Human Reproduction Open, 2023, **2023(2),** hoad018
https://doi.org/10.1093/hropen/hoad018
Advance Access Publication Date: May 4, 2023
Debate

Conventional outcome reporting per IVF cycle/embryo transfer may systematically underestimate chances of success for women undergoing ART: relevant biases in registries, epidemiological studies, and guidelines

Georg Griesinger [1,*] and Per Larsson[2]

[1]Department of Gynaecological Endocrinology and Reproductive Medicine, University Hospital of Schleswig-Holstein, Campus Lübeck, Lübeck, Germany
[2]Global Biometrics, Global Clinical Development, Ferring Pharmaceuticals, Kastrup, Denmark

*Correspondence address. University of Luebeck and University Hospital of Schleswig-Holstein, Campus Luebeck, Ratzeburger Allee 160, 23538 Luebeck. Germany. Tel: +49-451-505-778-212; Fax: +49-451-505-778-210; E-mail: georg.griesinger@uni-luebeck.de https://orcid.org/0000-0002-0606-5804

ABSTRACT

Pre-conception counselling and management of expectations about chance of success of IVF/ICSI treatments is an integral part of fertility care. Registry data are usually used to inform patients about expected success rates of IVF/ICSI treatment, as these data should best represent real-world populations and clinical practice. In registries, the success rate of IVF/ICSI treatments is conventionally reported per treatment cycle or per embryo transfer and estimated from data for which several treatment attempts per subject have been pooled (e.g. repetitive IVF/ICSI attempts or repetitive attempts of cryotransfer). This, however, may underestimate the true mean chance of success per treatment attempt, because treatment attempts of women with a poor prognosis will usually be over-represented in a pool of treatment cycle data compared to treatment events of women with a good prognosis. Of note, this phenomenon is also a source of potential bias when comparing outcomes between fresh transfers and cryotransfers, since women can undergo a maximum of only one fresh transfer after each IVF/ICSI treatment, but potentially several cryotransfers. Herein, we use a trial dataset from 619 women, who underwent one cycle of ovarian stimulation and ICSI, a Day 5 fresh transfer and/or subsequent cryotransfers (follow-up of all cryotransfers up to 1 year after the start of stimulation), to exemplify the underestimation of the live birth rate, when not accounting for repeated transfers in the same woman. Using mixed-effect logistic regression modelling, we show that the mean live birth rate per transfer per woman in cryocycles is underestimated by the factor 0.69 (e.g. live birth rate per cryotransfer of 36% after adjustment versus 25% unadjusted). We conclude that the average chance of success of treatment cycles of women of a given age, treated in a given centre, etc., when conventionally calculated per cycle or per embryo transfer from a pool of treatment events, do not apply to an individual woman. We suggest that patients are, especially at the outset of treatment, systematically confronted with mean estimates of success per attempt that are too low. Live birth rates per transfer from datasets encompassing multiple transfers from single individuals could be more accurately reported using statistical models accounting for the correlation between cycle outcomes within women.

Keywords: IVF outcome / IVF registry / IVF success / pregnancy rate / embryo transfer

Introduction

Pre-conception counselling about the chance of success of IVF/ICSI treatment and management of expectations is an integral part of fertility care. Various national and international registries have been set up to collect the number and types of IVF/ICSI treatments as well as the related clinical outcomes (Harton *et al.*, 2011; De Geyter *et al.*, 2015; Kadi and Wiesing, 2016; European IVF Monitoring Consortium (EIM), for the European Society of Human Reproduction and Embryology (ESHRE) *et al.*, 2022; Henderson *et al.*, 2023; Jain and Singh, 2022). The outcomes pregnancy and live birth are conventionally reported per treatment cycle or per embryo transfer (ET), calculated from a pool of treatment attempts performed within the timespan of a calendar year (ICMART, 2019). Examples of such reporting are abundant in the literature, such as a recent HFEA report on their 'One at a time' campaign (HFEA, 2022) that states that 'In 2008 the overall pregnancy rate [per embryo transfer] was 30% compared with 34% in 2013 [….]. Looking just at fresh transfers, the rate is slightly higher and shows the same increase over the years, from 32% to 36% between 2008 and 2013' (HFEA, 2015). The German IVF registry likewise uses outcome per cycle or per ET as their standard reporting and state in their latest published report that 'We are happy to note that the pregnancy rate per transfer in fresh cycles was 31.9%, whereas the pregnancy rate per transfer in cryocycles was 29.4%' (Bartnitzky *et al.*, 2021). The Centers for Disease Control and Prevention (CDC) of the USA aims at instructing patients and the general public on ART success rates as follows: 'The information in this section is provided to help consumers

