

Latent developmental-arrest burden after cleavage-stage transfer: a counterfactual reanalysis of a published IVF cohort

Enver Kerem Dirican

keremdirican@gmail.com

Akdeniz University

Şafak Olgan

Üsküdar University

Research Article

Keywords: Blastocyst, Cleavage-stage embryo, Embryo transfer, Extended embryo culture, Developmental arrest, Clinical pregnancy

Posted Date: October 13th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11174662/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Purpose

To estimate how many cleavage-stage (day-3) transfers in low-zygote IVF cycles occur where extended culture would have yielded no transferable blastocyst, and to compare clinical pregnancy after day-5 transfer with that in day-3 cycles inferred to be progressable.

Methods

Counterfactual secondary analysis of a published single-center cohort of 1116 first fresh cycles in which transfer day was assigned by the weekday of oocyte retrieval (584 day 3; 532 day 5, including 70 extended-culture cancellations). Zygote-specific day-5 cancellation rates were applied to the day-3 arm to estimate a latent-arrest fraction, and the day-5 per-protocol clinical pregnancy rate was compared with the implied day-3 progressable rate. The published arm-level comparison is not repeated as a finding.

Results

With ≤ 5 zygotes, an estimated 104 of 471 day-3 cycles (22.1%; 95% CI 17.3–26.8) were putatively futile transfers, falling from 37.3% at one zygote to 4.5% at five; 85% of this burden arose in cycles with one or two zygotes. The implied day-3 progressable rate was 32.7%, against 31.3% after day-5 per-protocol transfer (residual difference – 1.3 percentage points; 95% CI – 8.6 to 5.7). With ≥ 6 zygotes the residual difference was 23.6 points (95% CI 11.7–35.4). Covariate-adjusted g-computation gave similar estimates; the transfer-day-by-zygote-number interaction was significant (continuous model, $P = 0.004$).

Conclusions

About one in five cleavage-stage transfers in low-zygote cycles, concentrated in one- and two-zygote cycles, was estimated to occur where extended culture would have found no transferable blastocyst. These assumption-dependent estimates support counseling and prospective evaluation of cumulative live birth, not a universal transfer policy.

Introduction

Cleavage-stage transfer on day 3 (D3) has historically been favored when few embryos are available, because culture to the blastocyst stage on day 5 (D5) may leave nothing to transfer [1]. Current guidance addresses embryo assessment and the number of embryos transferred but does not define a universal embryo-number threshold for selecting the day of transfer [2]. The clinical intuition remains that a transfer of uncertain value is preferable to cancellation.

Blastocyst culture improves selection by allowing embryos to demonstrate developmental competence before transfer. Comparative studies consequently report higher pregnancy and live-birth rates per transfer, although the advantage per initiated cycle is less certain because extended-culture cancellations count as failures [3, 4]. This distinction is most important when only one or a few zygotes are available.

Developmental arrest after cleavage is associated with chromosomal abnormalities, failure of embryonic genome activation, and oocyte-derived deficits that are present before transfer but may not be morphologically apparent on D3 [5, 6]. If most embryos that would arrest in vitro also cannot establish clinical pregnancy after D3 transfer, early transfer does not rescue their potential: it converts an observable laboratory cancellation into an unrecognized transfer followed by luteal support, a wait and a negative pregnancy test. We call this counterfactual cycle-level event a putatively futile transfer; it does not imply that the fate of an individual embryo is known.

One cohort was assembled in a way that allows this accounting to be attempted [7]. Transfer day was assigned by the weekday of oocyte retrieval rather than by patient or embryo characteristics, and extended-culture cancellations were retained as failures rather than excluded. The published analysis of that cohort compared the arms directly and found no statistically significant clinical pregnancy difference when few zygotes were available, and a difference favoring D5 when several were. An arm-level contrast cannot say how many D3 cycles were placed where extended culture would have found nothing to transfer, because the arrested phenotype is never observed on the D3 side. We therefore estimated that quantity: we applied the directly observed D5 cancellation rate at each zygote count to the D3 arm, expressed the result as a putatively futile-transfer burden, compared pregnancy rates among the D3 cycles inferred to be progressable, and examined whether the association varied continuously with zygote number.

Materials and methods

Source cohort and participants. The source cohort and clinical protocols have been described previously [7]. Briefly, all fresh cycles performed at the Reproductive Endocrinology and Assisted Reproduction Unit of a university hospital in Türkiye (institution blinded for review), from January 2018 through December 2019 were screened. Fertility-preservation, preimplantation-genetic-testing, natural-cycle, and repeat cycles were excluded, leaving 1710 first treatments. Friday retrievals ($n = 217$), whose transfer day could vary with cycle characteristics, and cycles with total fertilization failure or arrest before D3 ($n = 39$) were excluded, leaving 1454. Excluding 338 freeze-all cycles, decided by the supervising physician at retrieval on clinical grounds, yielded the analytic cohort of 1116 first fresh cycles with at least one two-pronuclear (2PN) zygote: 584 D3 and 532 D5. Of the D5 cycles, 462 proceeded to transfer and 70 were canceled because no transferable blastocyst was available. Zygote count ranged from 1 to 15 and female age from 19 to 47 years. All cycles used autologous oocytes and ejaculated partner sperm. No a priori sample-size calculation was performed; all eligible cycles were analyzed. Every count in the archived analytic dataset was checked against the source report before analysis (Online Resource 1, Supplementary Methods S6 and Supplementary Table S6).

Relationship to the previous report. This is a secondary analysis of a cohort published in full [7]. No participants, cycles, or outcome assessments were added, and the outcome definition is unchanged. The source article asked whether cleavage transfer benefits patients by avoiding extended-culture cancellation and what the relative efficacy of the two strategies is across zygote numbers, and answered both as an arm-versus-arm comparison. Its published quantities appear here only where the decomposition takes them as input or is read against them.

Supplementary Methods S7 sets out, item by item, which quantities are re-reported from that article and which are new here; Table 1, Table 2 and the legend of Fig. 2 mark previously published values.

