

COMMENTARY

Measuring success in IVF is a complex multidisciplinary task: time for a consensus?

Laura Rienzi^{1,*}, Danilo Cimadomo¹, Alberto Vaiarelli¹, Gianluca Gennarelli², Jan Holte³, Claudia Livi⁴, Monica Aura Masip⁵, Petr Uher⁶, Gemma Fabozzi¹, Filippo Maria Ubaldi¹

ABSTRACT

The goal of an IVF cycle is the birth of at least one baby per intention to treat. However, IVF cannot confer competence on an embryo, but only can provide each couple with a safe treatment to meet a predetermined chance of success. This commentary highlights how clinical, financial and patient-centred perspectives should be included in the definition of success in IVF. The primary outcome, which is the cumulative live birth delivery rate per intention to treat, must always be complemented by analyses of risks, costs and time invested, as well as by measures of patient satisfaction. Finally, it is essential, whenever clinical conditions exist, to limit treatment discontinuation after failed attempts. Constant monitoring of the data is pivotal and must be adjusted for patient characteristics and compared with national and international registers. The authors aimed to review all these aspects and highlight the points that are still open for discussion. Is it time for a consensus?

INTRODUCTION

The aim of IVF is the birth of at least one baby per intention to treat (ITT). However, the chances of success are low ($\leq 30\%$ per cycle), and competence cannot be conferred on gametes retrieved from infertile patients. Therefore, the quality of treatment in reproductive medicine is of considerable importance, beyond its final outcome. The main purpose of a clinic must be to monitor and fulfil both the clinical and technical key performance indicators of IVF. IVF practitioners must rely solely on validated and reproducible practices.

Only once the efficacy of the IVF clinic at minimum fulfils the competence (if not

benchmark) results, as defined by the most important IVF scientific societies for the main key performance indicators, can additional tools and strategies be implemented to attain a live birth in the shortest possible time and with the lowest possible risks. However, any add-on involves another concern: cost-effectiveness. Is it worth investing a given amount of money to provide a protocol that differs from standard care? This financial perspective further complicates the definition of success in IVF, especially because a generally valid characterization of the financial burden is questionable.

Finally, there is the couple's perspective, which is often disregarded. Each of these perspectives is addressed in this commentary.

CLINICAL PERSPECTIVE

Wilkinson and colleagues reviewed 142 randomized controlled trials and stated that 'there is massive diversity in numerator and denominator selection in IVF trials due to its multistage nature' (Wilkinson *et al.*, 2016). They reported over 800 combinations and claimed that 'this causes methodological frailty in the evidence base'. Large variability was also shown among IVF clinics' self-reported outcomes across websites.

In general, only a minority of centres assess the cumulative live birth rate (CLBR) per ITT, which clearly stands out as the most valuable measure (Maheshwari *et al.*, 2015). In fact, this outcome entails a comprehensive view of

KEY WORDS

Cost-effectiveness
Cumulative live birth rate
IVF success
Patient-centeredness
Time to pregnancy

¹ Clinica Valle Giulia, GeneraLife IVF, Rome, Italy

² Livet, GeneraLife IVF, Turin, Italy

³ Carl von Linné Clinic, GeneraLife IVF, Uppsala, Sweden

⁴ Demetra, GeneraLife IVF, Florence, Italy

⁵ Ginefiv, GeneraLife IVF, Barcelona, Spain

⁶ FertiCare, GeneraLife IVF, Prague, Czech Republic

a treatment from the chosen stimulation regimen, passing through embryo manipulations, down to the sum of fresh and cryopreserved embryo transfers, and defines its efficacy. Notably, provided that the ideal number of oocytes is collected from each patient, this outcome cannot be improved; it is mostly dependent on the competence of the resulting embryos and the receptivity of the endometrium. Therefore, it is necessary for IVF practitioners to not impair such competence, through validated technologies and a safe environment *in vitro*.

Maheshwari and colleagues outlined a short-term (first live birth in a 2-year period), medium-term (all live births in a 5-year period from a single egg collection) and long-term (all live births in a 10-year period from up to three egg collections) view in defining the CLBR (Maheshwari *et al.*, 2015). The International Glossary on Infertility and Fertility Care (Zegers-Hochschild *et al.*, 2017a; Zegers-Hochschild *et al.*, 2017b) further refined the definition as the cumulative live birth delivery rate (CLBdR) per ITT: 'The number of deliveries with at least one live birth resulting from one initiated or aspirated ART cycle, including all cycles in which fresh and/or frozen embryos are transferred, until one delivery with a live birth occurs or until all embryos are used, whichever occurs first'. Also known as the 'take-home baby rate', this outcome has been credited as 'the single most relevant and consumer-friendly measure of success' as it also includes in the denominator patients who do not respond to stimulation or do not obtain embryos, thus providing an objective estimate of a couple's chance of success (Abdalla *et al.*, 2010; Chetkowski, 2014). Clearly, it is strictly dependent on the population of patients treated and may vary highly across clinics. In fact, the fulfilment of an expected CLBdR per ITT does not define a positive outcome. Conversely, whenever it is lower than expected, this raises concerns. The main values that must be met are outlined by both national and international registers and by scientific societies in the field of reproductive medicine.

