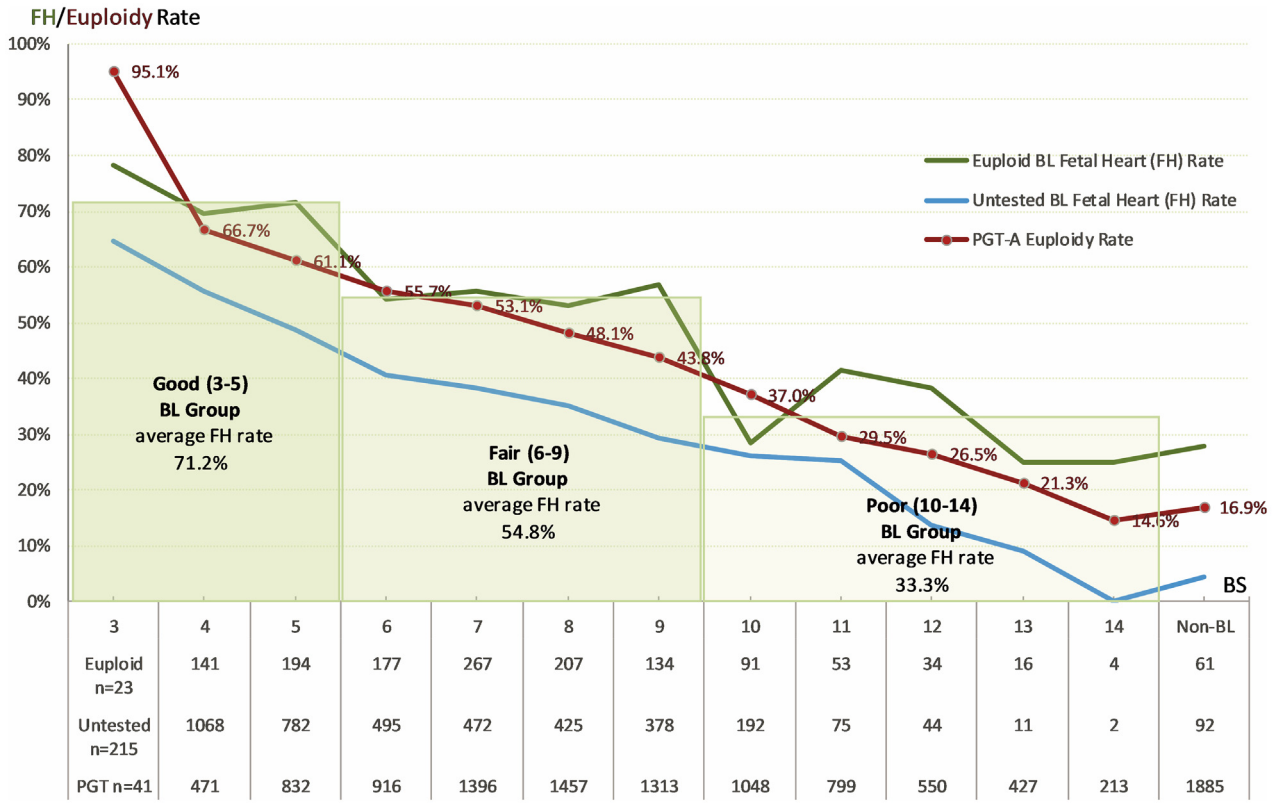
To our knowledge this study comprises the largest sample size ever used to explore the relationship between BL morphology grade, BL day, implantation, and ploidy status. We developed a numeric ranking system, the BS, successfully integrating the BL morphological grade with developmental speed, which is a predictor of BL implantation and euploidy.

Various studies have demonstrated that only the expansion (23, 24), or the ICM grade (36, 40–42), or the TE grade (23, 25, 43) is associated with pregnancy outcome, which was consistent with our finding. Besides that, all three components are significant implantation factors (Table 1). The TE grade has a higher predictive strength than the ICM grade for BL implantation, indicating the crucial function of the TE (26, 29). On a cautionary note, these data could be influenced by selection bias, as BL with a "C" ICM grade is rarely chosen for transfer in our practice. The lack of data on embryos exhibiting poor ICM grade BL could also affect the statistical significance of the ICM grade (7). Oocyte source, fertilization method, and DUC did not appear to have an impact on BL implantation, confirming that DUC only impacts embryo implantation in early stage but not blastocyst stage (44).