Table 1
Baseline characteristics by transfer-day arm and zygote stratum

Characteristic	Day 3	Day 5	SMD ^{a,b}
<i>Overall (day 3, n = 584; day 5, n = 532) – previously reported except where footnote a states otherwise</i>			
Female age, y	33.40 ± 5.11	32.40 ± 4.67	−0.20
Female age ≥ 35 y, n (%)	272 (46.6)	196 (36.8)	−0.20
Previous IVF failures, n	0.51 ± 0.78	0.40 ± 0.70	−0.15
Basal FSH, IU/L	10.05 ± 7.22	8.76 ± 4.28	−0.22
Antral follicle count, n	11.70 ± 8.12	13.82 ± 7.87	+ 0.27
Cumulus–oocyte complexes, n	9.07 ± 5.84	10.27 ± 5.55	+ 0.21
Metaphase II oocytes, n	6.30 ± 4.39	7.25 ± 3.62	+ 0.24
2PN zygotes, n	3.58 ± 2.70	4.39 ± 2.70	+ 0.30
<i>≤ 5 zygotes (day 3, n = 471; day 5, n = 383) – new to this report</i>			
Female age, y	33.92 ± 5.07	32.49 ± 4.75	−0.29
Female age ≥ 35 y, n (%)	242 (51.4)	145 (37.9)	−0.27
Previous IVF failures, n	0.52 ± 0.77	0.42 ± 0.72	−0.13
Basal FSH, IU/L	10.48 ± 7.14	9.15 ± 4.21	−0.23
Antral follicle count, n	10.35 ± 7.45	12.19 ± 6.99	+ 0.25
Cumulus–oocyte complexes, n	7.34 ± 4.65	8.75 ± 5.11	+ 0.29
Metaphase II oocytes, n	4.82 ± 3.10	5.86 ± 2.84	+ 0.35
2PN zygotes, n	2.46 ± 1.30	3.02 ± 1.32	+ 0.43
<i>≥ 6 zygotes (day 3, n = 113; day 5, n = 149) – new to this report</i>			
Female age, y	31.27 ± 4.75	32.17 ± 4.46	+ 0.20
Female age ≥ 35 y, n (%)	30 (26.5)	51 (34.2)	+ 0.17
Previous IVF failures, n	0.45 ± 0.79	0.33 ± 0.66	−0.17
Basal FSH, IU/L	8.27 ± 7.33	7.76 ± 4.30	−0.09
Antral follicle count, n	17.32 ± 8.39	18.01 ± 8.43	+ 0.08
Cumulus–oocyte complexes, n	16.28 ± 4.65	14.15 ± 4.69	−0.46
Metaphase II oocytes, n	12.45 ± 3.55	10.83 ± 2.86	−0.51
2PN zygotes, n	8.20 ± 1.99	7.91 ± 2.06	−0.14

Characteristic	Day 3	Day 5	SMD ^{a,b}
Values are mean $\pm$ standard deviation unless otherwise indicated. Allocation was by weekday of oocyte retrieval and was not randomization; the arms are not balanced.			
^a Every standardized mean difference in this table is new to this report. The day-3 and day-5 columns of the overall block correspond to Table 1 of the source article [7], which also reported fertilization, embryo-transfer, cryopreservation and embryo-utilization rates and the mean number of embryos transferred; none of those is used here. Two rows of the overall block were not in the source article and are new to this report: female age ≥ 35 y and 2PN zygote count. Three cells of the previous-IVF-failure row differ from the published ones – mean 0.51 versus 0.52 in the day-3 arm, and mean 0.40 versus 0.42 with standard deviation 0.70 versus 0.72 in the day-5 arm – and the difference is examined in Supplementary Methods S6 (Online Resource 1). The stratum-specific blocks are new to this report.			
^b Standardized mean difference (day 5 minus day 3), calculated as the difference in means divided by the square root of the average of the two variances. A positive value indicates a higher mean in the day-5 arm. An absolute value > 0.10 is conventionally read as meaningful imbalance; values > 0.10 occur for every variable in at least one stratum.			
FSH follicle-stimulating hormone, IVF in vitro fertilization, SMD standardized mean difference, 2PN two-pronuclear			

Table 2

Clinical pregnancy outcomes of the planned transfer strategy by zygote stratum, with the latent-arrest counterfactual decomposition

Stratum	Group	Source	n	Pregnancies, n	CPR, %	Adjusted OR (95% CI) ^a	P
≤ 5 zygotes	Day 3 (all)	Published	471	120	25.5	1.00 (reference)	—
	Day 5, intention-to-treat	Published	383	100	26.1	0.79 (0.57–1.10)	0.16
	Day 5, per protocol	New	319	100	31.3	1.00 (0.72–1.40)	0.99
	Implied day-3 progressable ^b	New	367.1	120	32.7	—	—
	Estimated futile day-3 transfers ^c	New	103.9 (22.1%; 17.3–26.8) ^e	0 (assumed)	0 (assumed)	—	—
Residual difference (day-5 per protocol – implied day-3 progressable) ^d : –1.3 pp (95% CI – 8.6 to + 5.7); covariate-adjusted, – 1.2 pp (95% CI – 8.7 to + 5.9)							
≥ 6 zygotes	Day 3 (all)	Published	113	29	25.7	1.00 (reference)	—
	Day 5, intention-to-treat	Published	149	72	48.3	3.22 (1.80–5.78)	< 0.001
	Day 5, per protocol	New	143	72	50.3	3.64 (2.00–6.61)	< 0.001
	Implied day-3 progressable ^b	New	108.4	29	26.8	—	—
	Estimated futile day-3 transfers ^c	New	4.6 (4.1%; 1.2–7.6) ^e	0 (assumed)	0 (assumed)	—	—
Residual difference (day-5 per protocol – implied day-3 progressable) ^d : +23.6 pp (95% CI + 11.7 to + 35.4); covariate-adjusted, + 24.0 pp (95% CI + 12.2 to + 35.4)							
Overall	Day 3 (all)	Published	584	149	25.5	1.00 (reference)	—
	Day 5, intention-to-treat	Published	532	172	32.3	1.21 (0.92–1.58)	0.17

