Trying to define the optimal progesterone elevation cut-off in fresh in vitro fertilization cycles: time to evolve our way of thinking

The effect of elevated progesterone on the day of human chorionic gonadotropin (hCG) administration on the probability of pregnancy after a fresh embryo transfer has been a matter of debate for more than 25 years. Available evidence has now convincingly demonstrated that a negative effect is indeed present (1), and it can be detected using appropriate analytical approaches (2). This effect seems to be exerted on the endometrium and this hypothesis is supported by both basic and clinical research but most importantly by data suggesting that the pregnancy rate of embryos obtained from cycles with elevated progesterone is not affected when these embryos are transferred in subsequent frozen-thawed embryo transfer cycles (1). This has led many clinicians and researchers to propose the freeze-all (or freeze-only) policy as an effective way of managing elevated progesterone at the end of the follicular phase.

One of the most controversial aspects of the progesterone elevation literature has been the selection (arbitrary or not) of the threshold above which the value of progesterone would be considered abnormal and an intervention would be reasonable. This has led to the introduction of multiple such thresholds which has been a source of confusion. Not surprisingly, currently, the burning question for any clinician measuring progesterone on the day of hCG is: what is the threshold that should be used to dictate the cancellation of a fresh transfer and the adoption of a freeze-only policy?

This is the question that Hill et al. (3) are trying to answer in their elaborate analysis published in this issue of *Fertility and Sterility*. They aspire to do so by exploring several different methodologies while taking into account the sensitivity and specificity of each proposed threshold and also the potential cost penalty that the application of a freeze-only strategy using this threshold would incur. Their analysis suggests that adopting a threshold between 1.5 ng/mL to 2.0 ng/mL seems to be the most cost-effective approach, although they do acknowledge that a statistically significant negative effect of progesterone on pregnancy rates is detected much earlier, using a threshold around 0.7 ng/mL.

The analysis of Hill et al. (3) is thought provoking and definitely useful for the readership of Fertility and Sterility since it describes the plethora of different variables one needs to consider when making a decision to change the criteria of a freeze-only policy in the clinic. The authors do acknowledge the limitations of their analysis, which cannot be stressed enough. Firstly, the potential effectiveness of the freeze-only strategy in these patients has never been proven through randomized controlled trials. Hence, the expected benefit of such a strategy, if any, has not yet been properly quantified and for the purpose of this analysis is indirectly calculated based on retrospective observational data. In addition, their analysis does not seem to account for significant confounders such as female age, numbers of oocytes retrieved, number, developmental stage and quality of embryos transferred. These have been shown to have significant impact on the association of elevated progesterone with live birth rates and seem to modify the observed effect size (2). Furthermore, their analysis has not incorporated any measure of uncertainty (such as 95% confidence intervals) and hence its output does not describe the range of potential outcomes that could be observed in the population. Variables such as the timing of progesterone measurement and the progesterone assay performance, the population mix, the efficiency of cryopreservation protocols and embryo transfer policies, all of which are specific to each clinic, can significantly modify the effect of elevated progesterone on pregnancy rates and alter the suitability of the thresholds. These issues are appropriately discussed by Hill and colleagues (3) and should be taken into account by the reader when interpreting the findings of their study.

The most significant problem, though, lies in the very essence of the research question that the authors try to answer, i.e. identifying the optimal threshold for elevated progesterone. Obviously defining something as abnormal has been in the core of medical thinking for thousands of years and is an approach that seems to facilitate clinical decision making. The tendency, though, to dichotomise the population in normal and abnormal based on a single threshold value of a continuous variable is, in most instances, highly problematic and leads to counter-intuitive clinical scenarios. Essentially, in the case of elevated progesterone, when one chooses a threshold, the population is divided in two distinct subgroups, with the members of each subgroup having the same probability of pregnancy (the average of the specific subgroup). Hence, if we select the highly popular 1.5 ng/mL threshold, then a patient with a serum progesterone of 1.45 ng/mL is unaffected by "elevated progesterone" whereas a patient with a serum progesterone of 1.52 ng/mL is considered to be fully affected (the "all or nothing" hypothesis). The affected patient is now expected to have a decrease of almost 10% in the probability of pregnancy as compared to the unaffected patient when (1) the difference in the progesterone concentration between these two patients is just 0.07 ng/mL which is well within the margin of error of most progesterone assays. Even more surprising is the fact that in this scenario, we assume that the difference in pregnancy rates between these two patients is the same as in the scenario where a patient has a progesterone value of 0.7 ng/mL and is compared with a patient with a progesterone value of 3.0 ng/mL.

Notwithstanding the obvious clinical irrationality of such an approach and the misclassification issues that it can generate, dichotomization also results in reduced statistical power which, when combined with inadequate adjustment for confounders, frequently leads to type II errors. Epidemiologists and biostatisticians have identified

this as an important issue in biomedical research and have named it "dichotomania" (4). The obvious answer to this problem is to recalibrate our thinking away from artificial dichotomies. Modern statistical techniques allow for a much better utilization of continuous data and can provide a significantly clearer picture of the association under investigation.

For these reasons, we propose that future clinical research in this field should move away from the "elevated progesterone" concept to the far more clinically and statistically sensible introduction of serum progesterone on the day of hCG as a continuous covariate in prediction models of live birth after fresh embryo transfer. This is of paramount importance since it is highly likely that serum progesterone concentration on the day of hCG represents the best currently available proxy variable of endometrial receptivity after ovarian stimulation.

So far, mostly due to our inefficient methodological approaches, we have failed to fully utilize the prognostic information that the measurement of progesterone can provide to optimize the outcomes of assisted reproductive technologies for our patients. The study by Hill et al. (3), represents an important contribution to the literature because it attempts to take the discussion around elevated progesterone to the next level. However, it also reveals the limitations of the current approach and the urgent need to evolve our way of thinking from trying to determine the optimal threshold of elevated progesterone to embracing the actual prognostic value of progesterone concentration on the day of hCG.

Christos A. Venetis, M.D., Ph.D.[a,b]
Basil C. Tarlatzis, M.D., Ph.D.[c]
[a] Centre for Big Data Research in Health and [b] School of Women's and Children's Health, University of New South Wales Medicine, Sydney, New South Wales, Australia; and [c] School of Medicine, Aristole University of Thessaloniki, Thessaloniki, Greece

https://doi.org/10.1016/j.fertnstert.2018.06.006

REFERENCES

1. Venetis CA, Kolibianakis EM, Bosdou JK, Tarlatzis BC. Progesterone elevation and probability of pregnancy after IVF: a systematic review and meta-analysis of over 60 000 cycles. Hum Reprod Update 2013;19:433–57.
2. Venetis CA, Kolibianakis EM, Bosdou JK, Lainas GT, Sfontouris IA, Tarlatzis BC, et al. Estimating the net effect of progesterone elevation on the day of hCG on live birth rates after IVF: a cohort analysis of 3296 IVF cycles. Hum Reprod 2015;30:684–91.
3. Hill MJ, Healy MW, Richter KS, Parikh T, Devine K, DeCherney AH, et al. Defining thresholds for abnormal premature progesterone levels during ovarian stimulation for assisted reproduction technologies. Fertil Steril 2018;110:671–9.e2.
4. Senn S. Dichotomania: an obsessive compulsive disorder that is badly affecting the quality of analysis of pharmaceutical trials. Proceedings of the International Statistical Institute, 55th Session, Sydney. 2005. The Hague, The Netherlands: ISI. Available at: http://isi.cbs.nl/iamamember/CD6-Sydney2005/ISI_Final_Proceedings.htm.