Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Sperm source and fertilisation after intracytoplasmic sperm injection in male-factor and combined-factor infertility: confounding by indication and unit of analysis in a retrospective single-centre cohort

Hesham Al-Inany
kaainih@yahoo.com
Cairo University
Amal Shohayeb
Cairo University
Mohamed Waly
Cairo University
Ahmed Adel
Ganna Research Center
Mohamed Said
Ganna Research Center

Research Article

Keywords: intracytoplasmic sperm injection, sperm retrieval, azoospermia, male infertility, fertilisation rate, confounding factors, epidemiologic

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11652069/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

Objective. To examine whether an apparent difference in fertilisation between ejaculated and surgically retrieved spermatozoa at intracytoplasmic sperm injection (ICSI) in routinely collected single-centre data reflects sperm source or is explained by diagnosis, the unit of analysis and small oocyte cohorts.

Materials and Methods. Retrospective single-centre cohort (January 2025–June 2026; STROBE/RECORD) of ICSI cycles coded as male-factor or combined-factor infertility, grouped by recorded sperm source (ejaculated vs surgically retrieved). The principal estimate of fertilisation (fertilised/injected oocytes) came from binomial logistic regression with patient-clustered robust variance, adjusted for maternal age and infertility category. The prespecified cycle-level mean difference (patient-level bootstrap), six sensitivity analyses and total fertilisation failure are also reported; pregnancy fields were assessed for completeness only.

Results. Of 118 cycles (110 patients), 113 were evaluable: 91 ejaculated and 22 surgically retrieved (19 testicular, 3 presumed epididymal); azoospermia was documented in 14/22 surgical and 2/96 ejaculated cycles. Pooled fertilisation was 70.1% (295/421) vs 60.0% (57/95); adjusted odds ratio 0.74 (95% CI 0.37–1.48), −6.5 percentage points (−21.8 to 8.8). The prespecified cycle-level difference was −17.7 percentage points (−35.6 to 0.0). Total fertilisation failure occurred in 5.5% vs 22.7% of cycles (P = 0.023), all with one to three injected oocytes; with four or more oocytes, fertilisation was 67.9% vs 64.1%. Pregnancy records were incomplete and not compared.

Conclusion. The lower fertilisation with surgically retrieved spermatozoa was imprecise, attenuated by adjustment for infertility category and confined to cycles with very few oocytes. Because sperm source was inseparable from azoospermia, these data cannot guide the choice of sperm source; they illustrate how confounding by indication and the cycle-level unit of analysis can inflate apparent differences in routinely collected ICSI data.

Introduction

ICSI rendered biological fatherhood possible in cases of very poor semen parameters and, when performed in conjunction with surgical sperm retrieval, also in many cases of azoospermia [1, 2]. Current guidelines suggest the use of surgical sperm retrieval for obstructive azoospermia and non-obstructive azoospermia, following proper assessment, with the mode of procedure—percutaneous or microsurgical epididymal aspiration, testicular sperm aspiration, conventional testicular sperm extraction (TESE) or microdissection TESE—to be determined by aetiology and phenotype of the testes [3–6]. Microdissection is superior in the latter case, but fertilisation and live birth success after ICSI with testicular spermatozoa may greatly depend on the cause of male infertility [7–11]. In obstructive azoospermia, spermatozoa from the epididymis and from the testis seem to provide equal live birth rates [12].

Comparisons of fertilisation after ICSI with ejaculated, epididymal and testicular spermatozoa date from the first years of the technique and include considerably larger series than the present one. In a retrospective series of 1,034 cycles, normal fertilisation was lower with epididymal and testicular than with ejaculated spermatozoa, although pregnancy rates were comparable [40]. In an Egyptian series of 366 cycles, fertilisation and pregnancy rates did not differ between abnormal ejaculates, epididymal spermatozoa and testicular spermatozoa from men with obstructive azoospermia, whereas results were poorest with testicular spermatozoa from men with non-obstructive azoospermia [41]. Larger series and meta-analyses have since shown that fertilisation is lower with testicular spermatozoa in non-obstructive than in obstructive azoospermia [42, 43], that epididymal and testicular spermatozoa give similar results in obstructive azoospermia [12, 44], and that live-birth rates are lower in non-obstructive azoospermia than in obstructive azoospermia or with ejaculated spermatozoa [45]. These studies indicate that the aetiology of azoospermia, rather than the anatomical source of the spermatozoa, is the main determinant of outcome.

The comparison between the two types of spermatozoa is, therefore, clinically relevant only if both procedures are realistic for the same type of patients. This applies primarily to cases of cryptozoospermia, near-azoospermia with occasional spermatozoa in the ejaculate and high sperm DNA fragmentation [13, 14]. In cases of cryptozoospermia, two meta-analyses of predominantly retrospective studies have yielded contradictory results: the first found no difference in fertilisation and pregnancy rates, while a more recent update suggested better results with the use of testicular spermatozoa [15–17]. Single-centre studies have also provided inconsistent findings [18, 19] and repeat ejaculate collection has been recommended as an alternative to surgery [20]. In cases of high sperm DNA fragmentation, testicular spermatozoa are less likely to be DNA damaged and are associated with higher pregnancy rates, although the evidence base is low [21–24].

On the other hand, comparing unselected patient groups with surgical sperm retrieval in cases of azoospermia and ejaculated spermatozoa in oligozoospermia entails a conflation of sperm source with diagnosis, woman's features and laboratory protocols. Nevertheless, such comparisons are easily generated from the historical databases held by most fertility centres, and they are often presented without attention to the unit of analysis, to cycles with very few oocytes or to diagnostic differences between the groups. The present study therefore does not aim to re-establish whether surgically retrieved spermatozoa can fertilise oocytes, a question already answered by the larger series cited above. Its added value is to quantify, transparently and reproducibly, how much of an apparent sperm-source difference in a typical single-centre database is explained by diagnosis, by the choice of the cycle or the oocyte as the unit of analysis, and by complete fertilisation failure in cycles with very few oocytes. We analyzed a historical single-center ICSI database for the following two objectives: first, to compare fertilisation by sperm source in male-factor and combined-factor cycles with the focus on the unit of analysis and possible confounding; and second, to evaluate the completeness of the embryo and pregnancy fields to allow further comparisons.

Material and Methods

Study design and reporting

This was a retrospective single-center cohort study using routinely collected clinical laboratory records. It is reported in accordance with the STROBE statement and its extension for routinely collected health data (RECORD) [25, 26]. Terminology follows the International Glossary on Infertility and Fertility Care [27]. The analysis was exploratory and was not registered.