Received: January 17, 2023. **Revised:** March 14, 2023. **Editorial decision:** April 25, 2023.

navigate and understand the information presented online, explore clinic services, see the types of patients each clinic treats, and understand fertility clinic success rates based on the latest data from the National ART Surveillance System' (CDC, 2021). For this purpose, the CDC reports success rates per clinic cumulatively per cycle ('cumulative percentages of intended egg retrievals that resulted in a live birth delivery after all associated egg or embryo transfers within 12 months of cycle start') and per ET ('percentages of transfers with at least one egg or embryo that resulted in a live-birth delivery').

During IVF/ICSI treatment, however, the majority of women undergo more than one treatment cycle before achieving a live birth or before discontinuing treatment. For example, in Germany, women on average underwent 1.9 treatment cycles during 2020 (Bartnitzky *et al.*, 2021). Furthermore, even after a single cycle of ovarian stimulation and IVF/ICSI, different numbers of ETs, fresh and frozen–thawed (cryotransfers) may follow in different women. As exemplified above, treatment success is still frequently estimated from data for which several treatment attempts per subject have been pooled (e.g. repetitive IVF/ICSI attempts with fresh transfers and/or repetitive attempts of cryotransfer).

Are we systematically underestimating our patients' chances of success per treatment cycle using pooled data?

The success rate per cycle or per ET cycle may be severely underestimated from pooled treatment cycle data, since treatment events in women with a poor prognosis will be over-represented in the dataset compared to treatment events in women with a good prognosis. This is because women achieving live birth early in the course of their treatment will contribute fewer treatment cycles and ETs, whereas women with later achievement or no achievement of live birth will contribute more treatment cycles and ETs to the pool of observations. Of note, this phenomenon is another source of potential bias when comparing outcome between fresh transfers and cryotransfers, since women can undergo a maximum of only one fresh transfer after IVF/ICSI treatment, but potentially several cryotransfers.

An example calculation: live birth rates underestimation in cryocycles

The complete dataset from the RAINBOW trial (Fernandez Sanchez *et al.*, 2022) is used herein to exemplify the underestimation of live birth per cryotransfer for individual women undergoing one ICSI cycle with or without fresh transfer and subsequent cryocycles. In total, 619 women underwent treatment with a fixed, individualized daily dose of follitropin delta with or without additional rhCG in a long GnRH agonist protocol followed by ICSI and a Day 5 fresh transfer. Surplus blastocysts were cryopreserved on Day 5 or Day 6 and subsequently used for cryotransfers up to 1 year after start of the stimulation.

Of the 619 women starting stimulation, 557 had at least one fresh or cryotransfer. A total of 927 ETs were performed, of which 520 were fresh transfers and 407 were cryotransfers. In total, 252 subjects had at least one cryotransfer, 102 had a second cryotransfer, 37 had a third, 12 had a fourth, three had a fifth, and one had a sixth cryotransfer. The timing of all fresh and cryotransfers and all live births are illustrated in Fig. 1, with digits indicating the order of the transfer. The live birth rates for each transfer are summarized in Fig. 2 and, as expected, the live birth