Stratum	Group	Source	n	Pregnancies, n	CPR, %	Adjusted OR (95% CI) ^a	<i>P</i>
	Day 5, per protocol	New	462	172	37.2	1.49 (1.13– 1.96)	0.005
	Implied day-3 progressable ^b	New	475.5	149	31.3	—	—
	Estimated futile day-3 transfers ^c	New	108.5 (18.6%; 14.7– 22.5) ^e	0 (assumed)	0 (assumed)	—	—

Stratum	Group	Source	n	Pregnancies, n	CPR, %	Adjusted OR (95% CI) ^a	P
Residual difference (day-5 per protocol – implied day-3 progressable) ^d : +5.9 pp (95% CI – 0.6 to + 12.0); covariate-adjusted, + 6.1 pp (95% CI – 0.1 to + 12.2)							
“Source” marks whether the row reproduces a value from the source article [7] or is derived here for the first time. All adjusted odds ratios in the table are new, because the source article’s models did not contain a zygote-count covariate. Clinical pregnancy was defined as an intrauterine pregnancy with fetal cardiac activity. Day-5 intention-to-treat retains the 70 extended-culture cancellations as failures; day-5 per protocol includes only the 462 cycles that reached transfer.							
^a Adjusted odds ratio for day 5 versus day 3 (reference), from logistic regression adjusted for female age, previous IVF failures, basal FSH, antral follicle count, cumulus–oocyte complexes, metaphase II oocytes and 2PN zygote count entered linearly.							
^b Implied subgroup remaining after removal of the expected latent-arrest fraction. Denominators are expected counts derived from the directly observed zygote-specific day-5 cancellation rates, not observed integers, and are given to one decimal place.							
^c Day-3 cycles that would have been canceled in extended culture under the primary specification ($\varphi = 0$) but instead proceeded to cleavage-stage transfer. These are modeled expectations, not observed individual outcomes; the zero pregnancy count is the $\varphi = 0$ assumption rather than an observation.							
^d Bootstrap 95% confidence intervals from 5000 stratified replications. The covariate-adjusted value is from g-computation, in which day-5 cancellation was modeled on zygote number and the covariates listed in footnote a, and predicted cancellation risk was standardized to the day-3 covariate distribution. A post hoc two one-sided tests analysis of the ≤ 5 -zygote stratum was compatible with equivalence at a ± 10 -percentage-point margin (both one-sided $P < 0.05$) but not at a stricter ± 7 -point margin (one-sided $P = 0.064$ against the lower bound); neither margin was prespecified. The residual difference is a descriptive counterfactual decomposition rather than an identified causal treatment effect; see Discussion.							
^e Percentage in parentheses is the estimated latent-arrest fraction of the day-3 stratum, with a 95% confidence interval from 20,000 parametric bootstrap replicates over the binomial sampling of the zygote-specific day-5 cancellation counts (Supplementary Methods S5, Online Resource 1). The interval covers sampling variability in those observed rates only; it does not cover the assumption that arrest risk is exchangeable between arms conditional on zygote count.							
The day-3 and day-5 intention-to-treat counts and rates in all three strata were reported in Table 2 of the source article [7], which also gave them at each zygote count from one to nine or more. The adjusted odds ratios reported there, 0.874 for ≤ 5 and 3.122 for ≥ 6 zygotes, came from a model without the zygote-count covariate and therefore differ from those shown here. The per-protocol counts and rates, the odds ratios shown here, the implied progressable subgroup, the estimated futile transfers and both residual differences are new to this report.							
CI confidence interval, CPR clinical pregnancy rate, FSH follicle-stimulating hormone, IVF in vitro fertilization, OR odds ratio, pp percentage points, 2PN two-pronuclear							

Allocation and laboratory procedures. Retrieval on Monday or Tuesday led to D3 transfer, on Wednesday or Thursday to D5. All oocytes underwent intracytoplasmic sperm injection. Embryos were cultured in single-step medium under oil in the same benchtop incubator at 6% carbon dioxide and 5% oxygen and were assessed using the Istanbul Consensus criteria [8]. One or two embryos were transferred. Luteal support

was vaginal progesterone from oocyte retrieval until a negative pregnancy test or documentation of a viable fetus.

Outcome and definitions. The outcome was clinical pregnancy, defined as an intrauterine pregnancy with fetal cardiac activity at 6–7 gestational weeks. Cancellation was defined at the cycle level as extended culture yielding no transferable blastocyst; all 70 D5 cancellations met it, none for a clinical or logistical reason. The dataset did not distinguish failure to form a blastocyst from formation of a blastocyst judged unsuitable for transfer. There were no missing values for the outcome, allocation, zygote count, or modeled covariates.

Counterfactual decomposition. D5 cancellations were retained as failures in the intention-to-treat (ITT) analysis. The observed D5 cancellation fraction at each zygote count was applied to the corresponding D3 cycles to estimate the number that would have been canceled under extended culture. This transport assumes that arrest risk under extended culture is exchangeable between arms conditional on zygote count alone; that is an assumption, not a finding. Under the primary specification these latent-arrest cycles were assigned a clinical pregnancy probability of zero; the remainder constituted the implied D3-progressable group, whose clinical pregnancy rate (CPR) was the observed D3 pregnancies divided by the estimated progressable D3 cycles. The residual difference was the observed D5 per-protocol CPR minus the implied D3-progressable CPR. An interval for the latent-arrest count was obtained by treating each zygote-specific cancellation count as binomial and drawing 20,000 parametric bootstrap replicates (Supplementary Methods S5). The decomposition is descriptive, not an identified causal effect within a directly observed principal stratum.

Statistical analysis. Results were reported overall and for the descriptive strata of five or fewer (≤ 5) and six or more (≥ 6) zygotes. These strata were not prespecified for the present analysis; they were inherited unchanged from the original cohort report, where they were themselves reported as a post hoc subgroup analysis. Between-arm imbalance was described using standardized mean differences. Logistic regression compared D5 with D3 under ITT and per-protocol definitions, adjusted for female age, previous in vitro fertilization failures, basal follicle-stimulating hormone, antral follicle count, cumulus–oocyte complexes, metaphase II oocytes and 2PN zygote count -- the source article's adjustment set, extended by zygote count because it is the effect modifier of interest. The residual difference and its 95% confidence interval (CI) were estimated from 5000 bootstrap replications stratified by transfer day, in which the zygote-specific cancellation rates were re-estimated within each replicate, so that uncertainty in the imputed denominator is propagated rather than held fixed.