Data source

Data were drawn from Gann IVF Centers' databases (Haram & Mohandseen) spanning the period from January 2025 to June 2026. During the analysis, original identifiers were replaced with consecutive pseudonymised identifiers, while dates and free text were removed.

Eligibility and exposure classification

A cycle eligible for the analysis had complete linkage key, infertility-category code of 1 (male-factor) or 3 (combined male- and female-factor), and unambiguous sperm-source code. Ejaculated spermatozoa were encoded EJ. Codes TESTIS and TESTIS/MD indicated testicular sperm source, whereas codes EPID.ASP or EPIS.ASP were considered epididymal aspirates (EPIS.ASP was recorded verbatim in the worksheet and was interpreted as a spelling variant of EPID.ASP, epididymal sperm aspiration); hence, the surgically retrieved group consisted of records labelled testis and presumed epididymal aspirate. Records with ambiguous or mixed sources were excluded. These codes identify the anatomical source but not the procedure used (e.g., TESE, microdissection TESE, testicular sperm aspiration, percutaneous or microsurgical epididymal sperm aspiration) nor the type of spermatozoa used (fresh or cryopreserved). In addition, free text was searched for literal terms AZOO and SEVER to indicate documentation on presence of azoospermia and severe oligozoospermia; lack of a match was not interpreted as their absence. The infertility category was the code entered in the clinical worksheet at the time of treatment. The clinical findings underlying the code (for example, the semen analysis on which a male factor was diagnosed) were not part of the extract, so the category could not be re-verified against semen parameters; it was used as recorded, and the AZOO text search served only as supporting documentation of azoospermia.

Outcomes

Fertilisation was defined as the number of fertilised oocytes divided by the number of injected oocytes; cycles without injected oocytes were excluded, and the recorded counts and percentages were checked for internal consistency (analysis code, Supplementary Material). The original analysis plan specified the mean cycle-level fertilisation rate as the primary comparison. Because this measure gives a cycle with one injected oocyte the same weight as a cycle with many, the oocyte-level estimate from the clustered regression described below is presented in this revision as the principal estimate, and the prespecified

cycle-level comparison is reported alongside it for transparency. It could not be confirmed whether the recorded number of fertilised oocytes counted only normal (two-pronuclear, 2PN) fertilisation or also abnormal (one-pronuclear or three-or-more-pronuclear) fertilisation, nor that all injected oocytes were at metaphase II. The outcome is therefore overall fertilisation per injected oocyte rather than the normal fertilisation rate per injected metaphase II oocyte recommended as a key performance indicator [28]; if abnormal fertilisation was counted, rates would be overestimated, and any between-group difference in abnormal fertilisation could bias the comparison in either direction. Secondary outcomes were total fertilisation failure (all injected oocytes failed to fertilise) and the completeness of the embryo-transfer and pregnancy fields.

Pregnancy status was based on the worksheet code (PREG or NEG); cycles with codes CLO (closed without transfer) and empty were considered with no evaluable pregnancy status. Ultrasound confirmation was assessed separately, from YES or BY U/S entries in the SEEN field. Embryo-category fields T1-T4 were compared with the total of embryos transferred. Day of assessment, grading definitions corresponding to T1-T4 fields and the number of embryos assessed per cycle were unavailable; delivery and miscarriage codes were not recorded completely and therefore it was not possible to estimate clinical pregnancy, miscarriage and live-birth rates. These fields were therefore summarised descriptively to document their completeness and were not used for comparison between groups.

Covariates

Maternal age (years), infertility category (male-factor vs combined factor) and the number of injected oocytes were available for all evaluable cycles. Male age, semen parameters, sperm DNA fragmentation, hormonal profile, testicular histology, ovarian stimulation regimen and laboratory conditions were not available. In particular, the stimulation protocol, gonadotrophin dose, trigger type and markers of ovarian reserve could not be accounted for; maternal age and the number of injected oocytes are only partial proxies for oocyte quality, so residual confounding by ovarian stimulation and oocyte quality cannot be excluded.

Statistical analysis

Continuous variables are expressed as median (interquartile range, IQR) and compared with the Mann–Whitney U test; categorical variables are compared with Fisher's exact test. Proportions are provided with exact (Clopper–Pearson) 95% confidence intervals (CI) [46]; Fig. 2B shows Wilson score intervals [29].

Some patients contributed more than one cycle and each cycle produces several oocytes, so the patient was treated as the independent unit in all analyses [30]. The prespecified comparison – the difference in mean fertilization per cycle (surgically retrieved minus ejaculated), as defined in the original analysis plan – was calculated with a 95% percentile interval based on 10,000 non-parametric bootstraps of

patient IDs (random seed 20261106) so that all cycles of each resampled patient were considered together [31, 32]. Similarly, the difference in pooled fertilization at oocyte level was estimated. As the difference in mean gives the same weight to each cycle regardless of number of oocytes, the oocyte-level fertilization was modelled with the binomial (events/trials) logistic regression with cluster-robust variance estimator, equivalently generalised estimating equations with independence working correlation [33]; this oocyte-level model, adjusted for maternal age and infertility category, provides the principal estimate. Models were fitted with no adjustment, with linear adjustment for maternal age, with adjustment for maternal age and infertility category, and further adjusted for the natural logarithm of oocytes injected. The odds ratios (OR) were transformed to standardised differences in fertilisation percentage points (model-based estimates with delta-method 95% CI).

Prespecified sensitivity analyses of the cycle-level comparison (i) limited the surgical group to the exact TESTIS classification; (ii) limited it to testicular classifications (TESTIS and TESTIS/MD); (iii) limited the whole cohort to cycles with male-factor coding; (iv) used only first cycle per patient; (v) limited the cohort to maternal age < 40 years; and (vi) limited it to cycles with four or more oocytes injected. Total fertilisation failure was also tabulated by the number of injected oocytes (one to three vs four or more). Embryo-transfer and pregnancy fields were summarised descriptively, without between-group testing. No imputation was done and no correction for multiple comparisons was made; all p-values and intervals are considered as descriptive measures of compatibility with the data rather than confirmatory tests. All analyses were conducted in Python 3 (NumPy, SciPy, pandas); the analysis code is provided as Supplementary Material.

Ethics

The study protocol was approved by Ganna Research Ethical Committee and informed consent was waived for this retrospective analysis of anonymized records.