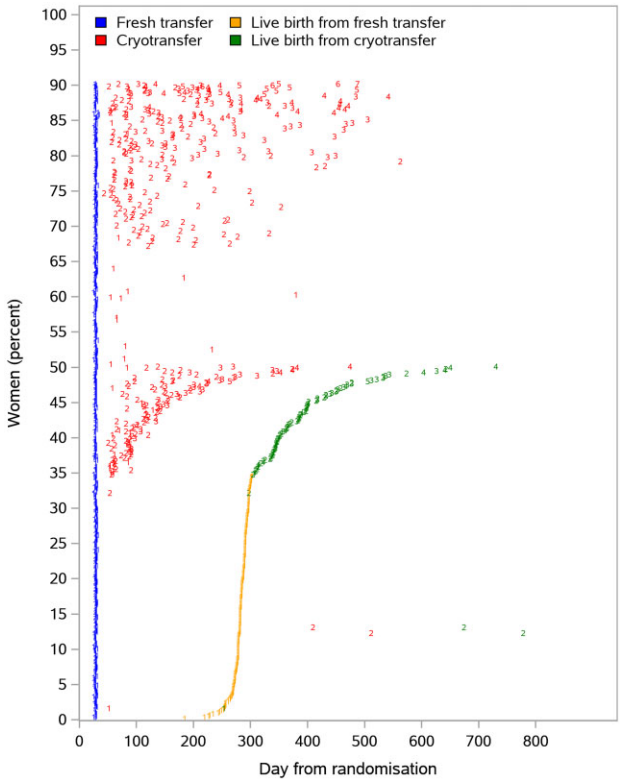

Figure 1. Timepoints of all fresh transfers, cryotransfers, and live births in the RAINBOW trial. Digits represent the order of the transfers, with one being the first transfer (fresh transfer, or the first cryotransfer for subjects without fresh transfer). The women are sorted by time to first live birth, and shown as percent of all women, so that the live birth digits form a cumulative live birth rate versus time curve.

rate decreases with increasing cycle rank. Of the 520 fresh transfers, 212 (41%) resulted in a live birth, whereas for the 407 cryotransfers, 100 (25%) resulted in a live birth (unadjusted difference, 16%).

To provide a more accurate comparison between fresh and cryotransfers, live birth rate per transfer was analysed using a mixed-effect logistic regression model, accounting for repeated transfers in the same woman and where women's log-odds of achieving live birth was assumed to be normally distributed within the trial population. The type of transfer (fresh or cryo) was included as a fixed factor in the model. The relationship between log-odds and probability (p = exp(log-odds)/(1 + exp(log-odds)) was subsequently used to obtain estimates of live birth rates for fresh and cryotransfers, but in order to account for the random woman factor in the model, the estimated live birth rates per transfer ($E(p_i)$) were calculated as follows:

$$E(p_i) = \frac{1}{\sqrt{2\pi}s}\int_{-\infty}^{\infty} \frac{e^x}{1+e^x} e^{-(1/2)\left(x-\text{logodds}_i/s\right)^2} dx,$$

where s denotes the standard deviation of the normal distribution in the mixed-effect model and i = (fresh or cryotransfer). The delta method (Doob, 1935) was used to calculate the confidence interval and p-value for the difference in mean live birth rate between fresh transfers and cryotransfers.

When using this analysis to adjust for repeated transfers in the same women, the live birth rates were estimated to be 41% per fresh transfer and 36% per cryotransfer. The difference between live birth rates per fresh transfer and per cryotransfer was estimated to be 5.3%, with a 95% confidence interval ranging

