Spontaneous Reduction in ART-Conceived Twin Pregnancies: Risk Factors, Perinatal Outcomes, and a Prediction Model

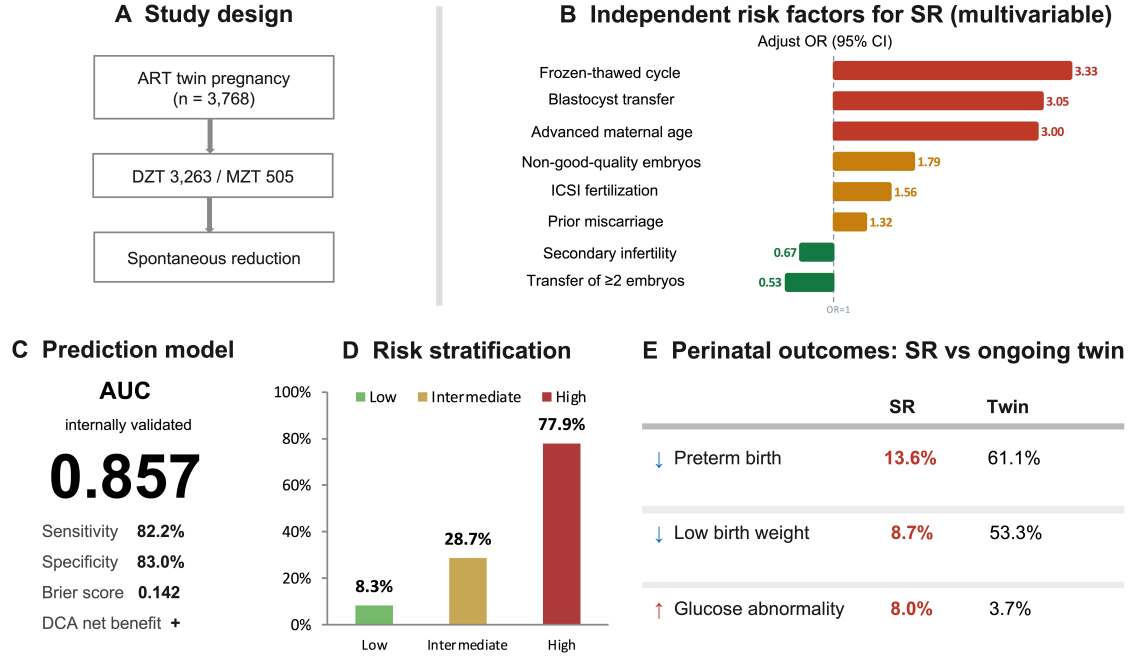
Authors

Shuqi Cai, Zhaolian Wei

Correspondence

weizhaolian@ahmu.edu.cn (Z. Wei)

Graphical Abstract



https://doi.org/10.71321/mapcd847

Spontaneous Reduction in ART-Conceived Twin Pregnancies: Risk Factors, Perinatal Outcomes, and a Prediction Model

Shuqi Cai [1], Zhaolian Wei [1,2,3*]

Received: 2026-05-30 | Accepted: 2026-05-31 | Published online: 2026-05-31

Abstract

Objective: To investigate the risk factors for spontaneous reduction (SR) in assisted reproductive technology (ART)-conceived twin pregnancies and compare perinatal outcomes between SR and non-SR pregnancies, and to develop and validate an SR risk prediction model in the dizygotic twin (DZT) cohort.

Methods: This single-center retrospective cohort study enrolled patients who achieved clinical twin pregnancies following ART at the Reproductive Center of the First Affiliated Hospital of Anhui Medical University between January 2015 and December 2024. Participants were classified into monozygotic twins (MZT, n=505) and dizygotic twins (DZT, n=3,263) based on the number of gestational sacs and fetal heartbeats detected on early pregnancy ultrasound, and each cohort was further divided into SR and Non-SR groups. Baseline characteristics, treatment parameters, and perinatal outcomes were compared between groups. Univariable and multivariable logistic regression analyses were performed with SR as the dependent variable, incorporating nine core predictors. A predictive model for SR was constructed in the DZT cohort and internally validated; model performance was evaluated using the receiver operating characteristic (ROC) curve with area under the curve (AUC), calibration curve, Brier score, and decision curve analysis (DCA). Risk stratification was performed by tertiles of predicted probability.

Results: In the DZT cohort, the SR group had significantly older maternal age, a higher proportion of advanced maternal age (≥35 years), and higher rates of blastocyst-stage transfer, intracytoplasmic sperm injection (ICSI), and frozen-thawed embryo transfer cycles than the Non-SR group (all $P<0.05$). No statistically significant differences were observed in any of these characteristics in the MZT cohort. Perinatal outcome trends were consistent across both cohorts: the SR group had later gestational age at delivery, lower rates of preterm birth and low birth weight, but higher rates of glucose metabolism abnormalities and lower rates of premature rupture of membranes. Multivariable logistic regression identified advanced maternal age (aOR=3.00), history of previous miscarriage (aOR=1.32), blastocyst-stage transfer (aOR=3.05), non-good-quality embryos (aOR=1.79), ICSI (aOR=1.56), and frozen-thawed cycles (aOR=3.33) as independent risk factors for SR, whereas secondary infertility (aOR=0.67) and transfer of ≥2 embryos (aOR=0.53) were associated with a reduced risk. The DZT-SR prediction model achieved an AUC of 0.857 on internal validation, with acceptable calibration and favorable net clinical benefit on DCA. SR incidence increased progressively across low-, intermediate-, and high-risk strata at 8.3%, 28.7%, and 77.9%, respectively.

Conclusion: SR in ART-conceived twin pregnancies is closely associated with maternal age and multiple embryo- and cycle-related factors. The DZT-SR prediction model, built on nine routinely available clinical variables, demonstrated good internal validation performance and may serve as a practical tool for early pregnancy risk stratification and individualized follow-up planning.

Keywords: Assisted reproductive technology; Twin pregnancy; Spontaneous reduction; Logistic regression; Prediction model

Introduction

According to the World Health Organization (WHO), approximately one in six people worldwide experiences infertility at some point during their lifetime [1]. The widespread adoption of assisted reproductive technology (ART) has substantially improved pregnancy outcomes for affected couples; however, the use of ovarian stimulation and multiple embryo transfer has led to a higher rate of iatrogenic multiple pregnancy compared with natural conception, making it one of the most common complications associated with ART [2-3]. Multiple pregnancy carries considerable maternal and neonatal risks: compared with singleton pregnancy, the risks of preterm birth, preterm birth before 32 weeks, stillbirth, and neonatal death are increased approximately 6-, 13-, 5-, and 7-fold, respectively, and the rates of severe maternal morbidity and mortality are

1 Department of Obstetrics and Gynecology, First Affiliated Hospital of Anhui Medical University, No 218 Jixi Road, Hefei, China.