In prior published studies (30, 32) of euploid BL transfers, with limited sample sizes, BL morphological features and the BL day were not associated with implantation potential. However, the BL day is the strongest predictor of euploid BL implantation, followed by ICM, expansion, and TE grades, supporting our previous finding (45). The value of traditional morphological assessment should not be overlooked, even for PGT-A BL selection. Maternal age is the strongest predictor of untested BL implantation, but insignificant in euploid BL, suggesting that the effect of maternal age on implantation is contributed mainly by ploidy status. Our previous study (36) and another report (32) indicate that slow-growing BL, which has higher aneuploidy and miscarriage rates, could be euploid and result in live birth. Although the effect of maternal age on euploid BL implantation is not significant, none of the euploid non-BL (mostly morula) embryos from patients older than 40 years implanted in this study (Supplemental Fig. 2, left).

Given that embryo ploidy is the principal determinant of implantation potential, the correlation between BL morphology and implantation suggests a relationship between BL morphology and ploidy as expected. Several studies (29–32) of smaller sample sizes have reported some association between BL morphological grade and aneuploidy. We, on the other hand, confirmed a strong correlation between the likelihood of euploidy and BL morphological features. Similar to another study (43), the TE grade has greater predictive power than the ICM grade.

FIGURE 2



Fetal heart (FH) implantation rate and euploidy rate by blastocyst (BL) score (BS). The FH rate in untested/euploid blastocysts and the euploidy rate in biopsied BLs gradually decrease with larger BS. Based on the similar FH rate of euploid BL, BLs can be further stratified into three groups: good ($3 \leq BS \leq 5$; average FH rate, 71.2%), fair ($6 \leq BS \leq 9$; average FH rate, 54.8%), and poor ($10 \leq BS \leq 14$; average FH rate, 33.3%). PGT-A = preimplantation genetic testing for aneuploidy.

Zhan. Blastocyst quality ranking score. Fertil Steril Rep 2020.

Some possible explanations include: [1] only TE cells were biopsied and analyzed for PGT-A. The ploidy result was only a direct reflection of TE ploidy, and may not have been concordant with the ICM due to the relatively high prevalence of mosaicism in human embryos, and perhaps even due to technical errors (12–14). [2] TE grading is relatively easier compared with ICM, which may lead to fewer variations between observers. The BL grading is subjective and reflected interobserver and intraobserver variance (7).

Although the Gardner-based BL grading is widely used, the complexity of grade combinations and the lack of numeric scores with which to assign ranking diminishes its clinical value. Some attempts have been made to address these limitations. One is a simplification of the BL grade into good, fair, and poor categories (28). The other is numeric transformation. In previous reports, the morphological grades were converted into numbers and a BL quality score (BQS; score 1–54, total of 17 scores) was defined as the production of multiplication (46). In addition, the embryo progression index (EPI) was calculated as the AUC of observed or estimated total cell number, which was estimated by the BL grade (46). This mean BQS was the stronger predictor of outcome than the mean EPI (46). Although the BQS was a useful attempt at creating a BL quality ranking system, there were some limitations, as follows: [1] the number assignment for each grade is not evidence-based; [2] the BQS was calculated by multiplication instead of addition, resulting in a wider spread of values between lower and higher quality embryos, which may also result in the nonlinear correlation with implantation; [3] the BL day was not taken into consideration; [4] the calculation is quite complicated (the EPI is AUC of estimated total cell number curve, which needs adjusting by BL grade); [5] the smaller sample size ($n = 777$); [6] the use of the clinical pregnancy rate per transfer instead of the real BL implantation rate; and [7] the use of the mean BQS and EPI instead of the individual value of BL due to multiple transfers undermined the conclusions.