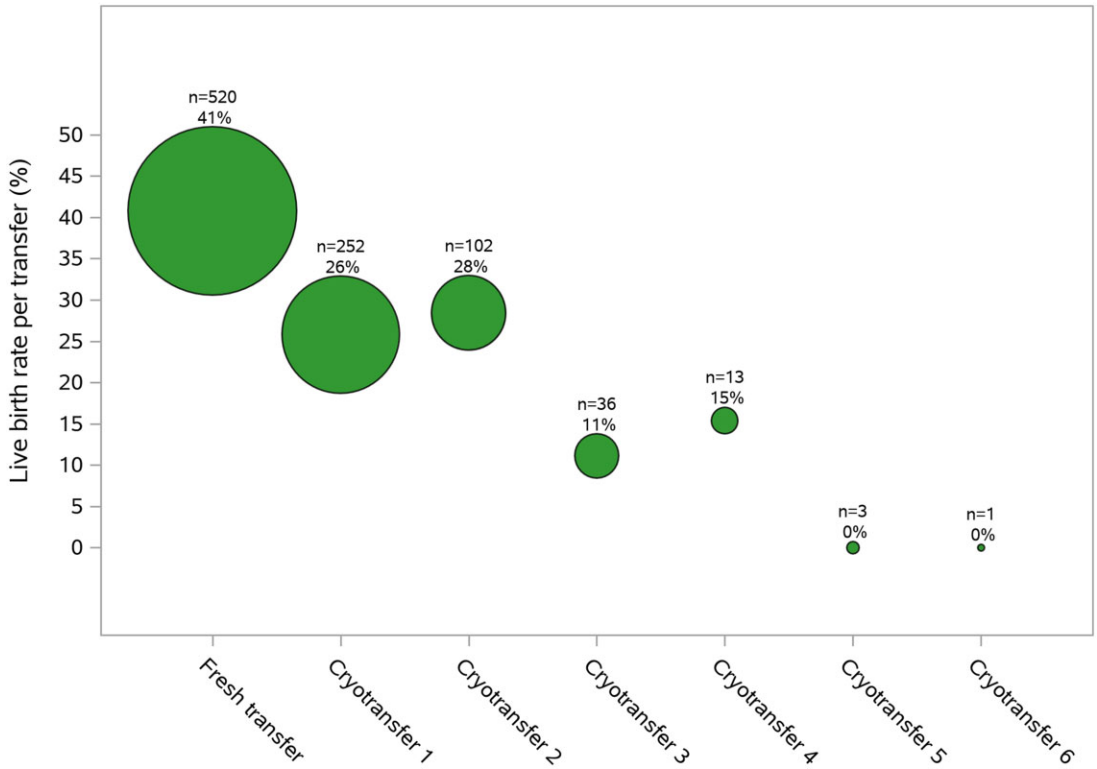

Figure 2. Live birth rate per transfer in the RAINBOW trial. The size of the bubbles indicates the number of women with fresh transfer or cryotransfer.

from –6.6% to 17.1%, and with no statistically significant difference between fresh and cryotransfers in chance of success ($P = 0.3827$). This example quantifies the underestimation of the live birth rate of women undergoing cryocycles by a factor of 0.69 (e.g. absolute 11%, relative 30% alluding to an individual woman's chance of live birth per cryocycle being 36% and not 25%).

Cryocycle success rates: better or worse by biology or by statistics?

Until 2014, success rates per cryotransfer were consistently reported to be lower as compared to fresh cycle success rates in registry data (European IVF Monitoring Consortium (EIM), for the European Society of Human Reproduction and Embryology (ESHRE) *et al.*, 2022). This phenomenon has so far mostly been explained by selection bias in an era of predominant fresh cycle IVF (European IVF Monitoring Consortium (EIM), for the European Society of Human Reproduction and Embryology (ESHRE) *et al.*, 2022), both due to patient factors (e.g. except for occasional freeze-all cycles, women will only undergo a cryocycle if not pregnant after the fresh transfer) and embryo factors (the top quality embryos have usually preferentially been transferred in the fresh cycle, leaving only lower quality embryos for cryopreservation), respectively. Of note, freezing–thawing per se appears not to induce damage, as randomized trials comparing a freeze-all and later cryotransfer versus immediate fresh transfer strategy have shown that freezing–thawing with state-of-the-art technology is not associated with worse outcomes (Bosdou *et al.*, 2019).

The selection biases when comparing live birth rates between fresh and cryotransfers also become apparent in the RAINBOW trial dataset. Figure 3 shows live birth rates versus embryo quality for fresh and cryotransfers as a bubble plot, with the size of bubbles indicating the percentage of transfers performed with each embryo quality. Blastocysts were scored by using the system of Gardner and Schoolcraft (1999) with the addition of D-categories for inner cell mass and trophectoderm scoring. Figure 3