2 Anhui Province Key Laboratory of Reproductive Health and Genetics, Anhui Medical University, No 81 Meishan Road, Hefei 230032, China.

3 NHC Key Laboratory of study on abnormal gametes and reproductive tract (Anhui Medical University), No 81 Meishan Road, Hefei 230032, China.

* Corresponding Author.

also markedly elevated [4-5]. In response, reproductive medicine societies worldwide have issued guidelines recommending a reduction in the number of embryos transferred and encouraging the adoption of elective single embryo transfer (eSET) to reduce multiple pregnancy rates [2-3, 6-7].

The routine use of first-trimester ultrasound has made vanishing twin syndrome (VTS) — or spontaneous reduction (SR) — increasingly recognizable in clinical practice. SR refers to the arrest of development of one or more embryos in an intrauterine multiple pregnancy that has already been confirmed clinically. In ART-conceived twin pregnancies, the reported incidence of SR ranges from approximately 12% to 38% [8-10]. Its etiology remains incompletely understood, but proposed contributing factors include abnormal embryonic development, suboptimal placental implantation, and an unfavorable intrauterine environment [9, 11-12]; maternal age, chorionicity, and certain ART-related procedures may also play a role [13-16], though the available evidence is limited and findings across studies are inconsistent.

The impact of SR on pregnancy outcomes also warrants attention. Perinatal outcomes in pregnancies that undergo SR and continue as singletons appear to fall somewhere between those of pregnancies that were singletons from the outset and those that remain twin pregnancies, but existing studies are limited by small sample sizes and heterogeneous findings [8, 10]. Furthermore, the clinical patterns and implications of SR may differ between monozygotic twins (MZT) and dizygotic twins (DZT), yet systematic stratified analyses are lacking, and data from Chinese populations remain scarce.

Against this background, the present study compared baseline characteristics, treatment parameters, and perinatal outcomes between SR and non-SR pregnancies in both MZT and DZT cohorts, examined the risk factors associated with SR, and developed and internally validated an SR risk prediction model within the DZT cohort, with the aim of informing early pregnancy risk assessment and follow-up management.

Materials and Methods

Study population

This was a single-center retrospective cohort study. We enrolled patients who achieved clinical pregnancy following ART at the Reproductive Center of the First Affiliated Hospital of Anhui Medical University between January 2015 and December 2024, and whose twin pregnancy was confirmed by transvaginal ultrasound in the first trimester. Participants were classified into MZT and DZT based on the number of gestational sacs and fetal heartbeats on early ultrasound, and each cohort was divided into an SR group and a Non-SR group according to whether spontaneous reduction occurred during first-trimester follow-up.

Definitions and diagnostic criteria

Twin pregnancy was confirmed approximately 30 days after embryo transfer by transvaginal ultrasound showing either two separate gestational sacs or a single sac containing two fetal heartbeats. MZT was defined as a single gestational sac with two fetal heartbeats, and DZT as two gestational sacs each with a fetal heartbeat. SR was defined as a spontaneous decrease in the number of fetal heartbeats from two to one on routine follow-up ultrasound in a previously confirmed twin pregnancy; the date of the first ultrasound documenting this change was recorded as the time of SR. Non-SR was defined as the absence of any such reduction in fetal heartbeats throughout the first trimester, with subsequent delivery of one or two live infants.

Inclusion and exclusion criteria

Patients were included if they: (1) achieved clinical pregnancy following ART with first-trimester ultrasound confirmation of a twin pregnancy (MZT or DZT); and (2) were followed through to delivery with complete obstetric and neonatal records.

Patients were excluded if they: (1) underwent multifetal pregnancy reduction; (2) had recurrent miscarriage or structural abnormalities of the reproductive tract; (3) had the pregnancy terminated due to systemic illness, uterine scarring, or personal reasons; (4) had pre-existing hypertension or diabetes mellitus before pregnancy; (5) conceived through donor oocytes, donor sperm, intrauterine insemination, or preimplantation genetic testing cycles; (6) were in a consanguineous relationship; (7) had a severe psychiatric disorder; or (8) had incomplete clinical records.

Embryo quality grading

Cleavage-stage embryos (D2/D3) were assessed morphologically according to the Istanbul Consensus and the ESHRE-ALPHA scoring criteria [17-19], taking into account blastomere number, fragmentation, and symmetry. On day 3, embryos meeting specific score combinations — such as 8CF4S1 — were classified as good-quality. Blastocysts (D5/D6) were graded using the Gardner-Schoolcraft system [18-23]. Day-5 blastocysts expanded to stage ≥3 with inner cell mass (ICM) and trophectoderm (TE) both graded A or B (i.e., AA, AB, BA, or BB) were considered good-quality; for day-6 blastocysts, the same grading criteria applied with a minimum expansion stage of 4. When all transferred embryos were good-quality, the cycle was categorized as the "good-quality" group; if any transferred embryo was below this threshold, the cycle was assigned to the "non-good-quality" group.

Study variables

Baseline and treatment characteristics collected included maternal age, body mass index (BMI), history of previous miscarriage, infertility type (primary or secondary), cycle type (fresh or frozen-thawed), fertilization method (conventional IVF or ICSI), embryonic developmental stage at transfer (cleavage-stage D2/D3 or blastocyst D5/D6), number of embryos transferred (1 or ≥2), embryo quality (good-quality or non-good-quality), and the underlying causes of infertility.

Perinatal outcomes included mode of delivery, gestational age at delivery, neonatal birth weight, presence of any low birth weight (any LBW, defined as at least one infant weighing <2,500 g), and pregnancy complications including hypertensive disorders of pregnancy, glucose metabolism abnormalities, and premature rupture of membranes.

Nine variables were entered into the logistic regression analyses as core predictors: advanced maternal age (≥35 years), maternal BMI, infertility type, history of previous miscarriage, embryonic developmental stage, number of embryos transferred, embryo quality, fertilization method, and cycle type.

Table 1. Baseline and treatment characteristics of the DZT group.