A viable embryo will reach full expansion with a prominent ICM and TE, consisting of many cells approximately at 116 ± 2 hours (D5) after insemination (7). The BL day can influence the BL ranking order, and its impact on implantation is conflicting. Recent studies indicate that D5 BL may be superior to D6 BL, although this difference disappeared in a euploid transfer (30). On the contrary, this study confirmed

our previous findings (45), indicating the superiority of D5 versus D6 in terms of overall BL implantation, euploid BL implantation (Fig. 1), and odds of euploidy (OR = 1.403).

In the present study we developed a protocol to convert BL grades into a numeric BL score. The keys to its successful development were, as follows: [1] the large sample size of our study (n = 5,653); [2] KID FH rate were used instead of pregnancy rates per transfer used in BQS, which is a better reflection of real implantation potential of BL (46); [3] scores were evident-based on the implantation range of each grade and/or statistical significance, which means that fine-tuning the protocol is possible to accommodate different grading systems and lab settings; [4] addition instead of multiplication for calculation, which contributes to the linearity; and [5] the impact of the BL day was considered and quantified successfully (2 points offset for D6). Evidence of our success is suggested by: [1] nearly linear correlation of BS with euploidy probability (Fig. 2); [2] nearly linear correlation with FH rates of untested and euploid BL (Fig. 2); [3] for implantation, the BS appears to be the most influential factor of overall BL and the sole factor of euploid BL (the effect of the BL day was insignificant [Table 2]); [4] the BS can predict FH implantation with AUC between 0.606 and 0.628, which is acceptable because the embryo quality is not the sole factor involved in implantation (e.g., uterine factor, ploidy status); and [5] the BS has better euploidy prediction ability (AUC = 0.683–0.698), indicating that ploidy is the main cause of the 10%–15% FH rate difference between euploid and untested BL (Fig. 2).

Clinical application

As a numeric measure of BL quality, the BS might also serve as a key performance indicator in an IVF laboratory quality control. The ability to stratify BL into simple groups, provide an approach for comparing results from different grading system. The linear correlation with implantation indicates that the BS is a valuable selection ranking tool in untested BL transfer, and is also critical for euploid BL transfer. Patients who used PGT-A will benefit from transferring the best quality BL selected by BS. The BS has a better ability to predict euploidy. The BS could be a potential PGT-A consultation tool. We could develop BS-based algorithms as noninvasive BL evaluation tools to avoid costly, invasive PGT-A in certain BL. Besides morphology, TLM provides another noninvasive assessment approach (16–20). The BL development is the outcome indicator in many morphokinetic models. However, the complexity and subjectivity of the BL grade are the major obstacles hindering the establishment of universal TLM models. The rapid development of image analysis using machine learning has recently provided a way to standardize and quantitate BL assessment objectively (47–49). However, this technique requires an enormous amount of training data (quantified images tagged with BL features). With the help of BS to convert and quantify existing BL grade data, standardized automatic BL assessment is possible. In the not distant future, combinations of machine learning, TLM, clinical and laboratory big data, and embryonic genomic data will make it possible to understand the critical parameters of assisted reproduction and to provide targeted and personalized treatments.

Limitations

The major drawback of this study is its retrospective nature and its use of KID due to double embryo transfers. Ideally, including only single embryo tranfers (SETs) and cycles from different patients could avoid the potential bias introduced by KID (50). However, if this dataset were limited to SET only, 4,477 BL remains; if excluding both, a maximum of 3,249 BL remains. Another problem is the very uneven distribution between different BS groups. For a BS between 12 and 14, the sample size ranges from 2 to 44. The impact of selection bias due to the lack of poor-quality BL could be exaggerated by reducing the total sample size (zero sample in BS = 14 if only 4,477 SET included; zero sample in BS = 13 or 14 if both excluded [n=3,249]). A further prospective SET study using the BS will provide more detailed analysis and allow for optimization of the score. Although the BS is a numeric conversion of the BL grade and the BL day, it is still a subjective measure. Also, many factors involved in the implantation (e.g., endometrium) were not included in the present study, which needs further investigation.

