The suggestion of excluding the articles of Khorram et al. (2006) and Sahin et al. (2004) is not correct. Khorram et al. randomized patients between metformin plus clomiphene and clomiphene alone. Metformin was started in the same cycle as was clomiphene. Sahin et al. did treat the patients with metformin 3 months prior to administrating clomiphene, but all patients also received clomiphene, which also makes the comparison justifiable. Both these studies together counted for 52 of 985 patients (5%) in this particular comparison.

Obviously, there is no argument against more data as the estimation of the truth will thereby improve.

References


Sahin Y, Yirmibes U, Kelestimur F, Aygen E. (2004) The effects of metformin clomiphene alone. Metformin was started in the same cycle as was clomiphene, which also makes the comparison justifiable. References received clomiphene, which also makes the comparison justifiable. Both these studies together counted for 52 of 985 patients (5%) in this particular comparison.

Sir,

I have read with great interest the systematic review and meta-analysis published in your journal by Venetis et al. (2007) in which the association between serum progesterone (P) elevation on the day of human chorionic gonadotrophin (hCG) administration with the probability of pregnancy in in vitro fertilization (IVF) is evaluated in depth.

First of all, I would like to complement the tremendous job the authors have done, with an initial screening of 1114 studies, and the final inclusion of 12 studies for the systematic review and of five studies for the meta-analysis, all of which meet appropriate strict and rigorous criteria for properly analysing such association. Nevertheless, there are some interesting data that merit further evaluation and comments.

In regard to the main and final conclusion of the study, the authors conclude that the available data do not support an association between P elevation on the day of hCG and the probability of clinical pregnancy. However, data obtained from the five selected studies for meta-analysis, which include a total of 700 cycles, show an OR = 0.75 (95% CI: 0.53–1.06), (P = 0.10), and a mean difference of −0.10 (95% CI: −0.22–0.0) (P = 0.11) between patients with P elevation versus those without it. Despite the fact that the different studies chose different serum P threshold levels to define premature luteinization, and that the variation coefficients of the hormonal assays obviously differ among the studies, these data reflect a trend to a negative association between P elevation and pregnancy rate. Although differences do not reach statistical significance, they are of a clear clinical relevance. Moreover, if analysis of crude data from the two GnRH antagonist cycles is performed (Ubaldi et al., 1996; Bosch et al., 2003), a statistically significant difference between patients with and without serum P elevation is observed (OR = 0.41; 95% CI: 0.17–0.97); P = 0.04, even considering the different P threshold level employed in each of the studies.

In this context, I completely agree with the authors comment that the use of the term ‘premature luteinization’ might not be appropriate, as serum P elevation is occurring despite the use of GnRH analogues, and therefore, under normal serum LH concentrations. Luteinization is the process, by which the mature ovarian follicle transforms into a corpus luteum, and it is mainly induced by LH (Murphy, 2000); thus, these two events are not related.

On the other hand, defining a single threshold for a detrimental serum P level seems imprecise: most of the studies failing to demonstrate an association employed a threshold level of 0.9 ng/ml. Probably, a trend analysis relating both variables would be of greater interest.

The second issue that needs to be clarified is the analysis of the factors related to serum P elevation. The authors performed a single variable analysis to study the different relationships of FSH requirements, duration of FSH stimulation, serum estradiol (E2) levels on the day of hCG administration and the number of oocytes retrieved with the presence of P elevation. Only E2 levels showed a significant relationship with P elevation. This simple approach provides a preliminary conclusion, but cannot be definitive. The most correct way of analysing the influence of diverse risk factors on a particular condition is a multivariate analysis, as co-linearity among the different variables could be present, and any of them might be a confounding variable, or could act by modifying the effect of one or the rest of them. In our study (Bosch et al., 2003), we performed a multivariate analysis with logistic regression in which all potentially related variables were included. As a result, the total dose of FSH employed and the serum E2 levels on the day of hCG showed a statistically significant relationship with the occurrence of P elevation. This finding is in consistency with others’ in which serum P levels correlated positively with the dose of FSH administered (Filicori et al., 2002) and with circulating FSH concentrations (Adonakis et al., 1998). Altogether, these data suggest that increased serum P is related to high circulating FSH concentrations that provide the development of more follicles augmenting total granulose cells’ activity. As a consequence, P


Etelka Moll1 and Fulcovan der Veen

Department of Obstetrics and Gynaecology,
Centre for Reproductive Medicine, Academic Medical Center,
University of Amsterdam,
PO Box 22700, 1100 DE Amsterdam,
The Netherlands

1Correspondence address. E-mail: e.moll@amc.uva.nl
doi:10.1093/humupd/dmm050

194
elevation appears associated to high serum E2 levels on the day of hCG.