Supporting analyses. Zygote number was modeled continuously with a natural cubic spline with three degrees of freedom interacted with transfer day; after the centering constraint the basis contributes two columns, so the interaction test carries 2 df. Marginal predictions were standardized over observed covariates. A covariate-adjusted g-computation modeled D5 cancellation and standardized predicted cancellation risk to the D3 covariate distribution. Sensitivity analysis allowed latent-arrest D3 cycles a clinical pregnancy probability, ϕ , from 0% to 30%. Post hoc two one-sided tests examined equivalence at margins of 7 and 10 percentage points, which against a comparator of about 31% are relative margins of

22% and 32%; neither was prespecified and no clinical justification is offered for either. An exploratory age-by-zygote-group-by-transfer-day interaction used age 35 years as the cut point, and alternative zygote thresholds were assessed with Holm correction. Specifications, diagnostics and bootstrap schemes are in Supplementary Methods S1–S5, with results in Supplementary Tables S1, S2 and S5. Analyses used Python 3.11 and are reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement. Supplementary Methods S1–S7, Supplementary Tables S1–S6 and Supplementary Figs. S1–S3 are provided in Online Resource 1, and the analysis code with derived outputs in Online Resource 2.

Use of generative artificial intelligence. Generative AI tools (ChatGPT, OpenAI; Claude, Anthropic) were used for language editing and condensation, formatting to journal requirements, drafting of figure-rendering and interval-computation scripts, and cross-checking of reported values and references against the archived analysis outputs and the source publications. They were not used to generate data. The authors reviewed and verified all output and take full responsibility for the content.

Results

Participants and baseline characteristics. The D3 arm included older women and fewer oocytes and zygotes overall, but in the ≥ 6 -zygote stratum the direction reversed: those D3 women were younger, with more cumulus–oocyte complexes and more metaphase II oocytes, yet a lower pregnancy rate (Table 1). Mean 2PN zygote count was 3.58 in the D3 arm and 4.39 in the D5 arm. The overall observed D5 cancellation rate was 13.2% (70/532). It was 16.7% (64/383) among cycles with ≤ 5 zygotes and 4.0% (6/149) among those with ≥ 6 zygotes. Cancellation fell steeply across the first three zygote counts, from 37.3% (22/59) at one and 27.6% (24/87) at two to 10.1% (9/89) at three, staying at or below 10.1% thereafter with a shallow rise across six to eight; none of the 46 day-5 cycles with nine or more zygotes was canceled (Fig. 2A).

Observed pregnancy rates. The D5 ITT CPR was 32.3% overall, 26.1% with ≤ 5 zygotes, and 48.3% with ≥ 6 zygotes. Among the 462 D5 cycles that reached transfer, corresponding per-protocol rates were 37.2%, 31.3%, and 50.3%. D3 CPRs were 25.5%, 25.5%, and 25.7%, respectively (Table 2). In adjusted ITT models including zygote count the association was unclear in the lower-zygote stratum (adjusted odds ratio [aOR], 0.79; 95% CI, 0.57–1.10; $P = 0.16$) and favored D5 in the higher-zygote stratum (aOR, 3.22; 95% CI, 1.80–5.78; $P < 0.001$) (Table 2).

Latent arrest and residual difference. Applying zygote-specific D5 cancellation rates to D3 cycles yielded an estimated 103.9 latent-arrest cycles among 471 D3 cycles with ≤ 5 zygotes (22.1%; 95% CI, 17.3 to 26.8), 4.6 among 113 with ≥ 6 zygotes (4.1%; 95% CI, 1.2 to 7.6), and 108.5 among all 584 cycles (18.6%; 95% CI, 14.7 to 22.5) (Fig. 1). Covariate-adjusted g-computation gave closely similar burdens of 21.8% (95% CI, 16.9 to 27.0), 2.5% (95% CI, 0.8 to 6.2) and 18.1% (95% CI, 14.2 to 22.3). The stratum figure conceals a steep gradient: the transported arrest fraction was 37.3% at one zygote, 27.6% at two, 10.1% at three, 7.3% at four and 4.5% at five, so 88 of the 103.9 cycles (85%) came from cycles with one or two zygotes. Under the

primary specification these are the estimated putatively futile transfers, and the implied D3-progressable CPR was 32.7% in the lower-zygote stratum, 26.8% in the higher-zygote stratum, and 31.3% overall.

The residual difference was – 1.3 percentage points (95% CI, – 8.6 to 5.7) with ≤ 5 zygotes, 23.6 points (95% CI, 11.7 to 35.4) with ≥ 6 zygotes, and 5.9 points (95% CI, – 0.6 to 12.0) overall (Table 2). Covariate-adjusted g-computation gave – 1.2, 24.0 and 6.1 points. The continuous transfer-day-by-zygote-number interaction was significant under ITT (likelihood-ratio $\chi^2 = 10.93$; $df = 2$; $P = 0.004$). The marginal D5-minus-D3 difference increased monotonically from – 3.8 points at one zygote to 26.7 points at 10 zygotes and crossed zero at approximately four zygotes (Fig. 2B; Supplementary Table S5; Supplementary Fig. S2). Likelihood-ratio tests for the alternative continuous and dichotomous specifications are in Supplementary Table S2; the dichotomy summarizes a gradient rather than a validated threshold.

Sensitivity and exploratory analyses. Across ϕ from 0% to 30% the higher-zygote residual difference stayed positive with its CI excluding zero, and the lower-zygote estimate shifted toward D5, its lower bound staying above – 10 points (Supplementary Table S3; Supplementary Fig. S3). Post hoc equivalence was supported at a margin of 10 points but not at 7 points. The age-by-zygote-group-by-transfer-day interaction was not significant ($P = 0.38$; Supplementary Table S4; Supplementary Fig. S1). Interaction tests at alternative zygote thresholds, with Holm adjustment, are given in Supplementary Table S1.