Results

Study cohort

The study cohort included 118 source-classified cycles of 110 patients (Fig. 1): 96 ejaculated cycles of 88 patients and 22 surgically labeled cycles of 22 patients (TESTIS 17, TESTIS/MD 2, EPID.ASP 2, EPIS.ASP 1), that is, 19 testicular and 3 presumed epididymal cycles. There were no patients contributing cycles to both groups; there were seven patients with 2–3 cycles within the ejaculated group. Five ejaculated cycles had no oocytes injected and were omitted from fertilization analysis, leaving 113 evaluable cycles (91 ejaculated, 22 surgically retrieved) with 516 injected oocytes. All 113 evaluable cycles passed the internal consistency checks.

Baseline characteristics

Characteristics of the two study groups are summarized in Table 1. Maternal age was 36 years (IQR 32–41) in the ejaculated and 39 years (35–41) in the surgical groups (P = 0.43), and 31.3% and 45.5% of the cycles included females who were 40 years of age or older. The median number of oocytes injected was small in both groups – 3 (2–7) and 4 (2–5) (P = 0.83), and 53.1% of all evaluable cycles included three or fewer oocytes injected. Combined-factor infertility was more common in the ejaculated (40.6% vs 18.2%) group (P = 0.053). Diagnostic text mentioning azoospermia was found in 14 of 22 surgical and 2 of 96 ejaculated cycles; diagnostic text mentioning severe oligoasthenoteratozoospermia was present in one ejaculated cycle. Thus, the two groups differed mainly in diagnosis: surgical retrieval was used predominantly in men with documented azoospermia, whereas the ejaculated group included more couples with combined-factor infertility.

Table 1
Characteristics at baseline according to the sperm source (total number of 118 source-classified cycles).

Characteristic	Ejaculated	Surgically retrieved	P value
Cycles, n	96	22	—
Patients (pseudonymised identifiers), n	88	22	—
Maternal age, years, median (IQR)	36 (32–41)	39 (35–41)	0.43
Maternal age ≥ 40 years, n (%)	30 (31.3)	10 (45.5)	—
Infertility category, n (%)			
Male factor (code 1)	57 (59.4)	18 (81.8)	0.053
Combined factor (code 3)	39 (40.6)	4 (18.2)	
Oocytes injected, median (IQR)[a]	3 (2–7)	4 (2–5)	0.83
Cycles with ≤ 3 oocytes injected, n/N (%)[a]	50/91 (54.9)	10/22 (45.5)	—
Diagnostic text, n (%)[b]			
Azoospermia-related text (AZOO)	2 (2.1)	14 (63.6)	—
Text indicating severe male factor (SEVER)	1 (1.0)	0	—
Sperm-source label, n			
EJ	96	—	
TESTIS / TESTIS/MD	—	17 / 2	
EPID.ASP / EPIS.ASP (presumed epididymal)	—	2 / 1	

IQR, interquartile range. P-values obtained from the Mann–Whitney U test (continuous data) or Fisher's exact test (categorical data). No correction for multiple testing applied. [a]Of evaluable cycles with at least one injected oocyte (ejaculated n = 91; surgically retrieved n = 22). [b]Exact text match in the diagnosis field. Absence of a match is not indicative of an absence of the diagnosis.

Fertilisation

Fertilisation results are summarised in Tables 2 and 3 and Fig. 2. In simple terms, 60.0% of oocytes injected with surgically retrieved spermatozoa were fertilised, compared with 70.1% of oocytes injected with ejaculated spermatozoa (57/95 vs 295/421; exact 95% CI 49.4–69.9 and 65.4–74.4), a crude difference of − 10.1 percentage points (patient-level bootstrap interval − 27.0 to 5.9). In the principal oocyte-level regression (Table 3), the unadjusted odds ratio (OR) for fertilisation with surgically retrieved vs ejaculated spermatozoa was 0.64 (95% CI 0.32–1.28). Adjustment for maternal age did not change the estimate (OR 0.64; 0.32–1.25), whereas additional adjustment for infertility category moved it

towards no difference (OR 0.74; 0.37–1.48; standardised difference − 6.5 percentage points, − 21.8 to 8.8). Combined-factor infertility was associated with higher fertilisation (OR 2.71; 1.54–4.77; pooled fertilisation 82.5% vs 63.1% in male-factor cycles). Further adjustment for the number of oocytes injected did not change the estimate (OR 0.73; 0.36–1.48).

Table 2
Fertilization rates according to the sperm source (113 evaluable cycles).

Outcome	**Ejaculated (n = 91)**	**Surgically retrieved (n = 22)**	**Difference or P value (95% CI)**
Oocytes injected, n	421	95	—
Oocytes fertilised, n	295	57	—
Pooled fertilisation, % (95% CI)[]	70.1 (65.4–74.4)	60.0 (49.4–69.9)	−10.1 pp (− 27.0 to 5.9)[]
Mean cycle fertilisation, % (prespecified)	74.0	56.3	−17.7 pp (− 35.6 to 0.0)[]
Median cycle fertilisation, % (IQR)	80.0 (56.3–100)	72.1 (25.9–97.9)	P = 0.067[]
Total fertilisation failure, n (%; 95% CI)[]	5 (5.5; 1.8–12.4)	5 (22.7; 7.8–45.4)	P = 0.023[]
By surgical label: fertilised/injected (%); mean cycle %[]			
TESTIS (17 cycles)	—	37/71 (52.1); 47.4	—
TESTIS/MD (2 cycles)	—	7/10 (70.0); 70.0	—
EPID.ASP or EPIS.ASP (3 cycles)	—	13/14 (92.9); 97.2	—

CI, confidence interval; IQR, interquartile range; pp, percentage points (surgically retrieved minus ejaculated). []Exact (Clopper–Pearson) interval. []Descriptive; subgroups too small for comparison. []95% percentile interval from 10,000 bootstrap replications of the patient ID (seed 20261106). []Mann–Whitney U test. []Fisher's exact test.

Table 3
Binomial regression of fertilisation at oocyte level with cluster-robust standard errors (principal analysis).

Model	OR (95% CI)	P	Standardised difference, pp (95% CI)
All surgical labels vs ejaculated (113 cycles, 106 patients)			
Unadjusted	0.64 (0.32–1.28)	0.21	−10.1 (− 26.3 to 6.1)
Adjusted for maternal age	0.64 (0.32–1.25)	0.19	−10.2 (− 26.0 to 5.5)
Adjusted for maternal age and infertility category	0.74 (0.37–1.48)	0.39	−6.5 (− 21.8 to 8.8)
As above + ln(oocytes injected)	0.73 (0.36–1.48)	0.39	−6.7 (− 22.3 to 8.9)
TESTIS label only vs ejaculated (108 cycles, 101 patients)			
Unadjusted	0.46 (0.24–0.91)	0.025	−18.0 (− 34.1 to − 1.8)
Adjusted for maternal age and infertility category	0.54 (0.28–1.07)	0.077	−13.7 (− 29.4 to 2.0)
Covariates in the adjusted all-labels model			
Maternal age, per year	0.99 (0.95–1.04)	0.82	—
Combined vs male factor	2.71 (1.54–4.77)	0.001	—

Binomial (fertilized/injected) logistic regression with robust variance using patient identifier as cluster variable (same as GEE method with independence correlation structure). OR, odds ratio of fertilization with surgically retrieved compared with ejaculated spermatozoa. Standardized differences are estimated by delta method based on model for marginal differences in pooled fertilization. pp, percentage points.