Nevertheless, we have described a numeric transformation system (the BS) that successfully integrates the BL morphology grade and the BL day, with a highly linear correlation with BL implantation and euploidy rate and ability to predict BL FH implantation and euploidy. As a ranking order tool, the BS has potential applications in BL selection and research. Also, it has the flexibility to stratify BL into simple groups.

In conclusion the three BL morphological grade components and the BL day were correlated with BL implantation and euploid probability, which can be converted successfully into a numeric BS. The BS is a predictor for BL FH implantation and ploidy status. It can stratify BL into simple groups.

Acknowledgments: The authors thank the clinical staff and embryology laboratory team of the Ronald O. Perelman and Claudia Cohen Center for Reproductive Medicine, Weill Cornell Medicine. We are especially thankful to Teja Kavi and Alexandra MacWade for editorial assistance.

REFERENCES

1. Kulkarni AD, Jamieson DJ, Jones HW Jr, Kissin DM, Gallo MF, Macaluso M, et al. Fertility treatments and multiple births in the United States. N Engl J Med 2013;369:2218–25.
2. Lemos EV, Zhang D, van Voorhis BJ, Hu XH. Healthcare expenses associated with multiple vs singleton pregnancies in the United States. Am J Obstet Gynecol 2013;209:586.e1–11.
3. Yang Z, Liu J, Collins GS, Salem SA, Liu X, Lyle SS, et al. Selection of single blastocysts for fresh transfer via standard morphology assessment alone and with array CGH for good prognosis IVF patients: results from a randomized pilot study. Mol Cytogenet 2012;5:24.
4. Scott RT Jr, Upham KM, Forman EJ, Hong KH, Scott KL, Taylor D, et al. Blastocyst biopsy with comprehensive chromosome screening and fresh embryo transfer significantly increases in vitro fertilization implantation and delivery rates: a randomized controlled trial. Fertil Steril 2013;100:697–703.
5. Rubio C, Bellver J, Rodrigo L, Castillon G, Guillen A, Vidal C, et al. In vitro fertilization with preimplantation genetic diagnosis for aneuploidies in