Discussion

This analysis does not revisit the published arm-level comparison [7]; it quantifies the accounting beneath it. Applying the observed D5 cancellation rate at each zygote count to the D3 arm, an estimated 103.9 of 471 D3 cycles with ≤ 5 zygotes, approximately one in five, would not have produced a transferable blastocyst. The burden is not evenly spread, so “one in five” describes no individual patient group. Once those cycles are removed from the D3 denominator the residual difference is – 1.3 percentage points, with an interval spanning – 8.6 to 5.7, so it remains imprecise. The earlier absence of a significant difference here is therefore compatible with a comparator of different composition, rather than requiring any advantage conferred by cleavage transfer. In the ≥ 6 -zygote stratum, where latent arrest was estimated at only 4.1%, the decomposition changes little: the residual difference of 23.6 points is close to the difference between the observed D3 rate of 25.7% and the D5 per-protocol rate of 50.3%. Across the whole range the continuous model described a gradient rather than a threshold.

The contribution is quantitative rather than directional. The published ITT design, counting cancellations as failures, is the right one for the clinical comparison but is silent on the composition of the D3 arm, because the cycles that would have arrested are never identified there. The decomposition estimates that composition from the only place it is observable, the D5 arm, and carries the estimate across. Exchangeability given zygote count is not exactly met: within the ≤ 5 stratum the D3 women were older, with higher basal follicle-stimulating hormone and fewer metaphase II oocytes (Table 1), each predicting more arrest, so the transported rate is more likely biased low than high and 22.1% should be read as a floor. Covariate-adjusted g-computation corrects the arrest model for measured imbalance and gives a similar 21.8%, but cannot test exchangeability, because the D3 arrest outcome is never observed. What is added is

a number for the futile-transfer burden, the progressable-subgroup comparison that follows from it, and an account of how both behave as the zero-pregnancy assumption is relaxed -- assumption-dependent throughout, and not a randomized causal effect.

The biological premise is plausible but not deterministic for an individual embryo, and is the one the source article invoked [7]: embryos arresting before blastulation are enriched for aneuploidy and developmental abnormalities, and relevant determinants may be established in the oocyte or before embryonic genome activation [5, 6, 9, 10]. The primary model therefore assigned them zero pregnancy probability; the φ analysis relaxed that, without reversing the higher-zygote finding.

These results complement literature showing higher pregnancy rates per transfer after blastocyst culture while cumulative outcomes depend on attrition, cryopreservation and the chosen denominator, and reporting a stepwise transition to blastocyst culture as feasible where laboratory performance is adequate [3, 4, 11–13]. The opposite position also has support. A retrospective cohort of women with a single viable embryo on D3 reported higher clinical pregnancy and live birth after immediate cleavage-stage transfer than after growing on [14]. That cohort is not interchangeable with this one: it began after D3 selection rather than at the zygote stage, transfer or grow-on was chosen by the clinician with the patient rather than by schedule, and its accrual spanned 2008 to 2016 with mixed fresh and frozen pathways, which secular change in practice makes hard to disentangle; published commentary further notes that patients undergoing embryo biopsy were not excluded and that gravidity and parity favored the day-3 group [15]. A recent commentary asked whether embryos are better nurtured in utero or in vitro and proposed that, in poor-prognosis patients, cleavage-stage transfer may rescue embryos that would otherwise arrest [16]; published correspondence replied that any such calculation must separate culture attrition from implantation failure and keep cancellations in the initiated-cycle denominator [17].

Rescue is what φ represents: the probability that a cycle destined to arrest nonetheless establishes clinical pregnancy after cleavage-stage transfer. Allowing it does not help the cleavage-stage case within this accounting. As φ rises, a larger share of the observed D3 pregnancies is attributed to cycles that would have arrested, the implied progressable rate falls, and the residual difference moves further in favor of D5, from – 1.3 points at $\varphi = 0$ to 7.2 points at $\varphi = 30\%$ (Supplementary Table S3; Supplementary Fig. S3). This is an accounting identity rather than a test: φ is not identified by these data, and the sweep bounds the consequences of assuming a value rather than estimating one. The informative anchor is $\varphi = 25.5\%$, the observed D3 rate, at which latent-arrest cycles are assumed to behave like every other D3 cycle; there the decomposition returns the raw arm comparison and adds nothing.

A target-trial emulation of 11,163 first cycles with a single zygote estimated higher per-couple live birth under planned cleavage-stage transfer (12.5% vs 10.1%; risk ratio, 1.24) [18]. Intended transfer stage was unrecorded for the 2442 cycles reaching no transfer, so the contrast was recovered with a multinomial mixture model using calendar time; it is a model-based estimate, not an observed comparison. Its g-computation assigns embryos failing in culture the live-birth probability of transferred cleavage-stage embryos with similar measured characteristics, and its sensitivity analyses varied the blastocyst development and live-birth rates rather than that assignment. φ is the parameter left unvaried there, and a

schedule-allocated cohort of the present kind allows it to be bounded empirically. The estimands also differ: that analysis addressed per-couple live birth in single-zygote cycles, whereas this one decomposes clinical pregnancy across the zygote range. At one zygote the marginally standardized day-5-minus-day-3 difference here was - 3.8 percentage points (95% CI, - 11.5 to 4.2), too imprecise to support equivalence and, the endpoint being clinical pregnancy, incapable of confirming or refuting a live-birth effect. The two are complementary rather than mutually confirming.

The patient experience also differs. A D5 cancellation communicates that no embryo met the laboratory's transfer criteria and avoids a transfer procedure, continued luteal support and the wait before a negative test; D3 transfer preserves a chance that cannot be directly observed, but in an estimated subset defers recognition of the same limitation. Counseling should present both without describing cancellation as inherently worse than any transfer.

Application is conditional on laboratory quality. The observed cancellation rates and the estimated latent burden are center-specific, and this cohort's blastocyst-development performance was not recomputed as a formal Vienna Consensus key performance indicator [19], so these data do not establish that the center met that benchmark. A laboratory considering broader extended culture should monitor blastocyst development against a specified denominator and review its own cancellation, cryosurvival and cumulative live-birth outcomes.