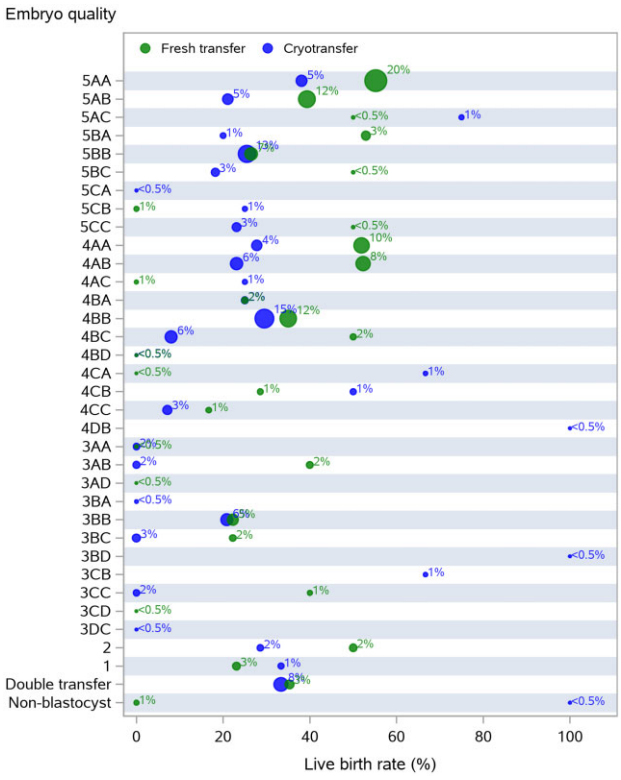

Figure 3. Live birth rate for different embryo qualities and for fresh and cryotransfers in the RAINBOW trial. The size of the bubbles (and the data labels) indicates the percentage of transfers performed with each embryo quality, with fresh and cryotransfers each summing up to 100%.

shows that the percentages of transfers using high-quality blastocysts were higher for fresh transfers compared with cryotransfers (i.e. 5AA: 20% vs 5%, 5AB: 12% vs 5%, 4AA: 10% vs 4%). In addition, the figure shows that for most quality gradings, the live birth rates were higher for fresh transfers (i.e. 5AA: 55% vs 38%, 5AB: 39 vs 21%, 4AA: 52% vs 28%), a difference largely attributable to the biases discussed above. In conclusion, comparisons of crude live birth rates between fresh and cryotransfers are, in a population with predominant fresh transfer after IVF/ICSI, already considerably biased by differences in both embryo quality and prognosis of populations.

Selection bias may, however, also favour cryocycles. More recently, freeze-all cycle usage has dramatically increased (European IVF Monitoring Consortium (EIM), for the European Society of Human Reproduction and Embryology (ESHRE) *et al.*, 2022). A freeze-all strategy is often done in high responders predicted to be at risk of ovarian hyperstimulation syndrome (Griesinger *et al.*, 2007, 2016; Bosch *et al.*, 2020) or in conjunction with pre-implantation genetic testing for aneuploidy (PGT-A) (Sermon *et al.*, 2016; Irani *et al.*, 2020). The good prognosis, high-response patients are thereby excluded from the fresh cycle pool of data, and, when PGT-A is applied, only positive selected embryos will be transferred in cryocycles, while the deselection of genetically abnormal embryos will decrease the number of cryocycles an individual can contribute to a pool of treatment cycles recorded in a registry.

We suggest that the lack of adjustment for multiple successive cryotransfers with increasingly poorer prognosis occurring in the same individual may represent yet another source of relevant bias. In our example, after adjusting for the number of ETs, the individual woman's success rate per cryocycle is similar to the fresh cycle success rate (Weiss *et al.*, 2023). Accordingly, the putative low pregnancy rate historically reported by registries for cryocycles may have led to an exaggerated concern about putative cryo-damage or futility of cryo-treatments. In addition, adjusted estimates would most likely not have lent support to multiple ETs for each cryocycle attempt in order to correct for a putative lower chance of success.