In summary, the primary outcome measure in IVF is a parameter that cannot be improved, but only reduced in cases of poor practice. The same concept does not apply to efficiency,

which encompasses the risks, time, costs and efforts invested to achieve a given CLBdR per ITT. Therefore, any time a new policy, protocol or tool is introduced in clinical practice, it should be ensured that the IVF efficacy is at the very least preserved.

FINANCIAL PERSPECTIVE

The European Society for Human Reproduction and Embryology (ESHRE) in 2015 underlined that a cost-effectiveness analysis in IVF is highly dependent on socioeconomic background and clinical setting, which strongly limits the generalizability of the obtained evidence (ESHRE, 2015). Previously, the ESHRE also released a position paper ('Good Clinical Treatment in Assisted Reproduction') highlighting two concepts: (i) the European Parliament has stated that the Member States should ensure 'the right of couples to universal access to infertility treatment', and (ii) 'it is important that there is coherence between a country's decision to support assisted reproduction and the financial and regulatory strategies affecting the quality of the service provided'. Therefore, a list of critical prerequisites that must be ensured besides cost-effectiveness was outlined: accessibility and short waiting time, appropriate investigation of male/female infertility, opportune counselling, careful choice of ovarian stimulation and treatment strategies, adoption of single-embryo transfer whenever possible, and long-term monitoring of outcomes.

Nevertheless, infertility is a societal issue and must be framed against the legal, social and economic backgrounds of each country. Therefore, each centre should conduct a proper investigation to fit its socioeconomic environment and specific expenses. For instance, Somigliana and co-workers recently investigated whether blastocyst stage preimplantation genetic-testing for aneuploidies (PGT-A) would be cost-effective in the setting of an Italian public IVF centre. However, in addition to outlining the mathematical model and its theoretical conditions, they also included an Excel file as supplementary material so that anyone could perform their own internal analysis (Somigliana *et al.*, 2019). In another study, the issue of cost-effectiveness of PGT-A was investigated by applying the expenses of different countries worldwide (from the

highest costs of American centres to the lowest cost scenarios of Chinese, Korean and Iranian centres) but in the context of polar body biopsy and array comparative genomic hybridization analysis (Neumann *et al.*, 2020). These studies, aimed at answering the same study question, were conducted using distinct strategies and resulted in different conclusions, thereby highlighting the complexity of performing a widely acceptable cost-effectiveness analysis in IVF.

More examples apply to other controversies, such as the choice between a single- or double-embryo transfer strategy, or between a freeze-all rather than fresh embryo transfer approach. In summary, although statistically sound and well conducted, all studies will always be limited by the application of national/regional costs and by the clinical strategies allowed or chosen in the context where they are performed. All conclusions should therefore be always confirmed internally.

REDUCTION OF IVF-RELATED RISKS

The European Observatory on Health Systems and Policies underlined in 2017 how critical it is to define the metrics to assess 'clinical efficiency' in health, whose scope, comparability among clinics, timeliness, quality and actual usefulness should be carefully evaluated (Cylus *et al.*, 2017). In this regard, the current authors think that 'clinical efficiency' in reproductive medicine and IVF should also encompass patient-centred factors such as time, risks and psycho-physical distress, all of which affect treatment discontinuation.

Maheshwari and colleagues recognized that the CLBR per ITT entails an intrinsic limitation: it does not account for the burden of treatment (Maheshwari *et al.*, 2015). Indeed, any negative experience should be clearly acknowledged when defining success in IVF. The minimization of IVF-related risks, such as ovarian hyperstimulation syndrome (OHSS), miscarriage and multiple pregnancies, also defines the quality of treatment. In particular, the implementation of technologies to minimize their prevalence, although involving higher costs in the first place, can spare expenditure in the medium or long term, and this must be carefully assessed. In addition, OHSS, miscarriage and multiple

pregnancies are also physically and psychologically demanding. A woman who delivers two newborns through a multiple gestation or two singleton gestations, just like a woman who delivers after a previous miscarriage or none, will contribute equally to the CLBdR per ITT, but through totally different experiences.

Health cannot be treated as a single concept, as well-being is not defined simply by the absence of illness (infertility in this field), but also by the pleasant or unpleasant moods and emotions of patients undergoing treatment. The World Health Organization (WHO) supported this view by including psychological well-being in the concept of health. In summary, the risks of any procedure should always be minimized, but the patient must be clearly informed of their prevalence.