Strengths and limitations. The weekday schedule reduced direct prognosis-based selection of transfer day, and the complete dataset allowed zygote-specific and covariate-adjusted analyses, stable across alternative standardizations. All cycles came from one laboratory in a 24-month window under a single culture system postdating the Vienna Consensus indicators [19], so the estimates characterize one laboratory over two years rather than a period of changing practice. It is a secondary analysis of a cohort already reported, so it is not independent evidence: it re-describes the same 1116 cycles and outcome ascertainment, its directional conclusions restate rather than confirm the source report, and any bias there is inherited. Schedule allocation was not randomization, so weekday effects and unmeasured confounding remain possible, and the clinical freeze-all decision after retrieval excluded 338 cycles, leaving a selected subset of initiated cycles in both arms. The D3 progressable group was inferred and the D5 per-protocol group defined by a postallocation event. The cancellation model had 70 events and moderate internal discrimination, and the equivalence, age and threshold analyses were exploratory. The dataset could not separate failure to form a blastocyst from formation of a blastocyst graded unsuitable, so the latent-arrest stratum mixes biology with laboratory policy and the zero-pregnancy assumption is conservative only for the first. At nine or more zygotes no cancellation was observed, so those D3 cycles were assigned an arrest probability of exactly zero rather than a shrunken estimate. The endpoint was clinical pregnancy rather than live birth, and the source report records a higher cryopreservation rate (26.4% vs 14.1%), a higher embryo utilization rate (46.0% vs 29.1%) and more embryos transferred per transfer (1.31 vs 1.10) in the D3 arm [7], none of it adjusted for here and all of it preventing inference about cumulative outcomes.

Prospective, multicenter studies should record intended transfer day before culture, follow all embryos from a retrieval to cumulative live birth, report laboratory performance, and measure the patient-reported effects

of cancellation and of negative transfer.

Conclusions

In cycles with five or fewer zygotes, an estimated 22.1% (95% CI, 17.3 to 26.8) of D3 transfers were placed in cycles that would not have produced a transferable blastocyst, almost all of them with one or two zygotes. Putting a number on that latent burden is what this reanalysis adds to the previously reported comparison of the two strategies in this cohort [7]. Once it is counted on both sides the residual difference spans a deficit of 8.6 points to an advantage of 5.7, so these data neither exclude nor establish a clinical pregnancy cost to extended culture in this stratum; and because that contrast removes cancellations from both arms, it does not by itself license a policy statement about initiated cycles. The estimates are assumption-dependent and are not a new arm-level effectiveness comparison. They support individualized counseling and prospective evaluation of cumulative live birth, not a universal transfer policy.

Declarations

Author Contribution

Author contributions: E.K.D. and S.O. wrote the main manuscript text and prepared the figures and tables. Both authors reviewed the manuscript.

Data Availability

The deidentified analytic dataset and the primary clinical database are available from the corresponding author on reasonable request, subject to institutional data-sharing policies and the terms on which the data were collected, which do not permit unrestricted public release. A synthetic dataset that reproduces the structure and value ranges of the analytic dataset, together with all derived numerical outputs and the analysis code, is provided in Online Resource 2 and permits end-to-end execution of the analysis pipeline.

References

1. Practice Committee of the American Society for Reproductive Medicine; Practice Committee of the Society for Assisted Reproductive Technology. Blastocyst culture and transfer in clinically assisted reproduction: a committee opinion. *Fertil Steril*. 2018;110:1246–52. <https://doi.org/10.1016/j.fertnstert.2018.09.011>.
2. National Institute for Health and Care Excellence. Fertility problems: assessment and treatment. NICE guideline NG257. London: NICE; 2026. <https://www.nice.org.uk/guidance/ng257>.
3. Martins WP, Nastri CO, Rienzi L, van der Poel SZ, Gracia C, Racowsky C. Blastocyst vs cleavage-stage embryo transfer: systematic review and meta-analysis of reproductive outcomes. *Ultrasound Obstet Gynecol*. 2017;49:583–91. <https://doi.org/10.1002/uog.17327>.

4. Glujovsky D, Cornelisse S, Blake D, Quinteiro Retamar AM, Alvarez Sedo CR, Ciapponi A. Blastocyst-stage versus cleavage-stage embryo transfer in assisted reproductive technology. *Cochrane Database Syst Rev*. 2026;1:CD002118. <https://doi.org/10.1002/14651858.CD002118.pub7>.
5. Braude P, Bolton V, Moore S. Human gene expression first occurs between the four- and eight-cell stages of preimplantation development. *Nature*. 1988;332:459–61. <https://doi.org/10.1038/332459a0>.
6. Magli MC, Jones GM, Gras L, Gianaroli L, Korman I, Trounson AO. Chromosome mosaicism in day 3 aneuploid embryos that develop to morphologically normal blastocysts in vitro. *Hum Reprod*. 2000;15:1781–6. <https://doi.org/10.1093/humrep/15.8.1781>.
7. Dirican EK, Olgan S, Sakinci M, Caglar M. Blastocyst versus cleavage transfers: who benefits? *Arch Gynecol Obstet*. 2022;305:749–56. <https://doi.org/10.1007/s00404-021-06224-2>.
8. Alpha Scientists in Reproductive Medicine; ESHRE Special Interest Group of Embryology. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. *Reprod Biomed Online*. 2011;22:632–46. <https://doi.org/10.1016/j.rbmo.2011.02.001>
9. Irani M, Zaninovic N, Canon C, O'Neill C, Gunnala V, Zhan Q, et al. A rationale for biopsying embryos reaching the morula stage on day 6 in women undergoing preimplantation genetic testing for aneuploidy. *Hum Reprod*. 2018;33:935–41. <https://doi.org/10.1093/humrep/dey053>.
10. Keefe D, Kumar M, Kalmbach K. Oocyte competency is the key to embryo potential. *Fertil Steril*. 2015;103:317–22. <https://doi.org/10.1016/j.fertnstert.2014.12.115>.
11. Yin Y, Chen G, Li K, Liao Q, Zhang S, Ma N, et al. Propensity score-matched study and meta-analysis of cumulative outcomes of day 2/3 versus day 5/6 embryo transfers. *Front Med*. 2017;11:563–9. <https://doi.org/10.1007/s11684-017-0535-6>.
12. Stimpfel M, Jancar N, Ban-Frangez H, Vrtacnik-Bokal E. When to transfer embryos if there is only 1 or 2? *Life (Basel)*. 2023;13:417. <https://doi.org/10.3390/life13020417>.
13. De Croo I, De Sutter P, Tilleman K. A stepwise approach to move from a cleavage-stage to a blastocyst-stage transfer policy for all patients in the IVF clinic. *Hum Reprod Open*. 2020;2020:hoaa034. <https://doi.org/10.1093/hropen/hoaa034>.
14. Xiao JS, Healey M, Talmor A, Vollenhoven B. When only one embryo is available, is it better to transfer on Day 3 or to grow on? *Reprod Biomed Online*. 2019;39:916–23. <https://doi.org/10.1016/j.rbmo.2019.08.003>.
15. Dirican EK. Overlooked details on cleavage versus blastocyst transfers when only one embryo is available. *J Gynecol Res Rev Rep*. 2024;6:1–2. [https://doi.org/10.47363/JGRRR/2024\(6\)196](https://doi.org/10.47363/JGRRR/2024(6)196).
16. Bartolacci A, de Girolamo S, Fraire-Zamora JJ, Pagliardini L, Papaleo E. The eternal dilemma: are embryos better nurtured in utero or in vitro? *Reprod Biomed Online*. 2025;51:105103. <https://doi.org/10.1016/j.rbmo.2025.105103>.
17. Dirican EK. Is the eternal dilemma dependent on embryo number? *Reprod Biomed Online*. 2026;52:105634. <https://doi.org/10.1016/j.rbmo.2026.105634>.
18. Fitzgerald O, Li W, Vallence C, Chambers GM, Rombauts L. Cleavage stage versus blastocyst stage transfers in patients with a single zygote: an emulated target trial. *Hum Reprod*. 2026;41:1106–14. <https://doi.org/10.1093/humrep/deag075>.