advanced maternal age: a randomized, controlled study. Fertil Steril 2017;107:1122–9.
6. Forman EJ, Hong KH, Ferry KM, Tao X, Taylor D, Levy B, et al. In vitro fertilization with single euploid blastocyst transfer: a randomized controlled trial. Fertil Steril 2013;100:100–107 e1.
7. Morbeck DE. Blastocyst culture in the Era of PGS and FreezeAlls: Is a 'C' a failing grade? Hum Reprod Open 2017;2017(3):hox017.
8. Munne S, Kaplan B, Frattarelli JL, Child T, Nakhuda G, Shamma FN, et al. Preimplantation genetic testing for aneuploidy versus morphology as selection criteria for single frozen-thawed embryo transfer in good-prognosis patients: a multicenter randomized clinical trial. Fertil Steril 2019;112:1071–9.e7.
9. Scott RT Jr, Upham KM, Forman EJ, Zhao T, Treff NR. Cleavage-stage biopsy significantly impairs human embryonic implantation potential while blastocyst biopsy does not: a randomized and paired clinical trial. Fertil Steril 2013;100:624–30.
10. Ten JR-AA, Diaz MC, Blanca H, Guerrero J, Lledo B, et al. Number of trophoectoderm cells removed for biopsy is correlated with first trimester miscarriage. Hum Reprod 2015;30:i1–501.
11. Gleicher N, Orvieto R. Is the hypothesis of preimplantation genetic screening (PGS) still supportable? A review. J Ovarian Res 2017;10:21.
12. Capalbo A, Treff NR, Cimadomo D, Tao X, Upham K, Ubaldi FM, et al. Comparison of array comparative genomic hybridization and quantitative real-time PCR-based aneuploidy screening of blastocyst biopsies. Eur J Hum Genet 2015;23:901–6.
13. Vera-Rodriguez M, Rubio C. Assessing the true incidence of mosaicism in preimplantation embryos. Fertil Steril 2017;107:1107–12.
14. Goodrich D, Tao X, Bohrer C, Lonczak A, Xing T, Zimmerman R, et al. A randomized and blinded comparison of qPCR and NGS-based detection of aneuploidy in a cell line mixture model of blastocyst biopsy mosaicism. J Assist Reprod Genet 2016;33:1473–80.
15. Paulson RJ. Hidden in plain sight: the overstated benefits and underestimated losses of potential implantations associated with advertised PGT-A success rates. Hum Reprod 2020;35:490–3.
16. Meseguer M, Herrero J, Tejera A, Hilligsoe KM, Ramsing NB, Remohi J. The use of morphokinetics as a predictor of embryo implantation. Hum Reprod 2011;26:2658–71.
17. Wong CC, Loewke KE, Bossert NL, Behr B, de Jonge CJ, Baer TM, et al. Non-invasive imaging of human embryos before embryonic genome activation predicts development to the blastocyst stage. Nat Biotechnol 2010;28:1115–21.
18. Liu Y, Chapple V, Feenan K, Roberts P, Matson P. Time-lapse deselection model for human day 3 in vitro fertilization embryos: the combination of qualitative and quantitative measures of embryo growth. Fertil Steril 2016;105:656–62.e1.
19. Campbell A, Fishel S, Bowman N, Duffy S, Sedler M, Hickman CF. Modelling a risk classification of aneuploidy in human embryos using non-invasive morphokinetics. Reprod Biomed Online 2013;26:477–85.
20. 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;31:2231–44.
21. Gardner DK, Schoolcraft WB. In Vitro Culture of Human Blastocyst. In: Jansen R, Mortimer D, editors. Towards Reproductive Certainty: Infertility and Genetics Beyond. Carnforth: Parthenon Press; 1999:377–88.
22. Gardner DK, Schoolcraft WB, Stevens J, Schlenker T, Schoolcraft WB. Blastocyst score affects implantation and pregnancy outcome: towards a single blastocyst transfer. Fertil Steril 2000;73:1155–8.
23. Ahlstrom A, Westin C, Wikland M, Hardarson T. Prediction of live birth in frozen-thawed single blastocyst transfer cycles by pre-freeze and post-thaw morphology. Hum Reprod 2013;28:1199–209.
24. Van den Abbeel E, Balaban B, Ziebe S, Lundin K, Cuesta MJ, Klein BM, et al. Association between blastocyst morphology and outcome of single-blastocyst transfer. Reprod Biomed Online 2013;27:353–61.
25. 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.
26. 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.
27. Wirleitner B, Schuff M, Stecher A, Murtinger M, Vanderzwalmen P. Pregnancy and birth outcomes following fresh or vitrified embryo transfer according to blastocyst morphology and expansion stage, and culturing strategy for delayed development. Hum Reprod 2016;31:1685–95.
28. Richardson A, Brearley S, Ahitan S, Chamberlain S, Davey T, Zujovic L, et al. A clinically useful simplified blastocyst grading system. Reprod Biomed Online 2015;31:523–30.