HONEST COUNSELLING, CARE AND SUPPORT

The effectiveness of any treatment should be complemented by its safety, timeliness, efficiency, equity and patient-centredness. Dealing with the latter, healthcare providers should

be responsive to individual patient preferences, needs and values, and ensure they guide all clinical decisions. Seven domains have been defined: access to care, patient engagement, information systems, care coordination, integrated and comprehensive team care, patient-centred care survey, and publicly available information. Direct engagement of patients in the decision-making process is key. As infertility is a matter of quality of life, each treatment should fulfil patients' expectations with the available strategies to meet their chances of success. A positive experience should be provided with clear counselling and informed consent.

Treatment discontinuation between failed attempts is an important issue affecting IVF success (*Gameiro et al., 2012*). The reasons for discontinuation might be mainly financial. However, many couples discontinue treatment in public settings as well, where their expenses are contained and several cycles might be free of charge. Therefore, their physical and psychological states represent a burden they might not bear so easily, especially in cases of poorer prognosis (advanced maternal age and/or longer

duration of infertility). In this regard, the American Society for Reproductive Medicine (ASRM) underlined the importance of patient perspective in the decision process whenever safe non-futile treatments can be offered (*ASRM, 2019*). However, defining a treatment as 'non-futile' is challenging, as these patients often have unrealistic expectations, even after receiving clear verbal or numerical information from clinicians. Clinicians should be prepared to address this possibility and properly manage these cases.

CONFOUNDERS

The definition of success in IVF should encompass all the perspectives previously described. However, these outcomes are misleading if not presented along with confounders, such as maternal and paternal age, duration and cause of infertility, sperm factor, body mass index, ovarian reserve markers and reproductive history. A recent opinion piece by Correia and collaborators (*Correia et al., 2020*) elegantly described this issue by stating that an incorrect adoption of confounders might lead to overestimation, underestimation,

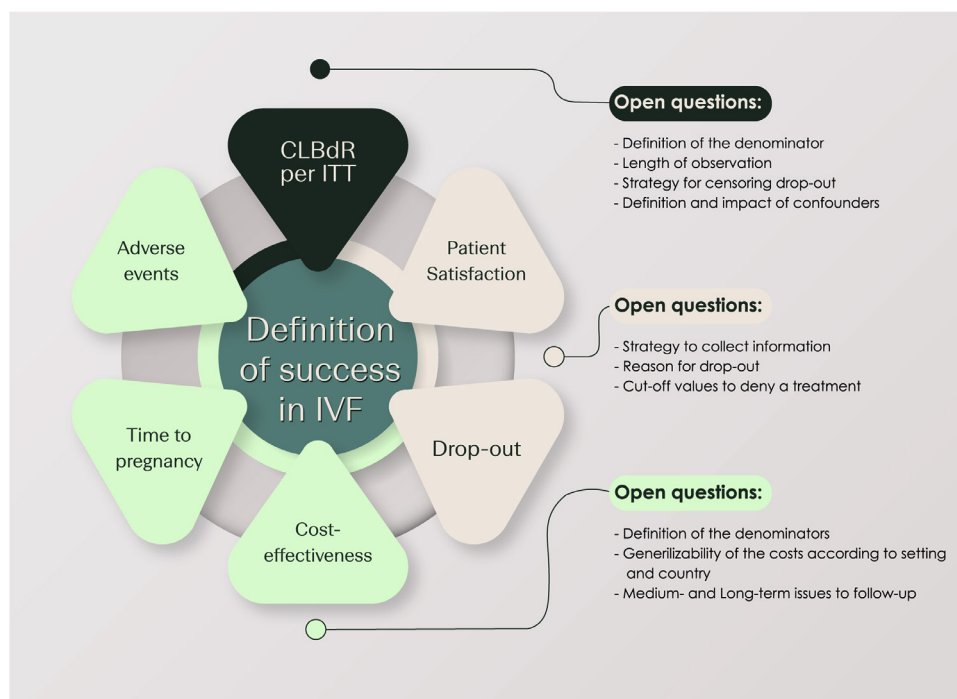


FIGURE 1 How do we define success in IVF? All outcomes must be evaluated according to the patient-related characteristics significantly associated with them. If dissociated from the population of patients they refer to and not displayed across relevant categories (e.g. maternal age and/or ovarian reserve ranges), the outcomes might be misleading. The national and international average outcomes reported in the IVF registers and the standards outlined by the scientific societies represent the benchmark values that should be met as a minimum by each centre. Ideally, the current measures should be integrated in a common scheme, but several questions remain open to reach a consensus. CLBdR, cumulative live birth delivery rate; ITT, intention to treat.

masking or reversal of the effect of a clinical strategy or tool, and suggested avoiding a 'kitchen sink' approach in adjusting IVF data. They recommended presenting the data among populations of patients clustered according to the most significant features affecting IVF outcomes (e.g. range of maternal age), and we agree with their view.