19. ESHRE Special Interest Group of Embryology; Alpha Scientists in Reproductive Medicine. The Vienna consensus: report of an expert meeting on the development of assisted reproductive technology laboratory performance indicators. Reprod Biomed Online. 2017;35:494–510.
<https://doi.org/10.1016/j.rbmo.2017.06.015>.

Figures

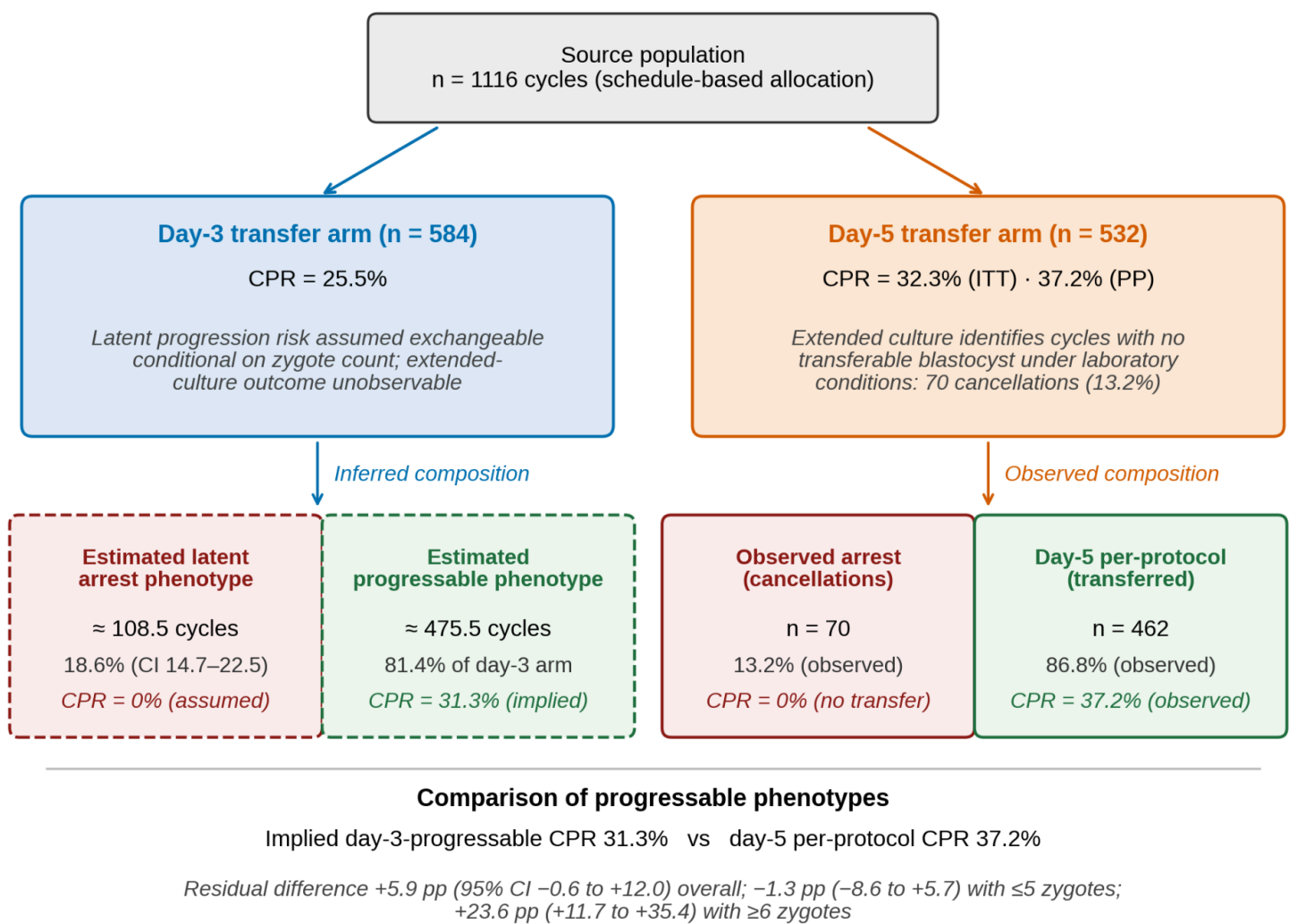


Figure 1

Counterfactual decomposition of the day-3 and day-5 transfer arms. The day-5 arm contains 70 observed extended-culture cancellations (13.2%) and 462 per-protocol transfers. Applying zygote-specific day-5 cancellation rates to the day-3 arm yields approximately 108.5 latent-arrest cycles (18.6%; 95% CI, 14.7 to 22.5) and 475.5 progressable cycles. The implied day-3 progressable-cycle clinical pregnancy rate is compared with the observed day-5 per-protocol rate. Dashed boxes indicate inferred quantities; solid boxes indicate observed quantities. CI = confidence interval; CPR = clinical pregnancy rate; ITT = intention-to-treat; PP = per protocol; pp = percentage points.