The couple as the correct 'unit-of-analysis'

Taking the multiple sources of bias into account, one would be very reluctant to draw causal inferences from registry data (Wilkinson *et al.*, 2019), as the distortion of the success rate when calculated from pooled treatment events is not limited to cryocycles. It will likewise occur when the outcome is calculated per initiated IVF treatment cycle, per oocyte retrieval event, or as a cumulation of success from all ETs after one ovarian stimulation cycle with IVF/ICSI treatment. Mean success rates, unadjusted for number of treatment attempts per patient, will thereby vary depending on the number of treatment attempts women will undergo. This in turn will depend on a multitude of factors, one of which is likely to be funding of IVF/ICSI cycles (e.g. without reimbursement, it is expected that on average fewer treatment attempts would be undertaken by each woman and therefore fewer cycles in higher treatment ranks will be contributing to the erroneous denominator). This also illustrates why international comparisons of pooled per-cycle outcome data do not make sense without a minimum understanding of the underlying number of cycles and/or ETs per woman (Fauser, 2019).

Efficacy and safety of a treatment should ideally be established in robust clinical trials with the appropriate unit-of-analysis being the patient, not the treatment cycle, and not the embryo (Vail and Gardener, 2003; Griesinger, 2016); however, instead there has been a call for using 'big data', 'algorithms to embrace confounding', and 'real-world evidence' (Cohen and Alikani, 2013). A recent review article has summarized the problems and perils of using large observational databases to address pertinent questions in reproductive medicine (Wilkinson *et al.*, 2019). Herein, we highlight just another one, namely the treatment cycle as the unit-of-analysis in most registries (Adamson *et al.*, 2018) and the ramifications of unadjusted success-rate reporting. Given the necessary high level of data protection, it may not be easy to collect non-anonymized data (e.g. for the number of cycles per patient) in institutions outside the one where the health-care providers practise. However, this would be a necessity for an appropriate adjustment of cycle data outcomes. Future efforts should also address the issue of infertile couple versus infertile female patient as the relevant unit-of-analysis, the issue of cross-over of patients between treatment centres, and the issue of reporting outcomes over longer time spans than a single calendar year.

Of note, the present analysis is not to be mistaken with an outcome prediction modelling approach, in which all known covariates must be considered while adjusting for multiple correlated treatment cycles within individuals in order to arrive at a best fit model of treatment success over time. A number of such models have already been developed and the methodological issues when analysing multiple-cycle data from couples undergoing IVF have been described in that context (Missmer *et al.*, 2011; Yland *et al.*, 2019).

What next for success-rate reporting from registry data?

Where feasible, registries could report adjusted mean success rates per treatment attempt, as exemplified herein. As a second-best option, registries could report success rates separately for treatment cycle ranks. As a minimum, and for easy patient counselling, at least first cycle rank cumulative success rates should be reported per female age stratum. This would help in setting patient expectations at the outset of treatment as well as explaining that success rates decrease per repetitive attempt. Where applicable, existing prediction algorithms should be used if based on the correct unit of analysis, i.e. the couple, and adjusted for treatment rank, among other relevant factors, such as female age, parity, or treatment type (SART Predictor; Wang *et al.*, 2022).

Conclusion

After 40 years of cycle data collection in ART registries (Felderbaum and Dahncke, 2020), progress should be made in our field in how we better inform patients and the public on treatment success perspectives. The correct unit-of-analysis should be the couple, not the cycle. Live birth rates per transfer from datasets encompassing multiple transfers from single individuals could be more accurately reported using statistical models accounting for the correlation between cycle outcomes within women.

Data availability

No new data were generated in support of this work.

Acknowledgements

We thank the investigators and participants of the RAINBOW trial.

Authors' roles

P.L. conducted the statistical analyses. P.L. and G.G. interpreted the analyses, wrote and reviewed the manuscript.

Funding

The original RAINBOW trial was funded by Ferring Pharmaceuticals. The authors declare no additional funding was provided for this work.

Conflict of interest

G.G. received no financial support for the research, authorship, and/or publication of this article. G.G. has, outside from this work, received honoraria and/or non-financial support (e.g. travel cost compensation) from Abbott, Ferring, Gedeon Richter, Guerbet, Merck, Organon, MSD, ObsEva, PregLem, ReprodWissen GmbH, Vifor and Cooper. P.L. is an employee of Ferring Pharmaceuticals.