CONCLUSION

All the macro-areas that characterize IVF should converge into a multidisciplinary network to provide patients with safe and efficient treatments. Such treatments must meet their predetermined chance of success and expectations involving the lowest possible risks, costs, time and psycho-physical burdens. Avoiding futile or unproven adds-on while maintaining sufficiently low treatment discontinuation rates is a key challenge. National IVF registries can accelerate this process by monitoring variables such as CLBdR (even across multiple cycles), as well as patient satisfaction (e.g. the Swedish registry).

FIGURE 1 summarizes the main outcomes and the points still open for discussion. We think it is time for the main scientific societies in reproductive medicine to join forces and gather experts to reach a consensus on the definition of success in IVF, as done by the ICMART and WHO when drafting the Glossary of Infertility and Fertility Care (Zegers-Hochschild *et al.*, 2017a; Zegers-Hochschild *et al.*, 2017b). Such a definition should be grounded on the thesis of Abdalla and colleagues that 'an ideal outcome measure in IVF ... incorporates clinical effectiveness, acceptability and safety. In addition, it should be reliable, valid, and easy to generate and understand' (Abdalla *et al.*, 2010). Conceiving a single measure is utopic, but a scheme that integrates the main outcomes described in this commentary should be our goal.

REFERENCES

- Abdalla, H.I., Bhattacharya, S., Khalaf, Y. **Is meaningful reporting of national IVF outcome data possible?** Hum. Reprod. 2010; 25: 9–13
- ASRM. **Fertility treatment when the prognosis is very poor or futile: an Ethics Committee opinion.** Fertil. Steril. 2019; 111: 659–663
- Chetkowski, R.J. **Consumer-friendly reporting of in vitro fertilization outcomes.** Fertil. Steril. 2014; 101: e7
- Correia, K.F., Dodge, L.E., Farland, L.V., Hacker, M.R., Ginsburg, E., Whitcomb, B.W., Wise, L.A., Missmer, S.A. **Confounding and effect measure modification in reproductive medicine research.** Hum. Reprod. 2020; 35: 1013–1018
- Cylus, J., Papanicolas, I., Smith, P.C. 2017 **How to make sense of health system efficiency comparisons?** World Health Organization, Regional Office for Europe
- ESHRE. **Economic aspects of infertility care: a challenge for researchers and clinicians.** Hum. Reprod. 2015; 30: 2243–2248
- Gameiro, S., Boivin, J., Peronace, L., Verhaak, C.M. **Why do patients discontinue fertility treatment? A systematic review of reasons and predictors of discontinuation in fertility treatment.** Hum. Reprod. Update 2012; 18: 652–669
- Maheshwari, A., McLernon, D., Bhattacharya, S. **Cumulative live birth rate: time for a consensus?** Hum. Reprod. 2015; 30: 2703–2707
- Neumann, K., Sermon, K., Bossuyt, P., Goossens, V., Geraedts, J., Traeger-Synodinos, J., Parriego, M., Schmutzler, A., Van Der Ven, K., Rudolph-Rothfeld, W., Vonthein, R., Griesinger, G. **An economic analysis of preimplantation genetic testing for aneuploidy by polar body biopsy in advanced maternal age.** BJOG 2020; 127: 710–718
- Somigliana, E., Busnelli, A., Paffoni, A., Viganò, P., Riccaboni, A., Rubio, C., Capalbo, A. **Cost-effectiveness of preimplantation genetic testing for aneuploidies.** Fertil. Steril. 2019; 111: 1169–1176
- Wilkinson, J., Roberts, S.A., Showell, M., Brison, D.R., Vail, A. **No common denominator: a review of outcome measures in IVF RCTs.** Hum. Reprod. 2016; 31: 2714–2722
- Zegers-Hochschild, F., Adamson, G.D., Dyer, S., Racowsky, C., De Mouzon, J., Sokol, R., Rienzi, L., Sunde, A., Schmidt, L., Cooke, I.D., Simpson, J.L., Van Der Poel, S. **The International Glossary on Infertility and Fertility Care, 2017.** Hum. Reprod. 2017; 32: 1786–1801
- Zegers-Hochschild, F., Adamson, G.D., Dyer, S., Racowsky, C., De Mouzon, J., Sokol, R., Rienzi, L., Sunde, A., Schmidt, L., Cooke, I.D., Simpson, J.L., Van Der Poel, S. **The**

International Glossary on Infertility and Fertility Care, 2017. Fertil. Steril. 2017; 108: 393–406

Received 29 April 2021; received in revised form 11 August 2021; accepted 12 August 2021.