29. Alfarawati S, Fragouli E, Colls P, Stevens J, Gutierrez-Mateo C, Schoolcraft WB, et al. The relationship between blastocyst morphology, chromosomal abnormality, and embryo gender. Fertil Steril 2011;95:520–4.
30. Capalbo A, Rienzi L, Cimadomo D, Maggiulli R, Elliott T, Wright G, et al. Correlation between standard blastocyst morphology, euploidy and implantation: an observational study in two centers involving 956 screened blastocysts. Hum Reprod 2014;29:1173–81.
31. Fragouli E, Alfarawati S, Spath K, Wells D. Morphological and cytogenetic assessment of cleavage and blastocyst stage embryos. Mol Hum Reprod 2014;20:117–26.
32. Minasi MG, Colasante A, Riccio T, Ruberti A, Casciani V, Scarselli F, et al. Correlation between aneuploidy, standard morphology evaluation and morphokinetic development in 1730 biopsied blastocysts: a consecutive case series study. Hum Reprod 2016;31:2245–54.
33. Campbell A, Fishel S, Bowman N, Duffy S, Sedler M, Thornton S. Retrospective analysis of outcomes after IVF using an aneuploidy risk model derived from time-lapse imaging without PGS. Reprod Biomed Online 2013;27:140–6.
34. Mumusoglu S, Yarali I, Bozdag G, Ozdemir P, Polat M, Sokmensuer LK, et al. Time-lapse morphokinetic assessment has low to moderate ability to predict euploidy when patient- and ovarian stimulation-related factors are taken into account with the use of clustered data analysis. Fertil Steril 2017;107:413–21.e4.
35. Baltaci V, Satiroglu H, Kabukcu C, Unsal E, Aydinuraz B, Uner O, et al. Relationship between embryo quality and aneuploidies. Reprod Biomed Online 2006;12:77–82.
36. Irani M, Reichman D, Robles A, Melnick A, Davis O, Zaninovic N, et al. Morphologic grading of euploid blastocysts influences implantation and ongoing pregnancy rates. Fertil Steril 2017;107:664–70.
37. Cobo A, de los Santos MJ, Castello D, Gamiz P, Campos P, Remohi J. Outcomes of vitrified early cleavage-stage and blastocyst-stage embryos in a cryopreservation program: evaluation of 3,150 warming cycles. Fertil Steril 2012;98:1138–46.e1.
38. Veeck LL, Zaninovic N. An atlas of human blastocysts 2003. Boca Raton, Fl: CRC Press; 2003.
39. Zhan Q, Clarke RN, Ye Z, Rosenwaks Z, Zaninovic N. Blastocyst score is strongly correlated with implantation outcome and ploidy. Fertil Steril 2015;104:e274.
40. Almagor M, Harir Y, Fieldust S, Or Y, Shoham Z. Ratio between inner cell mass diameter and blastocyst diameter is correlated with successful pregnancy outcomes of single blastocyst transfers. Fertil Steril 2016;106:1386–91.
41. Richter KS, Harris DC, Daneshmand ST, Shapiro BS. Quantitative grading of a human blastocyst: optimal inner cell mass size and shape. Fertil Steril 2001;76:1157–67.
42. Subira J, Craig J, Turner K, Bevan A, Ohuma E, McVeigh E, et al. Grade of the inner cell mass, but not trophectoderm, predicts live birth in fresh blastocyst single transfers. Hum Fertil (Camb) 2016;19:254–61.
43. Ahlstrom A, Westin C, Reismer E, Wikland M, Hardarson T. Trophectoderm morphology: an important parameter for predicting live birth after single blastocyst transfer. Hum Reprod 2011;26:3289–96.
44. Zhan Q, Ye Z, Clarke R, Rosenwaks Z, Zaninovic N. Direct unequal cleavages: embryo developmental competence, genetic constitution and clinical outcome. PloS One 2016;11:e0166398.
45. Irani M, O'Neill C, Palermo GD, Xu K, Zhang C, Qin X, et al. Blastocyst development rate influences implantation and live birth rates of similarly graded euploid blastocysts. Fertil Steril 2018;110:95–102.e1.

46. 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.
47. Paternot G, Debrock S, De Neubourg D, D'Hooghe TM, Spiessens C. Semi-automated morphometric analysis of human embryos can reveal correlations between total embryo volume and clinical pregnancy. Hum Reprod 2013;28:627–33.
48. Paternot G, Debrock S, D'Hooghe T, Spiessens C. Computer-assisted embryo selection: a benefit in the evaluation of embryo quality? Reprod Biomed Online 2011;23:347–54.
49. Rocha JC, Passalia FJ, Matos FD, Takahashi MB, Ciniciato DS, Maserati MP, et al. A method based on artificial intelligence to fully automatize the evaluation of bovine blastocyst images. Sci Rep 2017;7:7659.
50. Liu Y, Feenan K, Chapple V, Matson P. Assessing efficacy of day 3 embryo time-lapse algorithms retrospectively: impacts of dataset type and confounding factors. Hum Fertil (Camb) 2019;22:182–90.