When each cycle was given equal weight, as prespecified in the original analysis plan, mean cycle-level fertilisation was 74.0% with ejaculated and 56.3% with surgically retrieved spermatozoa, a difference of − 17.7 percentage points (95% bootstrap interval − 35.6 to 0.0). This difference is larger than the oocyte-level difference because cycles with one to three oocytes, in which a single unfertilised oocyte changes the cycle rate by 33 to 100 percentage points, carry the same weight as larger cycles. Within the surgically retrieved group, pooled fertilisation was 52.1% (37/71) with the TESTIS label, 70.0% (7/10) with TESTIS/MD and 92.9% (13/14) with the presumed epididymal labels; these subgroups were too small for separate comparison.

Sensitivity analyses

Results of sensitivity analyses are summarized in Table 4. Restriction of the surgically retrieved group to the exact TESTIS category increased the difference in mean cycle-level fertilization to -26.6 percentage points (− 46.1 to − 6.7); the adjusted OR was 0.54 (0.28–1.07). Restriction to male-factor-coded cycles (-14.4; −34.2 to 5.2) or to the first cycle per patient (-17.3; −35.2 to 0.2) gave estimates close to those of the prespecified analysis, while restriction to women younger than 40 years of age gave the larger difference (-25.8; −47.5 to − 4.1). For the cycles with four or more oocytes injected, the difference was small and imprecise (-3.4; −20.7 to 13.2). Stratification by the number of oocytes injected (Fig. 2B) demonstrated that the difference was confined to cycles with one to three oocytes (pooled fertilization 77.0% [77/100] vs 41.2% [7/17]), while it was small for cycles with four or more (67.9% [218/321] vs 64.1% [50/78]). In plain terms, the apparent disadvantage of surgically retrieved spermatozoa was seen almost entirely in cycles with very few oocytes.

Table 4
Sensitivity analysis of the difference in fertilisation (surgically retrieved minus ejaculated).

Analysis	Cycles (EJ / surgical)	Mean cycle difference, pp (95% interval)	Pooled difference, pp (95% interval)
Prespecified analysis (all surgical labels)	91 / 22	−17.7 (− 35.6 to 0.0)	−10.1 (− 27.0 to 5.9)
TESTIS label only	91 / 17	−26.6 (− 46.1 to − 6.7)	−18.0 (− 34.7 to − 1.3)
Testicular labels (TESTIS + TESTIS/MD)	91 / 19	−24.2 (− 43.0 to − 5.3)	−15.8 (− 32.0 to 0.6)
Male-factor-coded cycles only	55 / 18	−14.4 (− 34.2 to 5.2)	−8.2 (− 27.3 to 9.7)
First cycle per patient	84 / 22	−17.3 (− 35.2 to 0.2)	−8.6 (− 25.7 to 7.6)
Maternal age < 40 years	63 / 12	−25.8 (− 47.5 to − 4.1)	−11.3 (− 33.2 to 8.3)
≥ 4 oocytes injected	41 / 12	−3.4 (− 20.7 to 13.2)	−3.8 (− 22.5 to 14.8)

95% percentile confidence interval based on 10,000 bootstrap samples from patient identifier (random seed 20261106). EJ, ejaculated; pp, percentage points. Sensitivity analyses are unadjusted and not independent of the prespecified analysis.

Total fertilisation failure

Total fertilization failure occurred in 5 of 91 ejaculated cycles (5.5%; exact 95% CI 1.8–12.4) and in 5 of 22 surgically retrieved cycles (22.7%; 7.8–45.4; P = 0.023). All ten cases of total fertilization failure occurred in cycles with one to three oocytes injected (10 out of 60 such cycles), and none in the 53 cycles with four or more oocytes. Within cycles with one to three oocytes, failure occurred in 5 of 50 ejaculated (10.0%) and 5 of 10 surgically retrieved cycles (50.0%; P = 0.008, Fisher's exact test). All five failures in the surgically retrieved group had the TESTIS label.

Completeness of embryo-transfer and pregnancy records

Embryo transfer was recorded in 85 of 96 ejaculated and 17 of 22 surgical cycles (Table 5), with a median of 2 (IQR 1–3) and 3 (2–3) embryos transferred, respectively. The T1–T4 category counts matched the number of embryos transferred in all 102 transfers; category T4 was never used. A pregnancy status code was missing for five ejaculated transfers. A PREG code was recorded for 17 ejaculated and 2 surgical transfers, but an affirmative ultrasound entry existed for only 7 and 1 of these, respectively, and miscarriage and delivery were not recorded consistently. Because clinical pregnancy, miscarriage and live birth could not be ascertained, these fields are reported only to document their completeness, and no between-group comparison is made.

Table 5
Embryo transfer and pregnancy records according to sperm source (descriptive; completeness only).

Outcome	Ejaculated (n = 96)	Surgically retrieved (n = 22)	Comment
Cycles with embryo transfer, n	85	17	—
Embryos transferred, median (IQR)	2 (1–3)	3 (2–3)	—
Transferred embryos by category T1 / T2 / T3, n	70 / 73 / 64	18 / 13 / 15	—
Transfers with recorded pregnancy status, n	80	17	Status missing for 5 ejaculated transfers
Pregnancy coded (PREG), n/N (%)[1]	17/80 (21.3)	2/17 (11.8)	Not compared
Pregnancy coded per cycle, n/N (%)	17/96 (17.7)	2/22 (9.1)	—
Affirmative ultrasound entry among coded pregnancies, n/N	7/17	1/2	—
Cycles closed without transfer (CLO), n	11[2]	5[2]	—

Pregnancy status was the code from the worksheet and could not be verified separately; clinical pregnancy, miscarriage and live birth could not be ascertained. Grading T1-T4 definitions and date of evaluation were not known; T4 category was not used. [1]Of those transfers for which status was recorded; shown to document completeness, not for comparison. [2]5 with no oocytes injected, 5 with total fertilisation failure, and 1 with 9 fertilised oocytes without record of the transfer. [2]Of those with total fertilisation failure.