In conclusion, the available evidence suggests that the elevation of circulating P concentrations at the end of the follicular phase has a negative impact on GnRH antagonist cycles. On the other hand, this negative effect remains to be demonstrated in GnRH agonist cycles; the evaluation of larger series with an appropriate (higher) threshold P value may lead to a similar outcome. The total FSH dose employed for ovarian stimulation seems to be the factor most related to circulating P, through the increase of total granulosa cells’ activity.

References

Ernesto Bosch*
Instituto Valenciano de Infertilidad,
Plaza Policia Local, 3, 46015,
Valencia, Spain
*Corresponding address. Tel: +34-96-305-09-00; Fax: +34-96-305-09-99; E-mail: ebosch@ivi.es doió:10.1093/humupd/dmm046


Sir,

We would like to thank Dr Bosch for his interest in our work (Venetis et al., 2007). Dr Bosch raises certain issues regarding the interpretation of the results of this systematic review and meta-analysis, which deserve commenting.

(i) The concept of trend

In the systematic review and meta-analysis by Venetis et al. (2007), a clear research question was asked: is progesterone elevation on the day of hCG administration associated with the probability of pregnancy, in women undergoing ovarian stimulation using gonadotrophins and GnRH analogues for IVF? The results of the quantitative data synthesis, which was performed using a methodology described in detail in the published manuscript, failed to reject the null hypothesis at an alpha error level of 0.05. The conclusion derived from these data, as stated in the manuscript, was that ‘the best available evidence does not support an association between progesterone elevation on the day of hCG administration and the probability of clinical pregnancy in women undergoing ovarian stimulation with GnRH analogues and gonadotrophs for IVF’. The concept of ‘trend’ to which Dr Bosch refers to, is difficult to define and is not compatible with hypothesis testing and statistical inference. Therefore, it cannot be the result of a meta-analysis, in which the null hypothesis is either accepted or rejected.

(ii) Clear clinical relevance of a non-statistically significant result

Failing to reject the null hypothesis might be attributed to lack of power, however, it could also reflect the absence of a true difference.

Thus, it is not clear how differences that do not reach statistical significance can be of ‘clear’ clinical relevance, as Dr Bosch suggests in his letter.

(iii) Analysis of ‘crude data’

Such an approach, known as `treat-as-one-trial’, is considered statistically inappropriate, since it disregards study-to-study variation and it is prone to biased results and the well-known issue of Simpson’s Paradox (Bravata and Olkin, 2001; Altman and Deeks, 2002). For these reasons, combining the results from individual studies, where feasible, is performed with the use of proper meta-analytic models (Deeks et al., 2001). The application of such models in the current meta-analysis failed to support the presence of statistically significant differences in terms of pregnancy rates in patients with and those without progesterone elevation.

(iv) Association between effect size and threshold level of progesterone used to classify patients as those with and those without progesterone elevation

Analyzing the potential association between progesterone thresholds and effect size, might be of interest. However, it would require a much larger number of studies, in which various progesterone thresholds should have been used. This task was not feasible in the present systematic review and meta-analysis due to the limited number of eligible studies.

(v) Analysis of secondary outcomes

Prediction of the occurrence of progesterone elevation on a certain patient population might be performed by using multivariate logistic regression analysis, as Dr Bosch suggests. However, this was not the aim of the current systematic review and meta-analysis, and even in the context of a secondary, exploratory analysis, it would not have been feasible, due to the limited number of studies.

(vi) In his letter Dr Bosch states “In conclusion, the available evidence suggests that the elevation of circulating P concentrations at the end of the follicular phase has a negative impact on GnRH antagonist cycles.”

We do not share Dr Bosch’s certainty regarding the negative effect of elevated progesterone on the day of hCG on pregnancy rates in GnRH antagonist cycles. Clearly, based on the results of