Variable	Non-SR(n=2456)	SR(n=807)	*P* value
Maternal age (years)	29.99 ± 3.67	31.86 ± 4.25	< 0.001
BMI(kg/m²)	22.35 ± 3.14	22.61 ± 3.26	0.052
No. of previous miscarriages	0.45 ± 0.80	0.51 ± 0.84	0.084
Advanced maternal age (≥35 years), n (%)			< 0.001
No	2163 (88.1)	594 (73.6)	
Yes	293 (11.9)	213 (26.4)	
Infertility type, n (%)			0.104
Primary infertility	1332 (54.2)	464 (57.5)	
Secondary infertility	1124 (45.8)	343 (42.5)	
History of previous miscarriage, n (%)			0.305
No	1643 (66.9)	524 (64.9)	
Yes	813 (33.1)	283 (35.1)	
Infertility factors, n (%)			
Pelvic factor	1585 (64.5)	403 (49.9)	< 0.001
Tubal factor	1451 (59.1)	316 (39.1)	< 0.001
Endometriosis	143 (5.8)	46 (5.7)	0.891
Ovulatory disorder / ovarian factor	566 (23.0)	135 (16.7)	< 0.001
Uterine factor	110 (4.5)	45 (5.6)	0.206
Other factors	190 (7.7)	280 (34.7)	< 0.001
No female factor	365 (14.9)	40 (5.0)	< 0.001
Male factor infertility	675 (27.5)	336 (41.6)	< 0.001
Embryonic developmental stage, n (%)			< 0.001
D2/D3 (cleavage stage)	317 (12.9)	36 (4.4)	
D5/D6 (blastocyst stage)	2139 (87.1)	771 (95.6)	
No. of embryos transferred, n (%)			0.019
1	42 (1.7)	25 (3.0)	
≥2	2414 (98.3)	782 (97.0)	
Embryo quality, n (%)			< 0.001
Good quality	2154 (87.7)	648 (80.3)	
Non-good quality	302 (12.3)	159 (19.7)	
Fertilization method, n (%)			< 0.001
IVF	1793 (73.0)	514 (63.7)	
ICSI	663 (27.0)	293 (36.3)	
Cycle type, n (%)			< 0.001
Fresh cycle	550 (22.4)	67 (8.3)	
Frozen-thawed cycle	1906 (77.6)	740 (91.7)	

Note: Continuous variables are presented as mean ± SD; categorical variables as n (%). Pelvic factors include pelvic inflammatory disease and adhesions; tubal factors include tubal obstruction, hydrosalpinx, and prior tubal surgery; ovulatory disorder/ovarian factors include polycystic ovary syndrome, ovulatory dysfunction, and diminished ovarian reserve; uterine factors include fibroids, adenomyosis, intrauterine adhesions, and prior uterine surgery; other factors include thyroid disorders and systemic autoimmune conditions; male factor infertility includes oligozoospermia, asthenozoospermia, teratozoospermia, and azoospermia.

Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics 25, and the prediction model was built and validated in RStudio (v4.5.1). Continuous variables with approximately normal distributions are presented as mean ± standard deviation (SD) and compared between groups using the independent-samples t-test (or Welch's t-test when variances were unequal). Categorical variables are presented as counts and percentages and compared using the Pearson chi-square test; Fisher's exact test or the continuity-corrected chi-square test was used when any expected cell count was less than 5. All tests were two-sided, and P<0.05 was considered statistically significant. Risk factors for SR were examined by univariable and multivariable logistic regression, with effect sizes expressed as odds ratios (OR) with 95% confidence intervals (95% CI).

An SR risk prediction model was developed in the DZT cohort and internally validated using a 70:30 training-to-test split. Discrimination was assessed by the receiver operating characteristic (ROC) curve and the area under the curve (AUC); the optimal classification threshold was determined by the Youden index, from which sensitivity, specificity, and related metrics were derived. Calibration was evaluated using a calibration curve and the Brier score. Clinical utility was assessed by decision curve analysis (DCA). Risk stratification was performed by dividing the test-set predicted probabilities into tertiles, and SR incidence was compared across the resulting low-, intermediate-, and high-risk groups.

Results

Study population

A total of 3,768 patients with ART-conceived twin pregnancies were included: 3,263 in the DZT cohort (SR group, n=807; Non-SR group, n=2,456) and 505 in the MZT cohort (SR group, n=62; Non-SR group, n=443). The DZT cohort underwent full comparative analysis, logistic regression, and prediction model development; findings from the MZT cohort are presented descriptively.

Baseline and treatment characteristics in the DZT cohort

Baseline and treatment characteristics are shown in Table 1. Several differences were observed between the SR and Non-SR groups.

Regarding continuous variables, maternal age was significantly higher in the SR group than in the Non-SR group (31.86±4.25 vs. 29.99±3.67 years, P <0.001), and BMI was also slightly higher (22.61±3.26 vs. 22.35±3.14 kg/m², P =0.052). The number of previous miscarriages did not differ significantly between groups (0.51±0.84 vs. 0.45±0.80, P =0.084).

Among categorical variables, the proportion of women of advanced maternal age (≥35 years) was significantly higher in the SR group (26.4% vs. 11.9%, P <0.001). Infertility type (pri-

mary vs. secondary) was similar between groups (P =0.104). A history of previous miscarriage did not differ significantly between groups (SR 35.1% vs. Non-SR 33.1%, P = 0.305). The distribution of infertility causes also differed significantly, with notable differences in the proportions of pelvic, tubal, and ovarian factors (all P <0.001 or P =0.001), suggesting that the underlying cause of infertility may be related to SR risk. With respect to embryo and cycle characteristics, blastocyst-stage transfer (D5/D6) was more frequent in the SR group (95.6% vs. 87.1%, P <0.001). The distribution of the number of embryos transferred also differed between groups, with ≥2 embryos being the predominant strategy in both (P =0.001). The SR group had a higher rate of ICSI (36.3% vs. 27.0%, P <0.001) and a markedly higher proportion of frozen-thawed cycles (91.7% vs. 77.6%, P <0.001).

Risk factors for SR in the DZT cohort: logistic regression analysis

Univariable and multivariable logistic regression analyses were performed with SR as the outcome variable, incorporating the nine prespecified core predictors. The multivariable model was fitted using complete-case analysis.

Univariable analysis

On univariable analysis (Table 2), advanced maternal age (≥35 years) was significantly associated with SR (OR=2.65, 95% CI: 2.17–3.23, P<0.001). Each 1 kg/m² increase in BMI was associated with a modest increase in SR risk (OR=1.03, 95% CI: 1.00–1.05, P=0.048). History of previous miscarriage and infertility type did not reach statistical significance or showed unstable effect estimates in univariable analysis. Transfer of ≥2 embryos was associated with a lower likelihood of SR (OR=0.55, 95% CI: 0.33–0.91, P=0.021). Non-good-quality embryos were associated with higher SR risk (OR=1.67, 95% CI: 1.35–2.06, P<0.001), as were ICSI compared with conventional IVF (OR=1.54, 95% CI: 1.30–1.82, P<0.001) and frozen-thawed cycles compared with fresh cycles (OR=2.95, 95% CI: 2.03–4.29, P<0.001).