References

Adamson GD, De Mouzon J, Chambers GM, Zegers-Hochschild F, Mansour R, Ishihara O, Banker M, Dyer S. International Committee for Monitoring Assisted Reproductive Technology: world report on assisted reproductive technology, 2011. Fertil Steril 2018;**110**:1067–1080.

Bartnitzky S, Blumenauer V, Czeromin U, Fehr D, Gnoth C, Grewe C, Krüssel J-S, Kupka MS, Tandler-Schneider A, Tauchert S. D·I·R Annual 2020—the German IVF-Registry. *J Reprod Med Endocrinol* 2021;**18**:204–247.

Bosch E, Broer S, Griesinger G, Grynberg M, Humaidan P, Kolibianakis E, Kunicki M, Marca L, Lainas A, Le Clef G *et al.* ESHRE guideline: ovarian stimulation for IVF/ICSI. *Hum Reprod Open* 2020;**2020**:hoaa009.

Bosdou JK, Venetis CA, Tarlatzis BC, Grimbizis GF, Kolibianakis EM. Higher probability of live-birth in high, but not normal, responders after first frozen-embryo transfer in a freeze-only cycle strategy compared to fresh-embryo transfer: a meta-analysis. *Hum Reprod* 2019;**34**:491–505.

CDC. *How to Access and Interpret Fertility Clinic Success Rates*. Centers for Disease Control and Prevention, 2021. Available at: https://www.cdc.gov/art/reports/2019/how-to-access-interpret.html (May 2023, date last accessed).

Cohen J, Alikani M. Evidence-based medicine and its application in clinical preimplantation embryology. *Reprod Biomed Online* 2013;**27**:547–561.

De Geyter C, Fehr P, Moffat R, Gruber IM, von Wolff M. Twenty years' experience with the Swiss data registry for assisted reproductive medicine: outcomes, key trends and recommendations for improved practice. *Swiss Med Wkly* 2015;**145**:w14087.

Doob JL. The limiting distributions of certain statistics. *Ann Math Stat* 1935;**6**:160–169.

European IVF Monitoring Consortium (EIM), for the European Society of Human Reproduction and Embryology (ESHRE); Wyns C, De Geyter C, Calhaz-Jorge C, Kupka MS, Motrenko T, Smeenk J, Bergh C, Tandler-Schneider A, Rugescu IA *et al.* ART in Europe, 2018: results generated from European registries by ESHRE. *Hum Reprod Open* 2022;**2022**:hoac022.

Fauser BC. Towards the global coverage of a unified registry of IVF outcomes. *Reprod Biomed Online* 2019;**38**:133–137.

Felderbaum R, Dahncke W. Das Deutsche IVF-Register als Instrument der Qualitätssicherung und zur Beratung der Patienten. *Der Gynäkologe* 2020;**33**:800–811.

Fernandez Sanchez M, Visnova H, Larsson P, Yding Andersen C, Filicori M, Blockeel C, Pinborg A, Khalaf Y, Mannaerts B; Rainbow Study Group. A randomized, controlled, first-in-patient trial of choriogonadotropin beta added to follitropin delta in women undergoing ovarian stimulation in a long GnRH agonist protocol. *Hum Reprod* 2022;**37**:1161–1174.

Gardner DK, Schoolcraft WB. In vitro culture of human blastocysts. In Jansen R, Mortimer D (eds). *The 11th World Congress on in Vitro Fertilization and Human Reproductive Genetics*. Carnforth, UK: The Panthenon Publishing Group LTD, 1999, 378–388.

Griesinger G. Beware of the 'implantation rate'! Why the outcome parameter 'implantation rate' should be abandoned from infertility research. *Hum Reprod* 2016;**31**:249–251.

Griesinger G, Verweij PJM, Gates D, Devroey P, Gordon K, Stegmann BJ, Tarlatzis BC. Prediction of ovarian hyperstimulation syndrome in patients treated with Corifollitropin alfa or rFSH in a GnRH antagonist protocol. *PLoS One* 2016;**11**:e0149615.