Discussion

Principal findings

In this sample of male-factor and combined-factor ICSI cycles, fertilisation was lower with surgically retrieved than with ejaculated spermatozoa, but the estimate was imprecise and depended on how it

was defined. In the principal oocyte-level analysis, the difference was − 10 percentage points before and − 6.5 percentage points after adjustment for infertility category, with a confidence interval compatible with no difference. The prespecified cycle-level difference was larger (− 17.7 percentage points) but also imprecise, with an interval that reached zero. The excess of low fertilisation in the surgical group was confined to cycles with three or fewer oocytes, where total fertilisation failure accounted for most of the difference. The embryo and pregnancy records were internally consistent but incomplete and were not used to compare clinical outcomes.

Interpretation

Three features of the dataset explain why the estimates differ. First, the unit of analysis matters [30]. More than half of the cycles had three or fewer injected oocytes, so the cycle-level mean is dominated by cycles whose fertilisation rate can take only a few extreme values; the oocyte-level estimate, which accounts for clustering within patients, is more stable and is the more appropriate summary. The high proportion of small cohorts is consistent with a population enriched for advanced maternal age or low oocyte yield, but markers of ovarian reserve were unavailable, so poor ovarian response could not be defined according to accepted criteria [34].

Second, the groups differ by diagnosis, so the comparison is effectively between men with azoospermia and men with spermatozoa in the ejaculate rather than between two interchangeable sperm sources. Surgically retrieved spermatozoa were used predominantly in men with documented azoospermia, whereas the ejaculated group was not demonstrably restricted to severe male-factor infertility and contained a larger share of combined-factor cycles, which had substantially higher fertilisation. Adjustment for infertility category reduced the association by about one-third, but this binary code captures only a small part of the diagnostic difference. The surgical group was itself heterogeneous: it probably included both obstructive and non-obstructive azoospermia, whose fertilisation outcomes differ [41–43], as well as fresh and cryopreserved spermatozoa and different retrieval techniques, none of which were recorded. Residual confounding by the aetiology of azoospermia, testicular histology, sperm cryopreservation, male age, ovarian stimulation and ovarian reserve is therefore likely. Regression cannot remove this confounding by indication, because the dataset contains no population in which both sperm sources were clinical options.

Third, the total fertilisation failures clustered in cycles with very few oocytes. With one to three oocytes, a single unfavorable sperm-oocyte interaction results in complete failure of the fertilisation, and such failures have been attributed to oocyte-activation deficiency, immaturity of spermatozoa and their decondensation problems as well as purely technical reasons [35]. The testicular spermatozoa of men with impaired spermatogenesis may be more prone to such processes [10, 11]; however, the current data set does not allow distinction between the sperm-source effect and chance, diagnosis or oocyte quality. In addition, because fertilisation could not be confirmed as normal (2PN) fertilisation, any between-group difference in abnormal pronuclear patterns could not be assessed and may have biased the comparison in either direction.

Comparison with previous studies

Our findings are directionally consistent with earlier and much larger series. In 1,034 ICSI cycles, normal fertilisation was lower with epididymal and testicular than with ejaculated spermatozoa, although pregnancy rates were comparable [40]. In a Cairo series of 366 cycles, fertilising ability was highest with normal ejaculates and lowest with testicular spermatozoa from men with non-obstructive azoospermia, whereas abnormal ejaculates, epididymal spermatozoa and testicular spermatozoa from men with obstructive azoospermia gave similar fertilisation and pregnancy rates [41]. Large series of testicular-sperm cycles have shown lower fertilisation in non-obstructive than in obstructive azoospermia [42, 43], and a systematic review found lower live-birth rates in non-obstructive azoospermia than in obstructive azoospermia or with ejaculated spermatozoa [45]. ICSI outcomes with ejaculated spermatozoa are largely unrelated to the basic semen parameters [36], and in obstructive azoospermia epididymal and testicular spermatozoa give comparable results [12, 44]. Taken together, these studies indicate that the aetiology of azoospermia, rather than the anatomical source of the spermatozoa, is the main determinant of outcome. Lower fertilisation in a small surgical group consisting mostly of men with azoospermia, and near-complete fertilisation in the even smaller presumed epididymal subgroup (13 of 14 oocytes), are compatible with this literature, but neither subgroup is large enough to support conclusions, and the absence of data on aetiology and retrieval technique prevented the epididymal-versus-testicular and technique-specific comparisons that would be clinically more informative.

The present data do not address the clinically relevant comparisons in which both sperm sources are realistic options. In cryptozoospermia, meta-analyses of predominantly retrospective cohorts have found either no difference between testicular and ejaculated spermatozoa or more favourable outcomes with testicular spermatozoa [15–17]. In men with high sperm DNA fragmentation, testicular spermatozoa carry less DNA damage and have been associated with higher pregnancy rates, although the evidence is of low certainty [21–23]. Paternal effects related to DNA damage are thought to become evident after embryonic genome activation rather than at fertilisation [37, 38], so fertilisation alone is an insensitive outcome for these questions.

Strengths and limitations

The strength of this analysis is the transparency of data integrity checks, reconciling of the source worksheet with its earlier version, analysis at the patient-level, triangulation at cycle-level, oocyte-level, and using adjusted estimates, prespecified sensitivity analyses, and access to the pseudonymised dataset and code for reproduction.

The study has important limitations. Male characteristics (age, semen parameters, sperm DNA fragmentation, hormonal profile and testicular histology), the aetiology of azoospermia, the retrieval technique, the use of fresh or cryopreserved spermatozoa, the ovarian stimulation protocol and markers of ovarian reserve were not available. The infertility category was used as coded in the worksheet and could not be re-verified against semen analyses. Fertilisation could not be confirmed as normal (2PN) fertilisation per metaphase II oocyte. The selection criteria of the source worksheet were unknown.

Embryo-category codes referred only to transferred embryos, pregnancy codes were not independently verified, ultrasound documentation was scant, and miscarriage and live-birth data were lacking, so clinically important reproductive outcomes could not be assessed. Finally, the small surgical group leads to low precision and makes subgroup and sensitivity analyses susceptible to chance findings; the study was not designed or powered to test a clinical hypothesis.