Multivariable analysis

After adjustment for all nine predictors (Table 2), advanced maternal age remained an independent risk factor for SR (aOR=3.00, 95% CI: 2.40–3.76, P<0.001). BMI was no longer significant in the multivariable model (aOR=1.01, 95% CI: 0.99–1.04, P=0.344). A history of previous miscarriage was independently associated with increased SR risk (aOR=1.32, 95% CI: 1.06–1.64, P=0.013). Secondary infertility was inversely associated with SR (aOR=0.67, 95% CI: 0.54–0.83, P<0.001). Embryonic developmental stage was independently associated with SR, with blastocyst-stage (D5/D6) transfer showing a substantially higher risk than cleavage-stage transfer (aOR = 3.05, 95% CI: 1.92–4.82, P < 0.001). Transfer of ≥2 embryos remained associated with lower SR risk (aOR=0.53, 95% CI: 0.31–0.90, P=0.019). Non-good-quality embryos (aOR=1.79, 95% CI: 1.42–2.26, P<0.001), ICSI (aOR=1.56, 95% CI: 1.30–1.88, P<0.001), and frozen-thawed cycles (aOR=3.33, 95% CI: 2.23–4.99, P<0.001) were each independently associated with increased SR risk; frozen-thawed cycle type showed the largest effect size among all predictors in the model.

Development and validation of the DZT-SR prediction model

A multivariable logistic prediction model was constructed using the nine core predictors described above. The regression equation is as follows:

logit(p) = −4.072 + 1.304 × (advanced maternal age ≥35 years) + 0.056 × BMI − 0.077 × (secondary infertility) + 0.439 × (history of previous miscarriage) + 1.243 × (blastocyst transfer D5/D6) − 1.119 × (≥2 embryos transferred) − 0.621 × (non-good-quality embryos) + 0.317 × (ICSI) + 3.016 × (frozen-thawed cycle)

p = 1 / [1 + exp(−logit(p))] represents the predicted probability of SR.

Discrimination

ROC curve analysis showed good discriminative performance for the model (Figure 1): AUC was 0.868 in the training set and 0.857 in the test set. Using the Youden index, the optimal clas-

Table 2. Factors associated with SR: univariate and multivariable logistic regression

Variable	OR (95% CI)	*P* (univariable)	aOR (95% CI)	*P* value (multivariable)
Advanced maternal age (≥35 vs <35 years)	2.65 (2.17-3.23)	< 0.001	3.00 (2.40-3.76)	< 0.001
Maternal BMI (per 1 kg/m² increase)	1.03 (1.00-1.05)	0.048	1.01 (0.99-1.04)	0.344
Infertility type (secondary infertility)	0.87 (0.74-1.03)	0.104	0.67 (0.54-0.83)	< 0.001
Embryonic stage (D5/D6 blastocyst)	3.23 (2.29-4.57)	< 0.001	3.05 (1.92-4.82)	< 0.001
History of previous miscarriage (yes)	1.09 (0.92-1.29)	0.305	1.32 (1.06-1.64)	0.013
No. of embryos transferred (≥2)	0.55 (0.33-0.91)	0.021	0.53 (0.31-0.90)	0.019
Embryo quality (non-good quality)	1.67 (1.35-2.06)	< 0.001	1.79 (1.42-2.26)	< 0.001
Fertilization method (ICSI)	1.54 (1.30-1.82)	< 0.001	1.56 (1.30-1.88)	< 0.001
Cycle type (frozen-thawed cycle)	2.95 (2.03-4.29)	< 0.001	3.33 (2.23-4.99)	< 0.001

sification threshold was 0.501, at which the model achieved a sensitivity of 0.822 and a specificity of 0.830 in the test set. The positive predictive value (PPV) was 0.750, the negative predictive value (NPV) was 0.883, and overall accuracy was 0.827.

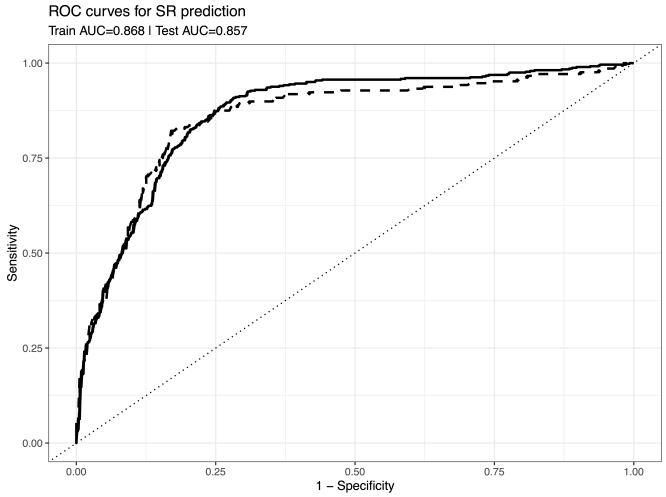


Figure 1. ROC curves of the DZT-SR prediction model in the training and test sets.

ROC curves for the spontaneous reduction (SR) prediction model in dizygotic twin (DZT) pregnancies, shown separately for the training set (solid line) and the test set (dashed line). The area under the curve (AUC) was 0.868 in the training set and 0.857 in the test set. The optimal classification threshold determined by the Youden index was 0.501, at which the model achieved a sensitivity of 0.822 and a specificity of 0.830 in the test set. The diagonal reference line represents a non-discriminating classifier.

Calibration

The calibration curve for the test set showed good overall agreement between predicted probabilities and observed SR rates across risk strata, with the curve tracking close to the 45° reference line and no obvious systematic over- or underestimation across most of the risk range (Figure 2). The Brier score was 0.142, indicating that the overall probability prediction error was within an acceptable range.

Clinical net benefit

Decision curve analysis showed that, across a reasonable range of threshold probabilities, the net benefit of using the model exceeded both the "treat all" and "treat none" reference strategies (Figure 3). This suggests that applying the model to identify high-risk patients and guide follow-up decisions could help avoid unnecessary interventions while maintaining the ability to detect truly high-risk individuals

Risk stratification

To make the model more applicable in practice, predicted probabilities in the test set were divided into tertiles to form low-, intermediate-, and high-risk groups. Observed SR rates increased markedly across strata: 8.3% (15/181) in the low-risk group, 28.7% (52/181) in the intermediate-risk group, and 77.9% (141/181) in the high-risk group (Figure 4).