Griesinger G, von Otte S, Schroer A, Ludwig AK, Diedrich K, Al-Hasani S, Schultze-Mosgau A. Elective cryopreservation of all pronuclear oocytes after GnRH agonist triggering of final oocyte maturation in patients at risk of developing OHSS: a prospective, observational proof-of-concept study. *Hum Reprod* 2007;**22**:1348–1352.

Harton GL, De Rycke M, Fiorentino F, Moutou C, SenGupta S, Traeger-Synodinos J, Harper JC; European Society for Human Reproduction and Embryology PGD Consortium. ESHRE PGD consortium best practice guidelines for amplification-based PGD. *Hum Reprod* 2011;**26**:33–40.

Henderson I, Lacey L, Akhtar MA, Quenby S. Ethnic group and reason for assisted reproductive technology failure: analysis of the Human Fertilisation and Embryology Authority registry data from 2017 to 2018. *Fertil Steril* 2023;**119**:241–249.

HFEA. *Improving Outcomes for Fertility Patients: Multiple Births*. 2015. Available at: https://ifqlive.blob.core.windows.net/umbraco-website/1169/multiple_births_report_2015.pdf (May 2023, date last accessed).

HFEA. Our campaign to reduce multiple births 2022. Available at: https://www.hfea.gov.uk/about-us/our-campaign-to-reduce-multiple-births/ (May 2023, date last accessed).

ICMART. *The ICMART Tool Box for ART Data Collection*. ICMART. 2019. Available at: https://www.icmartivf.org/wp-content/uploads/ICMART-toolbox.pdf (May 2023, date last accessed).

Irani M, Canon C, Robles A, Maddy B, Gunnala V, Qin X, Zhang C, Xu K, Rosenwaks Z. No effect of ovarian stimulation and oocyte yield on euploidy and live birth rates: an analysis of 12 298 trophectoderm biopsies. *Hum Reprod* 2020;**35**:1082–1089.

Jain M, Singh M. Assisted reproductive technology (ART) techniques. In: *StatPearls [Internet]*. Treasure Island (FL) StatPearls Publishing 2022.

Kadi S, Wiesing U. The German IVF register as an instrument to document assisted reproductive technologies. *Geburtshilfe Frauenheilkd* 2016;**76**:680–684.

Missmer SA, Pearson KR, Ryan LM, Meeker JD, Cramer DW, Hauser R. Analysis of multiple-cycle data from couples undergoing in vitro fertilization: methodologic issues and statistical approaches. *Epidemiology* 2011;**22**:497–504.

SART Predictor. Society for Assisted Reproductive Technology. What are My Chances With ART. https://www.sartcorsonline.com/Predictor/Patient (May 2023, date last accessed).

Sermon K, Capalbo A, Cohen J, Coonen E, De Rycke M, De Vos A, Delhanty J, Fiorentino F, Gleicher N, Griesinger G *et al.* The why, the how and the when of PGS 2.0: current practices and expert opinions of fertility specialists, molecular biologists, and embryologists. *Mol Hum Reprod* 2016;**22**:845–857.

Vail A, Gardener E. Common statistical errors in the design and analysis of subfertility trials. *Hum Reprod* 2003;**18**:1000–1004.

Wang CW, Kuo CY, Chen CH, Hsieh YH, Su EC. Predicting clinical pregnancy using clinical features and machine learning algorithms in in vitro fertilization. *PLoS One* 2022;**17**:e0267554.

Weiss MS, Luo C, Zhang Y, Chen Y, Kissin DM, Satten GA, Barnhart KT. Fresh vs. frozen embryo transfer: new approach to minimize the limitations of using national surveillance data for clinical research. *Fertil Steril* 2023;**119**:186–194.

Wilkinson J, Brison DR, Duffy JMN, Farquhar CM, Lensen S, Mastenbroek S, van Wely M, Vail A. Don't abandon RCTs in IVF. We don't even understand them. *Hum Reprod* 2019;**34**:2093–2098.

Yland J, Messerlian C, Minguez-Alarcon L, Ford JB, Hauser R, Williams PL; EARTH Study Team. Methodological approaches to analyzing IVF data with multiple cycles. *Hum Reprod* 2019;**34**:549–557.