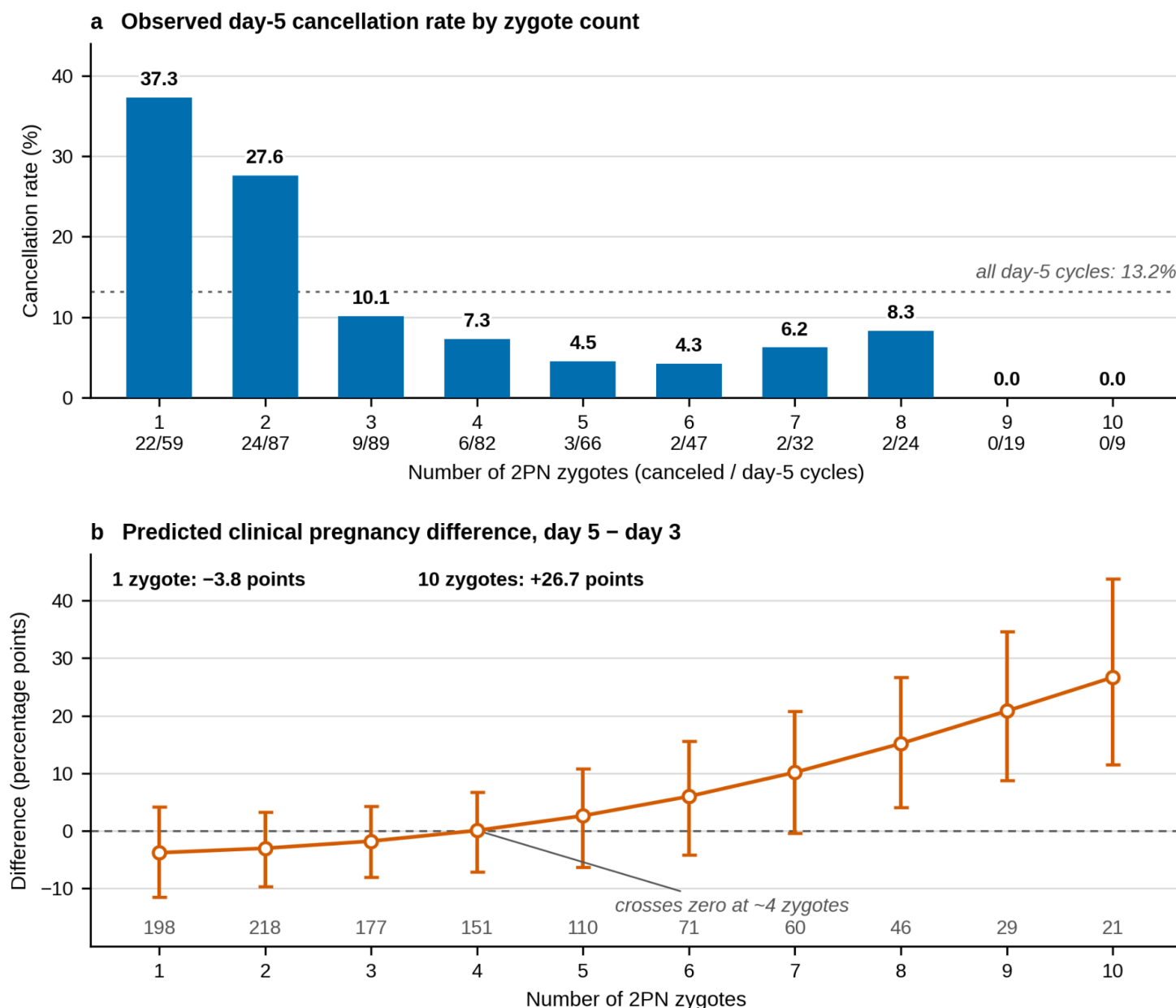


Figure 2

Cancellation and the transfer-day difference across zygote number. **(a)** Directly observed day-5 cancellation rate (extended culture yielding no transferable blastocyst) by two-pronuclear zygote count. These are the rates applied to the day-3 arm in the decomposition; the panel is shown as the input to the analysis rather than as an outcome comparison. The rate is printed above each bar and the number canceled over the number of day-5 cycles beneath each zygote count; the dotted line marks the overall day-5 cancellation rate of 13.2%. The 18 day-5 cycles with more than 10 zygotes, none of which was canceled, are not shown. The day-5 denominators at one to eight zygotes were reported in the source article [7]; the cancellation count at each zygote count, and therefore every rate plotted here, is new. **(b)** Predicted day-5-minus-day-3 clinical pregnancy difference by zygote count, from an intention-to-treat logistic model with a natural cubic spline (3 df) in zygote number interacted with transfer day and adjusted for female age, previous in vitro fertilization failures, basal follicle-stimulating hormone, antral follicle count, cumulus–oocyte complexes, and

metaphase II oocytes. Points are marginally standardized over the observed covariate distribution and bars are bootstrap 95% confidence intervals from 2000 replications. The number of cycles contributing at each zygote count is printed along the axis. The difference rose monotonically and crossed zero at approximately four zygotes, indicating a gradient rather than a discrete threshold (likelihood-ratio $\chi^2 = 10.93$; $df = 2$; $P = 0.004$). CI = confidence interval.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [OnlineResource1blinded.pdf](#)