Perinatal outcomes in the DZT cohort

Perinatal outcome analyses were restricted to live births with complete delivery records. Compared with the Non-SR group, the SR group had a significantly later gestational age at delivery (38.42±1.99 vs. 36.10±2.04 weeks, P<0.001) and a marked-

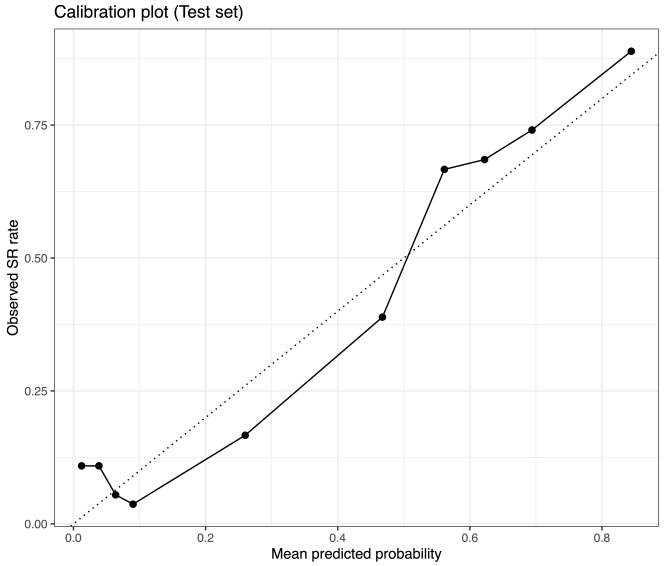


Figure 2. Calibration curve of the DZT-SR prediction model in the test set.

The calibration curve plots predicted probability of SR against the observed SR rate across risk strata in the test set. The diagonal dashed line represents perfect calibration (predicted probability = observed rate). Points falling close to this reference line indicate good agreement between model predictions and actual outcomes. The Brier score was 0.142, indicating acceptable overall prediction error.

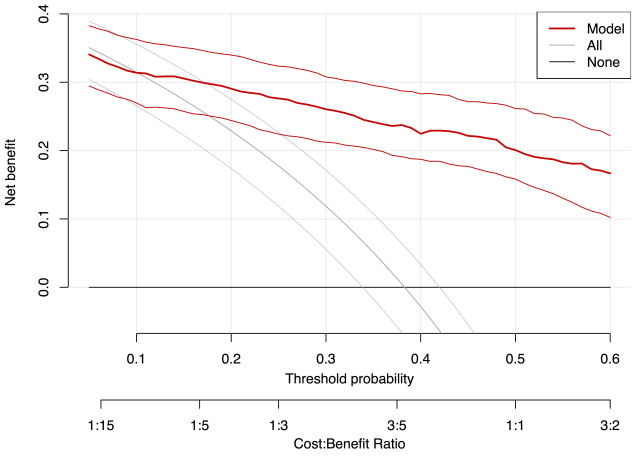


Figure 3. Decision curve analysis (DCA) for the DZT-SR prediction model.

Decision curve analysis evaluating the clinical net benefit of the SR prediction model across a range of threshold probabilities in the test set. The model curve (solid line) is compared with two reference strategies: treating all patients as high risk ("Treat all", gray line) and treating no patients as high risk ("Treat none", horizontal line at zero). The model provides greater net benefit than either reference strategy across a clinically meaningful range of threshold probabilities, supporting its potential utility for guiding individualized follow-up decisions in early pregnancy.

ly lower rate of preterm birth (13.6% vs. 61.1%, P<0.001). The smaller of the two birth weights was significantly higher in the SR group (P<0.001), and the proportion of pregnancies with any low birth weight infant (any LBW, <2,500 g) was substantially lower (8.7% vs. 53.3%, P<0.001). Birth weight discordance of ≥20% (discord20) was observed in 15.6% of Non-SR deliveries; as most SR pregnancies resulted in singleton deliveries, this measure was effectively zero in the SR group (P<0.001). Overall rates of any pregnancy complication were similar between groups (21.2% vs. 20.8%, P=0.816). The rate of hypertensive disorders of pregnancy was slightly lower in the SR group but did not reach significance (6.7% vs. 8.8%, P=0.060). Glucose metabolism abnormalities were more frequent in the SR group (8.0% vs. 3.7%, P<0.001), whereas premature rupture of membranes was less common (2.2% vs. 5.1%, P<0.001). Detailed perinatal outcomes are presented in Table 3.

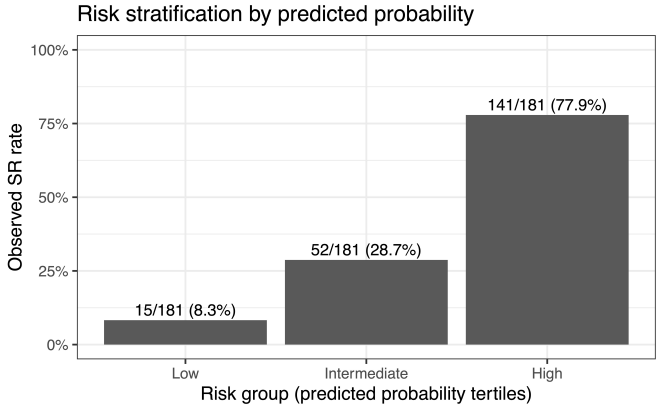


Figure 4. Risk stratification by predicted probability tertiles in the test set.

Bar chart showing the observed SR incidence across three risk strata defined by tertiles of the model-predicted probability in the test set (n=543). The low-risk group (first tertile) had an SR rate of 8.3% (15/181), the intermediate-risk group (second tertile) 28.7% (52/181), and the high-risk group (third tertile) 77.9% (141/181). The stepwise increase in SR incidence across strata confirms the model's risk stratification ability.

Baseline and treatment characteristics in the MZT cohort

The MZT cohort was smaller, comprising 443 women in the Non-SR group and 62 in the SR group. Among continuous variables, maternal age was slightly higher in the SR group but the difference did not reach significance (31.73±4.45 vs. 30.76±4.28 years, P=0.098); BMI (P=0.887) and number of previous miscarriages (P=0.131) were also similar between groups. Among categorical variables, the proportion of women with advanced maternal age was higher in the SR group (25.8% vs. 17.4%), though again without statistical significance (P=0.109). Infertility type, history of previous miscarriage, and most infertility-related factors including pelvic and tubal causes showed no significant differences. One exception was uterine factor infertility, which was more common in the SR group (14.5% vs. 6.3%, P=0.033). Embryonic developmental stage, number of embryos transferred, embryo quality, fertilization method, and cycle type were all comparable between groups (all P>0.05). Full details are provided in Table 4.

Table 3. Comparison of perinatal outcomes in the dizygotic twin group.