Implications for practice and research

These data cannot help to choose between ejaculated and surgically retrieved spermatozoa in individual couples, for whom diagnosis-specific guidelines should be followed [3, 4]. These data show how the problems associated with choice of the unit of analysis and diagnostic imbalance can increase the apparent difference between two types of spermatozoa in routinely collected ICSI data. In future studies, it is essential to identify a population in which both sperm types are clinically appropriate (for instance, cryptozoospermia or high levels of sperm DNA fragmentation), document the aetiology of azoospermia, the retrieval technique and whether fresh or cryopreserved spermatozoa were used, present data on normal fertilisation rate per injected metaphase II oocyte, assess embryos using consensus criteria [39], and follow up pregnancies until live birth. The analysis models should take into consideration clustering of oocytes within cycles and cycles within couples.

Conclusion

In this single-centre dataset, ICSI with surgically retrieved spermatozoa was associated with lower and more variable fertilisation than ICSI with ejaculated spermatozoa, but the difference was imprecise, was confined to cycles with very few oocytes and diminished after adjustment for infertility category. Because surgical retrieval was used almost exclusively in men with azoospermia, these data cannot inform the choice of sperm source; their main contribution is to show how confounding by indication and the cycle-level unit of analysis can inflate apparent sperm-source differences in routinely collected ICSI data.

Declarations

Ethics approval

and consent. The study protocol was approved by the Ganna Research Ethical Committee; informed consent was waived for this retrospective analysis of anonymised records.

Use of AI-assisted technologies.

[ChatGPT was used to review and edit the manuscript. Authors are fully responsible for the data present in this manuscript]

Funding.

[None]

Disclosure of Potential Conflicts **of Interest.** [The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.]

Author Contribution

[Conceptualization was done by Professor Al-Inany; Data collection; Formal analysis; Methodology; Writing-original draft preparation; Writing-review and editing was done by all authors. All authors have approved the submitted version.]

Availability of data.

The pseudonymised analytical dataset and analysis scripts are available from the corresponding author upon reasonable request and subject to approval by the institutional data custodian.

References

1. Palermo G, Joris H, Devroey P, Van Steirteghem AC (1992) Pregnancies after intracytoplasmic injection of a single spermatozoon into an oocyte. Lancet 340(8810):17–18. 10.1016/0140-6736(92)92425-F
2. Agarwal A, Baskaran S, Parekh N et al (2021) Male infertility. Lancet 397(10271):319–333. 10.1016/S0140-6736(20)32667-2
3. Minhas S, Bettocchi C, Boeri L, Capogrosso P, Carvalho J, Cilesiz NC, EAU Working Group on Male Sexual and Reproductive Health et al (2021) European Association of Urology guidelines on male sexual and reproductive health: 2021 update on male infertility. Eur Urol 80(5):603–620. 10.1016/j.eururo.2021.08.014
4. Schlegel PN, Sigman M, Collura B, De Jonge CJ, Eisenberg ML, Lamb DJ et al (2021) Infertility in males – diagnosis and management of infertility in men: AUA/ASRM Guideline Part II. Fertil Steril 115(1):62–69. 10.1016/j.fertnstert.2020.11.016
5. Practice Committee of the American Society for Reproductive Medicine in (2018) collaboration with the Society for Male Reproduction and Urology. Evaluation of the azoospermic male: a committee opinion. Fertil Steril 109:777–782. 10.1016/j.fertnstert.2018.01.043
6. Esteves SC, Miyaoka R, Agarwal A (2011) Sperm retrieval techniques for assisted reproduction. Int Braz J Urol 37(5):570–583. 10.1590/S1677-55382011000500002
7. Schlegel PN (1999) Testicular sperm extraction: microdissection improves sperm yield with minimal tissue excision. Hum Reprod 14(1):131–135. 10.1093/humrep/14.1.131

8. Dabaja AA, Schlegel PN (2013) Microdissection testicular sperm extraction: an update. Asian J Androl 15:35–39. 10.1038/aja.2012.141
9. Bernie AM, Mata DA, Ramasamy R, Schlegel PN (2015) Comparison between microdissection TESE, conventional TESE, and TESA in non-obstructive azoospermia: a systematic review and meta-analysis. Fertil Steril 104:1099–1103e3
10. Corona G, Minhas S, Giwercman A, Bettocchi C, Dinkelman-Smit M, Dohle G et al (2019) Sperm retrieval and ICSI results in men with non-obstructive azoospermia: a systematic review and meta-analysis. Hum Reprod Update 25(6):733–757. 10.1093/humupd/dmz028
11. Vloeberghs V, Verheyen G, Haentjens P, Goossens A, Polyzos NP, Tournaye H (2015) How effective is TESE-ICSI in patients with non-obstructive azoospermia? Hum Reprod 30(8):1790–1796
12. van Wely M, Barbey N, Meissner A, Repping S, Silber SJ (2015) Is there a difference in live birth rates after MESA or TESE in men with obstructive azoospermia? Hum Reprod 30(4):761–766. 10.1093/humrep/dev032
13. Esteves SC (2015) Clinical management of infertile men with nonobstructive azoospermia. Asian J Androl 17:459–470
14. Bendikson KA, Neri QV, Takeuchi T, Toschi M, Schlegel PN, Rosenwaks Z et al (2008) Outcomes of intracytoplasmic sperm injection utilizing occasional spermatozoa from the ejaculate of men with spermatogenic failure. J Urol 180:1060–1064. 10.1016/j.juro.2008.05.025
15. Abhyankar N, Kathrins M, Niederberger C (2016) Use of testicular versus ejaculated spermatozoa for intracytoplasmic sperm injection among patients with cryptozoospermia: a meta-analysis. Fertil Steril 105:1469–1475e1. 10.1016/j.fertnstert.2016.02.013
16. Kang YN, Hsiao YW, Chen CY, Wu CC (2018) Superiority of testicular spermatozoa over ejaculated spermatozoa for ICSI in cryptozoospermia: an update systematic review and meta-analysis. Sci Rep 8:7874. 10.1038/s41598-018-26280-0
17. Ku FY, Wu CC, Hsiao YW, Kang YN (2018) Association between sperm source and miscarriage and take-home baby after ICSI in cryptozoospermia: a meta-analysis of testicular and ejaculated spermatozoa. Andrology 6:882–889. 10.1111/andr.12546
18. Ben-Ami I, Raziel A, Strassburger D, Komarovsky D, Ron-El R, Friedler S (2013) Intracytoplasmic sperm injection outcomes of ejaculated versus extracted testicular spermatozoa in cryptozoospermic men. Fertil Steril 99(7):1867–1871. 10.1016/j.fertnstert.2013.02.025
19. Hauser R, Bibi G, Yogev L, Carmon A, Azem F, Botchan A et al (2011) Virtual azoospermia and cryptozoospermia—fresh/frozen testicular or ejaculate sperm for better IVF outcome? J Androl 32:484–490. 10.2164/jandrol.110.011353
20. Marinaro JA, Brant A, Kang C, Punjani N, Xie P, Zaninovic N et al (2023) Successful cryptozoospermia management with multiple semen specimen collection. Fertil Steril 120:996–1003. 10.1016/j.fertnstert.2023.07.019
21. Esteves SC, Roque M, Bradley CK, Garrido N (2017) Reproductive outcomes of testicular versus ejaculated sperm for intracytoplasmic sperm injection among men with high levels of DNA

fragmentation in semen: systematic review and meta-analysis. Fertil Steril 108:456–467e1. 10.1016/j.fertnstert.2017.06.018