Variable	Non-SR	SR	*P* value
Gestational age at delivery (weeks)	36.10 ± 2.04	38.42 ± 1.99	< 0.001
Preterm birth, n (%)	1500 (61.1)	109 (13.6)	< 0.001
Cesarean delivery, n (%)	2326 (94.7)	572 (70.9)	< 0.001
Smaller birth weight (g)	2398.70 ± 462.32	3284.94 ± 596.27	< 0.001
Any LBW (<2500 g), n (%)	1309 (53.3)	70 (8.7)	< 0.001
Birth weight discordance ≥20%, n (%)	383 (15.6)	—	—
Any complication, n (%)	511 (20.8)	171 (21.2)	0.816
Hypertensive disorders of pregnancy, n (%)	216 (8.8)	54 (6.7)	0.060
Glucose metabolism abnormality, n (%)	91 (3.7)	64 (8.0)	< 0.001
Premature rupture of membranes, n (%)	125 (5.1)	17 (2.2)	< 0.001

Note: Continuous variables are presented as mean ± SD; categorical variables as n (%). Analyses were restricted to live births with complete delivery records. LBW, low birth weight; any LBW, at least one infant with birth weight <2500 g; birth weight discordance ≥20%, defined as (larger birth weight − smaller birth weight) / larger birth weight ≥ 0.20.

Perinatal outcomes in the MZT cohort

In the MZT cohort, the SR group delivered at a significantly later gestational age (37.49±3.22 vs. 35.63±1.78 weeks, P<0.001) and had a markedly lower rate of preterm birth (30.6% vs. 83.5%, P<0.001). The cesarean delivery rate was also significantly lower in the SR group (71.0% vs. 96.8%, P<0.001). The smaller birth weight was higher in the SR group (3,063.33±716.92 vs. 2,251.06±508.81 g, P<0.001), and the rate of any LBW was substantially lower (17.7% vs. 65.2%, P<0.001). The overall rate of any pregnancy complication did not differ between groups (24.2% vs. 20.5%, P=0.508), nor did rates of hypertensive disorders of pregnancy (P=0.339) or premature rupture of membranes (P=0.501). Glucose metabolism abnormalities could not be formally tested due to the small number of events. Perinatal outcomes for the MZT cohort are summarized in Table 5.

Discussion

This study analyzed risk factors for SR and compared perinatal outcomes in a large single-center cohort of ART-conceived twin pregnancies, and developed and internally validated an SR prediction model in the DZT cohort. The findings suggest that SR is not a random occurrence but is closely tied to maternal age, embryo characteristics, and treatment cycle factors.

Interpretation of risk factors for SR

Advanced maternal age (≥35 years) was the strongest independent risk factor for SR in this study (aOR=3.00), consistent with previous reports describing older maternal age in VTS populations [9, 11]. The biological rationale is reasonably straightforward: as women age, the rate of oocyte aneuploidy

rises and the proportion of chromosomally abnormal embryos increases, raising the likelihood of early embryonic arrest. In a twin pregnancy, this may manifest as one embryo failing to develop further. Older age may also be accompanied by reduced endometrial receptivity, which could further compromise early placental establishment.

Frozen-thawed cycle type had the largest effect size in the model (aOR=3.33). These cycles are often used when fresh transfers have previously failed or when endometrial preparation is needed, and the patients selected for them may already carry a higher background risk of early pregnancy complications. Beyond patient selection, individual variation in embryo-endometrial synchrony during frozen-thawed transfer may affect the stability of early placentation [24]. Blastocyst-stage transfer (aOR=3.05) was also associated with higher SR risk. Two explanations seem plausible: first, blastocysts implant more efficiently and are more likely to be identified as twin pregnancies on the initial ultrasound, making subsequent SR easier to detect on follow-up; second, patients selected for blastocyst transfer may represent more complex treatment histories that independently carry higher SR risk [18, 22].

Non-good-quality embryos (aOR=1.79) and ICSI (aOR=1.56) were both independently associated with SR, pointing toward embryo developmental potential as a common underlying theme. Embryos of lower morphological quality may have limited capacity to sustain development after implantation. For ICSI, the association may partly reflect the characteristics of the patient population — particularly those with severe male factor infertility — rather than the procedure itself [18, 22]. A history of previous miscarriage (aOR=1.32) may reflect underlying endometrial dysfunction that compromises early implantation stability.

The inverse association between transferring ≥2 embryos and SR risk (aOR=0.53) requires some clarification. This study enrolled only patients with confirmed twin pregnancies, so the comparison concerns who among established twin pregnancies is more likely to undergo SR, not who is more likely to form a multiple pregnancy in the first place. When two embryos are transferred, clinicians may select a better-matched embryo combination, or follow-up management for these patients may differ in ways that contribute to greater twin pregnancy stability. BMI showed a weak association with SR on univariable analysis but lost significance after adjustment, suggesting it may act through shared pathways with age or cycle strategy rather than as an independent contributor. Secondary infertility was associated with lower SR risk (aOR=0.67), possibly because these patients have had at least one prior live birth and

Table 4. Baseline and treatment characteristics of the monozygotic twin group.

Variable	Non-SR (n = 443)	SR (n = 62)	*P* value
Maternal age (years)	30.76 ± 4.28	31.73 ± 4.45	0.098
BMI(kg/m²)	22.49 ± 3.25	22.43 ± 2.76	0.887
No. of previous miscarriages	0.44 ± 0.72	0.69 ± 1.29	0.131
Advanced maternal age (≥35 years), n (%)			0.109
No	366 (82.6)	46 (74.2)	
Yes	77 (17.4)	16 (25.8)	
Infertility type, n (%)			0.301
Primary infertility	224 (50.6)	27 (43.5)	
Secondary infertility	219 (49.4)	35 (56.5)	
History of previous miscarriage, n (%)			0.518
No	297 (67.0)	39 (62.9)	
Yes	146 (33.0)	23 (37.1)	
Infertility factors, n (%)			
Pelvic factor	239 (54.0)	28 (45.2)	0.222
Tubal factor	153 (34.5)	20 (32.3)	0.723
Endometriosis	41 (9.3)	7 (11.3)	0.609
Ovulatory disorder / ovarian factor	81 (18.3)	9 (14.5)	0.468
Uterine factor	28 (6.3)	9 (14.5)	0.033
Other factors	55 (12.4)	9 (14.5)	0.641
No female factor	46 (10.4)	10 (16.1)	0.177
Male factor infertility	100 (22.6)	18 (29.0)	0.260
Embryonic developmental stage, n (%)			0.332
D2/D3 (cleavage stage)	18 (4.0)	4 (6.5)	
D5/D6 (blastocyst stage)	425 (96.0)	58 (93.5)	
No. of embryos transferred, n (%)			0.434
1	266 (60.1)	34 (54.8)	
≥2	177 (39.9)	28 (45.2)	
Embryo quality, n (%)			0.652
Good quality	401 (90.5)	55 (88.7)	
Non-good quality	42 (9.5)	7 (11.3)	
Fertilization method, n (%)			0.128
IVF	307 (69.3)	37 (59.7)	
ICSI	136 (30.7)	25 (40.3)	
Cycle type, n (%)			0.392
Fresh cycle	80 (18.1)	14 (22.6)	
Frozen-thawed cycle	363 (81.9)	48 (77.4)	

Note: Continuous variables are presented as mean ± SD; categorical variables as n (%). Infertility factor categories are defined as in Table 1.

Table 5. Comparison of perinatal outcomes in the monozygotic twin group.

Variable	Non-SR	SR	*P* value
Gestational age at delivery (weeks)	35.63 ± 1.78	37.49 ± 3.22	< 0.001
Preterm birth, n (%)	370 (83.5)	19 (30.6)	< 0.001
Cesarean delivery, n (%)	429 (96.8)	44 (71.0)	< 0.001
Smaller birth weight (g)	2251.06 ± 508.81	3063.33 ± 716.92	< 0.001
Any LBW (<2500 g), n (%)	289 (65.2)	11 (17.7)	< 0.001
Birth weight discordance ≥20%, n (%)	53 (12.0)	——	——
Any complication, n (%)	91 (20.5)	15 (24.2)	0.508
Hypertensive disorders of pregnancy, n (%)	40 (9.0)	3 (4.8)	0.339
Glucose metabolism abnormality, n (%)	27 (6.1)	3 (4.8)	——
Premature rupture of membranes, n (%)	21 (4.7)	1 (1.6)	0.501

may have a more favorable baseline endometrial environment, though the underlying mechanism warrants further study.

Perinatal outcomes after SR

The SR group delivered at a later gestational age and had substantially lower rates of preterm birth and low birth weight compared with the ongoing twin group. This is broadly consistent with published systematic reviews showing that perinatal outcomes in SR-survivor pregnancies tend to be better than those in continuing twin pregnancies, while remaining somewhat worse than in pregnancies that were singletons from the start [9-10]. Early exposure to a twin implantation environment and the more complex placental bed remodeling that accompanies it may have lasting effects on placental function in the surviving fetus. The higher proportions of advanced maternal age and frozen-thawed cycles in the SR group may also independently contribute to perinatal risk. The higher rate of glucose metabolism abnormalities in the SR group (8.0% vs. 3.7%) alongside the lower rate of premature rupture of membranes (2.2% vs. 5.1%) suggests that the complication profile of SR-survivor pregnancies may have some distinctive features that deserve further investigation.

Prediction model performance and clinical implications

The DZT-SR prediction model was built entirely from variables available through routine clinical assessment. Internal validation yielded an AUC of 0.857, with sensitivity and specificity both exceeding 0.82 at the optimal threshold. Calibration and DCA net benefit were both acceptable. After stratifying by predicted probability tertiles, observed SR rates rose from 8.3% in the low-risk group to 77.9% in the high-risk group, providing a practical basis for tailoring follow-up intensity in early pregnancy. Women in the low-risk group can be managed with standard follow-up, while those in the high-risk group may benefit from shorter ultrasound intervals, closer monitoring for gestational sac development and intrauterine hematoma, and additional psychological support. The model should be used as a risk stratification aid rather than a definitive diagnostic tool, and clinical judgment remains essential.

Limitations and future directions

Several limitations should be noted. First, the retrospective single-center design introduces the potential for selection and information bias. Second, SR was not subclassified by whether only a gestational sac disappeared or whether cardiac activity had already been detected before arrest, which may affect incidence estimates. Third, data on chorionicity, ovarian stimulation protocol, luteal support regimen, first-trimester intrauterine hematoma, and embryo genetic testing results were unavailable, limiting mechanistic interpretation. Fourth, the MZT cohort was too small to support reliable multivariable modeling. Fifth, only internal validation was performed, and the model's generalizability needs to be tested in external cohorts. Future work could address these gaps through multicenter prospective studies with standardized SR definitions and follow-up protocols; by incorporating chorionicity, intrauterine hematoma, and embryo genetic information to explore biological mechanisms; by externally validating the model and translating it into more accessible formats such as a nomogram or online calculator; and by accumulating larger MZT samples to enable a more rigorous comparison of SR patterns and perinatal outcome profiles between MZT and DZT pregnancies.

Conclusion

This study found that SR in ART-conceived twin pregnancies is independently associated with advanced maternal age, history of previous miscarriage, blastocyst-stage transfer, non-good-quality embryos, ICSI, and frozen-thawed cycles, while secondary infertility and transfer of ≥2 embryos were associated with lower SR risk. Perinatal outcomes in SR-survivor pregnancies were better than those in ongoing twin pregnancies overall, though the higher rate of glucose metabolism abnormalities points to a somewhat distinct complication profile. The DZT-SR prediction model, built on nine routinely available clinical variables, showed good internal validation performance (AUC=0.857) and clear risk stratification ability, and may support individualized follow-up planning in early pregnancy. Multicenter prospective studies and external validation are needed to strengthen the evidence base and improve the model's broader applicability.

Abbreviations

SR: Spontaneous reduction; ART: Assisted reproductive technology; DZT: Dizygotic twins; MZT: Monozygotic twins; ROC: Receiver operating characteristic; AUC: Area under the curve; DCA: Decision curve analysis; ICSI: Intracytoplasmic sperm injection; aOR: Adjusted odds ratio; WHO: World Health Organization; eSET: Elective single embryo transfer; VTS: Vanishing twin syndrome; BMI: Body mass index; IVF: In vitro fertilization; ICM: Inner cell mass; TE: Trophectoderm; LBW: Low birth weight; SD: Standard deviation; OR: Odds ratio; CI: Confidence interval; PPV: Positive predictive value; NPV: Negative predictive value.

Author Contributions

Zhaolian Wei: Conceptualization, Project administration, Resources, Supervision, Writing – review & editing. Shuqi Cai: Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Writing – original draft.

Acknowledgements

The authors would like to thank the staff of the Reproductive Center of the First Affiliated Hospital of Anhui Medical University for their assistance with data collection and clinical follow-up.

Funding Information

Not Applicable.

Ethics Approval and Consent to Participate

This study was approved by the Committee on Medical Ethics of the First Affiliated Hospital of Anhui Medical University (Reference number: Quick-PJ 2018-07-05). The study was conducted in accordance with the Declaration of Helsinki and relevant ethical guidelines for biomedical research involving human subjects. As this was a retrospective study based on routinely collected clinical data, the requirement for individual informed consent was waived by the ethics committee.