22. Greco E, Scarselli F, Iacobelli M, Rienzi L, Ubaldi F, Ferrero S et al (2005) Efficient treatment of infertility due to sperm DNA damage by ICSI with testicular spermatozoa. Hum Reprod 20:226–230
23. Moskovtsev SI, Jarvi K, Mullen JB, Cadesky KI, Hannam T, Lo KC (2010) Testicular spermatozoa have statistically significantly lower DNA damage compared with ejaculated spermatozoa in patients with unsuccessful oral antioxidant treatment. Fertil Steril 93:1142–1146
24. Alrabeeah K, Yafi F, Flageole C, Phillips S, Wachter A, Bissonnette F et al (2014) Testicular sperm aspiration for nonazoospermic men: sperm retrieval and intracytoplasmic sperm injection outcomes. Urology 84:1342–1346. 10.1016/j.urology.2014.08.032
25. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP et al (2007) The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 370(9596):1453–1457. 10.1016/S0140-6736(07)61602-X
26. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I et al (2015) The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med 12(10):e1001885. 10.1371/journal.pmed.1001885
27. Zegers-Hochschild F, Adamson GD, Dyer S, Racowsky C, de Mouzon J, Sokol R et al (2017) The International Glossary on Infertility and Fertility Care, 2017. Hum Reprod 32(9):1786–1801. 10.1093/humrep/dex234
28. ESHRE Special Interest Group of Embryology and Alpha Scientists in Reproductive Medicine (2017) The Vienna consensus: report of an expert meeting on the development of ART laboratory performance indicators. Hum Reprod Open 2017(2):hox011. 10.1093/hropen/hox011
29. Wilson EB (1927) Probable Inference, the Law of Succession, and Statistical Inference. J Am Stat Assoc 22(158):209–212. 10.1080/01621459.1927.10502953
30. Vail A, Gardener E (2003) Common statistical mistakes in the design and analysis of subfertility trials. Hum Reprod 18(5):1000–1004
31. Efron B, Tibshirani RJ (1993) An Introduction to the Bootstrap. Chapman & Hall, New York
32. Davison AC, Hinkley DV (1997) Bootstrap Methods and their Application. Cambridge University Press, Cambridge
33. Liang KY, Zeger SL (1986) Longitudinal data analysis using generalized linear models. Biometrika 73(1):13–22. 10.1093/biomet/73.1.13
34. Ferraretti AP, La Marca A, Fauser BC, Tarlatzis B, Nargund G, ESHRE working group on Poor Ovarian Response Definition (2011) Consensus on the definition of 'poor response' to ovarian stimulation for in vitro fertilization: the Bologna criteria. Hum Reprod 26(7):1616–1624. 10.1093/humrep/der092
35. Flaherty SP, Payne D, Matthews CD (1998) Fertilization failures and abnormal fertilization after intracytoplasmic sperm injection. Hum Reprod 13(Suppl 1):155–164

36. Nagy ZP, Liu J, Joris H, Verheyen G, Tournaye H, Camus M et al (1995) The outcome of intracytoplasmic sperm injection is unrelated to any of the three basic sperm parameters. Hum Reprod 10:1123–1129. 10.1093/oxfordjournals.humrep.a136104
37. Tesarik J, Greco E, Mendoza C (2004) Late, but not early, paternal effects on human embryo development is related to sperm DNA fragmentation. Hum Reprod 19:611–615. 10.1093/humrep/deh127
38. Colaco S, Sakkas D (2018) Paternal factors contributing to embryo quality. J Assist Reprod Genet 35:1953–1968. 10.1007/s10815-018-1304-4
39. 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. Hum Reprod 40(6):989–1035. 10.1093/humrep/deaf021
40. Nagy Z, Liu J, Janssenswillen C, Silber S, Devroey P, Van Steirteghem AC (1995) Using ejaculated, fresh, and frozen-thawed epididymal and testicular spermatozoa gives rise to comparable results after intracytoplasmic sperm injection. Fertil Steril 63(4):808–815. 10.1016/s0015-0282(16)57486-x
41. Aboulghar MA, Mansour RT, Serour GI, Fahmy I, Kamal A, Tawab NA, Amin YM (1997) Fertilization and pregnancy rates after intracytoplasmic sperm injection using ejaculate semen and surgically retrieved sperm. Fertil Steril 68(1):108–111. 10.1016/s0015-0282(97)81484-7
42. De Croo I, Van der Elst J, Everaert K, De Sutter P, Dhont M (2000) Fertilization, pregnancy and embryo implantation rates after ICSI in cases of obstructive and non-obstructive azoospermia. Hum Reprod 15(6):1383–1388. 10.1093/humrep/15.6.1383
43. Vernaeve V, Tournaye H, Osmanagaoglu K, Verheyen G, Van Steirteghem A, Devroey P (2003) Intracytoplasmic sperm injection with testicular spermatozoa is less successful in men with nonobstructive azoospermia than in men with obstructive azoospermia. Fertil Steril 79(3):529–533. 10.1016/S0015-0282(02)04809-4
44. Nicopoullos JDM, Gilling-Smith C, Almeida PA, Norman-Taylor J, Grace I, Ramsay JWA (2004) Use of surgical sperm retrieval in azoospermic men: a meta-analysis. Fertil Steril 82(3):691–701. 10.1016/j.fertnstert.2004.02.116
45. Esteves SC, Agarwal A (2013) Reproductive outcomes, including neonatal data, following sperm injection in men with obstructive and nonobstructive azoospermia: case series and systematic review. Clin (Sao Paulo) 68(Suppl 1):141–150
46. Clopper CJ, Pearson ES (1934) The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 26(4):404–413. 10.1093/biomet/26.4.404