Competing Interests

The authors declare that they have no existing or potential commercial or financial relationships that could create a conflict of interest at the time of conducting this study.

Data Availability

Not Applicable.

References

[1] Organization WH. (2023). Infertility prevalence estimates, 1990–2021. World Health Organization. http://digitallibrary.un.org/record/4012749

[2] Multiple gestation associated with infertility therapy: an American Society for Reproductive Medicine Practice Committee opinion. (2012). Fertility and Sterility, 97(4), 825–834. https://doi.org/https://doi.org/10.1016/j.fertnstert.2011.11.048

[3] Guidance on the limits to the number of embryos to transfer: a committee opinion. (2021). Fertility and Sterility, 116(3), 651–654. https://doi.org/https://doi.org/10.1016/j.fertnstert.2021.06.050

[4] Multifetal Gestations: Twin, Triplet, and Higher-Order Multifetal Pregnancies: ACOG Practice Bulletin, Number 231. (2021). Obstet Gynecol, 137(6), e145–e162. https://doi.org/10.1097/aog.0000000000004397

[5] Santana DS, Cecatti JG, Surita FG, Silveira C, Costa ML, Souza JP, et al. (2016). Twin Pregnancy and Severe Maternal Outcomes: The World Health Organization Multicountry Survey on Maternal and Newborn Health. Obstetrics & Gynecology, 127(4), 631–641. https://doi.org/10.1097/aog.0000000000001338

[6] Alteri A, Arroyo G, Baccino G, Craciunas L, De Geyter C, Ebner T, et al. (2024). ESHRE guideline: number of embryos to transfer during IVF/ICSI†. Hum Reprod, 39(4), 647–657. https://doi.org/10.1093/humrep/deae010

[7] Gerris JM. (2005). Single embryo transfer and IVF/ICSI outcome: a balanced appraisal. Hum Reprod Update, 11(2), 105–121. https://doi.org/10.1093/humupd/dmh049

[8] Sukur YE, Altun T, & Pal L. (2012). Predictors of spontaneous reduction in multiple pregnancies conceived following assisted reproductive technology. Eur J Obstet Gynecol Reprod Biol, 162(2), 174–177. https://doi.org/10.1016/j.ejogrb.2012.02.031

[9] Zhou F, Li Y, Wang H, Liu S, & Wang X. (2020). Perinatal outcomes in vanishing twin pregnancies following assisted reproductive technology (ART) - a systematic review and meta-analysis. J Perinat Med, 48(7), 639–647. https://doi.org/10.1515/jpm-2020-0088

[10] Sun L, Jiang LX, & Chen HZ. (2017). Obstetric outcome of vanishing twins syndrome: a systematic review and meta-analysis. Arch Gynecol Obstet, 295(3), 559–567. https://doi.org/10.1007/s00404-017-4289-9

[11] Carlson K, Mikes BA. Vanishing Twin Syndrome. [Updated 2025 Nov 8]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2026 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK563220/

[12] Ge Y, Lai S, Li X, Shi J, Ma C, & Zhao J. (2022). Vanishing twin syndrome is associated with first-trimester intrauterine hematoma in twin pregnancies after in vitro fertilization. Front Endocrinol (Lausanne), 13, 1062303. https://doi.org/10.3389/fendo.2022.1062303

[13] Vitthala S, Gelbaya TA, Brison DR, Fitzgerald CT, & Nardo LG. (2009). The risk of monozygotic twins after assisted reproductive technology: a systematic review and meta-analysis. Hum Reprod Update, 15(1), 45–55. https://doi.org/10.1093/humupd/dmn045

[14] Busnelli A, Dallagiovanna C, Reschini M, Paffoni A, Fedele L, & Somigliana E. (2019). Risk factors for monozygotic twinning after in vitro fertilization: a systematic review and meta-analysis. Fertil Steril, 111(2), 302–317. https://doi.org/10.1016/j.fertnstert.2018.10.025

[15] Hviid KVR, Malchau SS, Pinborg A, & Nielsen HS. (2018). Determinants of monozygotic twinning in ART: a systematic review and a meta-analysis. Hum Reprod Update, 24(4), 468–483. https://doi.org/10.1093/humupd/dmy006

[16] Kanter JR, Boulet SL, Kawwass JF, Jamieson DJ, & Kissin DM. (2015). Trends and correlates of monozygotic twinning after single embryo transfer. Obstet Gynecol, 125(1), 111–117. https://doi.org/10.1097/aog.0000000000000579

[17] Lundin K, & Ahlström A. (2015). Quality control and standardization of embryo morphology scoring and viability markers. Reprod Biomed Online, 31(4), 459–471. https://doi.org/10.1016/j.rbmo.2015.06.026

[18] The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. (2011). Hum Reprod, 26(6), 1270–1283. https://doi.org/10.1093/humrep/der037

[19] Coticchio G, Ahlström A, Arroyo G, Balaban B, Campbell A, De Los Santos MJ, et al. (2025). The Istanbul Consensus update: a revised ESHRE/ALPHA consensus on oocyte and embryo static and dynamic morphological assessment(† ‡). Reprod Biomed Online, 50(6), 104955. https://doi.org/10.1016/j.rbmo.2025.104955

[20] Schoolcraft WB, Gardner DK, Lane M, Schlenker T, Hamilton F, & Meldrum DR. (1999). Blastocyst culture and transfer: analysis of results and parameters affecting outcome in two in vitro fertilization programs. Fertil Steril, 72(4), 604–609. https://doi.org/10.1016/s0015-0282(99)00311-8

[21] De los Santos MJ, Apter S, Coticchio G, Debrock S, Lundin K, Plancha CE, et al. (2016). Revised guidelines for good practice in IVF laboratories (2015). Hum Reprod, 31(4),

685–686. https://doi.org/10.1093/humrep/dew016
[22] Gardner DK, Schoolcraft WB, Wagley L, Schlenker T, Stevens J, & Hesla J. (1998). A prospective randomized trial of blastocyst culture and transfer in in-vitro fertilization. Hum Reprod, 13(12), 3434–3440. https://doi.org/10.1093/humrep/13.12.3434
[23] Gardner DK, Lane M, Stevens J, Schlenker T, & Schoolcraft WB. (2000). Blastocyst score affects implantation and pregnancy outcome: towards a single blastocyst transfer. Fertil Steril, 73(6), 1155–1158. https://doi.org/10.1016/s0015-0282(00)00518-5
[24] Committee Opinion No. 719: Multifetal Pregnancy Reduction. (2017). Obstet Gynecol, 130(3), e158–e163. https://doi.org/10.1097/aog.0000000000002302