Figures

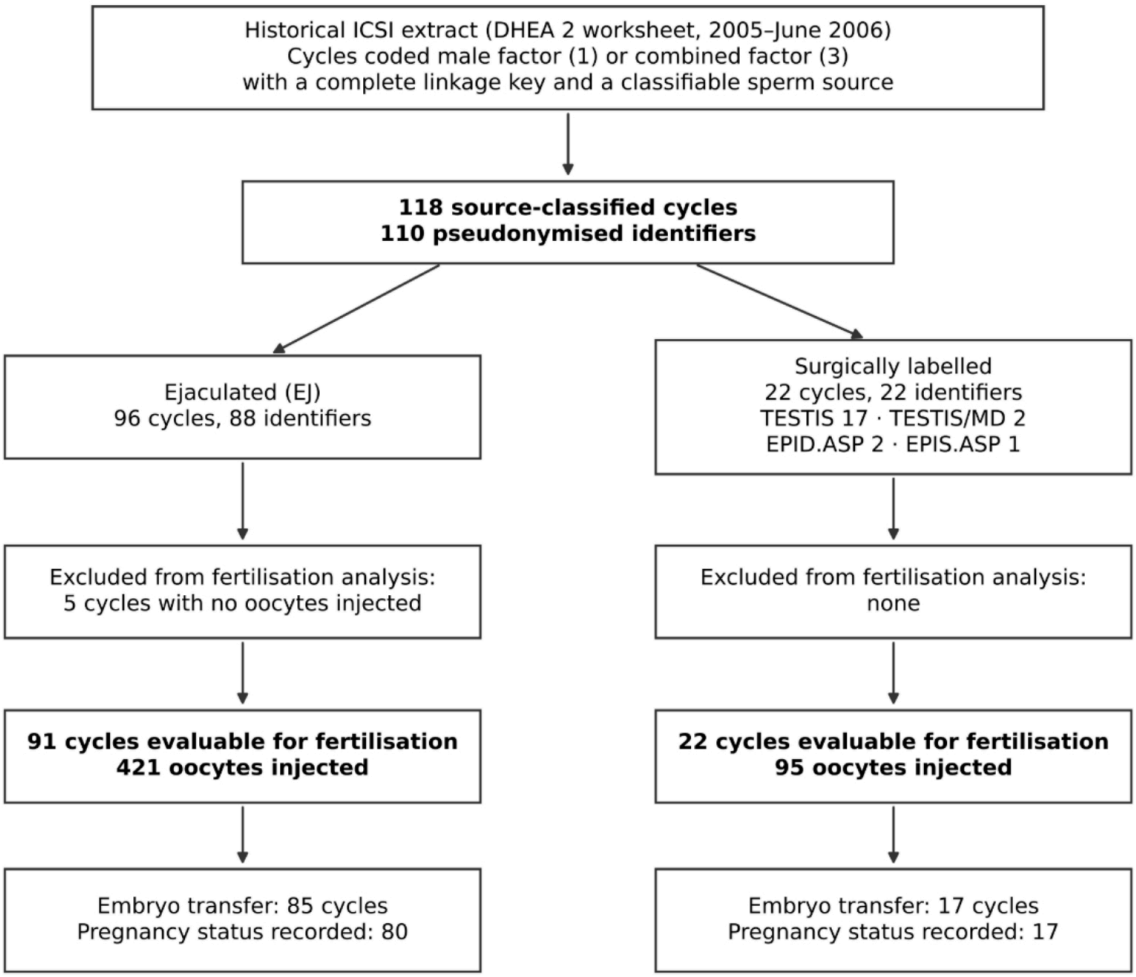


Figure 1

Flow diagram of cycles participating in the study. Selection criteria of the source worksheet are unknown. Pregnancy status refers to the worksheet code (PREG/NEG) among cycles with embryo transfer.

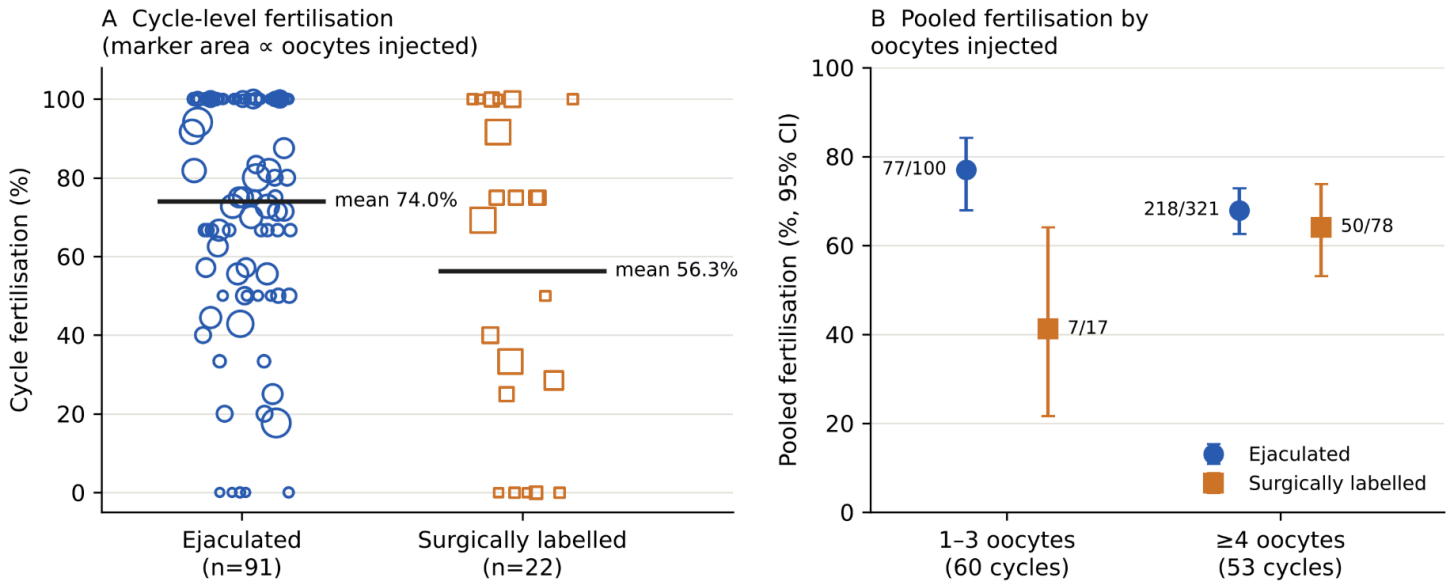


Figure 2

(A) Fertilisation (number of fertilised/oocytes injected) by sperm source; each point represents one cycle, size of point corresponds to number of oocytes injected, horizontal bar is mean. (B) Fertilisation (95% Wilson interval) by sperm source stratified by number of oocytes injected (1-3 vs. ≥4); labels indicate fertilised/oocytes injected.

Supplementary Files

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

- SpermSourceICSICycles.xls