COMMENT

Comment on "Accuracy of the FMF Bayes theorem-based model for predicting preeclampsia at 11–13 weeks of gestation in a Japanese population"

Antonio Farina[1] · Paolo Ivo Cavoretto[2]

Preeclampsia (PE) is a major cause of maternal and perinatal morbidity or mortality, and early onset disease requiring iatrogenic preterm delivery is associated with even higher risks of maternal and neonatal complications. Identification of pregnancies at high risk of subsequent PE development is beneficial so that the focus on therapy and prevention (including prophylactic use of low-dose aspirin, closer surveillance and earlier delivery) can be more appropriate. Therefore, in recent years, many efforts have been made to screen for the disease before onset to minimize its impact on pregnancy outcome [1].

In their paper, Goto et al. performed an external validation of the Fetal Medicine Foundation (FMF) algorithm for PE risk calculation in a Japanese population.

Many papers have previously validated the FMF algorithm. The statistical analysis is based on a very complicated and excellent competing risks model that is based on a survival-time analysis for the gestational age at delivery with PE. This approach assumes that if the pregnancies were to continue indefinitely, all women would experience PE, and whether they do so before a specified gestational age depends on "competition" between delivery before or after the development of PE. In the original papers by the FMF in London [2–4], the competing risks model is visually expressed with a Gaussian distribution having the mean gestational week at delivery with PE on the X-axis. Consistent with the incidence of the disease, only 3.6% of

low-risk patients will develop PE within 42 weeks, and in this case, the Gaussian distribution is completely shifted to the right [4]. In the presence of specific maternal risk factors, the whole Gaussian distribution (mean and relative standard deviation (SD)) shifts to the left, thus anticipating PE onset. As a result of the analysis, 50% of the low-risk and high-risk patients developed PE at 54.4 weeks and 44.0 weeks, respectively [4]. An extra and more accurate shift to the left for PE prediction purposes is given by the aberrant values of biochemical and biophysical markers.

In Goto's paper, the prediction model essentially revealed a good degree of reproducibility in terms of detection rate (DR) compared with the original results already shown by other local and independent studies and obtained mainly in European populations (Caucasian in the vast majority of studies). In particular, the DR for preterm PE at a fixed 10% false positive rate (FPR) was 91% when all the available markers (maternal factors, mean arterial pressure (MAP), uterine artery pulsatility index (UtA-PI) and placental growth factor (PlGF)) were included in the model. On the other hand, the prediction for term PE was poorer with 60% DR at a fixed FPR of 10%. The calibration curve for early PE (expected vs. observed risk) was basically in line with the original estimation [2, 3]. In fact, in a calibration plot, the very high expected and observed risks (1:1–1:10) were quite similar, but a less correct calibration (with underestimation of the validation model) was obtained for the observed value. In fact, for an estimation of 1:600, the corresponding expected risk value was ~1:90. It is unclear whether this abnormal calibration could affect DR and/or FPR.

In contrast, a previous multicenter study by Chaemsaithong et al. [5] on a series in which the vast majority of subjects were East Asian women (94% of cases from China, Japan, Indonesia) found a lower DR than that reported in the study covered by this commentary. In Chaemsaithong's paper, the DR was only 64% for the same FPR of 10% [5]. The reason for this difference is not clear, but from the visual

✉ Antonio Farina
antonio.farina@unibo.it

1   Division of Obstetrics and Gynecology, St. Orsola-Malpighi Polyclinic, Department of Medical and Surgical Sciences (DIMEC), University of Bologna, Bologna, Italy

2   Gynecology and Obstetrics Department, IRCCS San Raffaele Hospital, Vita-Salute University, Milan, Italy

inspection of the box plots relative to the biochemical and biophysical markers, in Goto's paper (considering exclusively the Japanese population), the mean multiple of median (MoM) values for PlGF and UtA-PI in early PE are higher than those described in Chaemsaithong's paper. Therefore, the values of these markers in non-Japanese East Asian populations may present smaller deviations from their controls.

It is also very interesting, in our opinion, that all the biophysical and biochemical markers present a wider SD in Chaemsaithong's paper [5] than in Goto's paper. This is consistent with the presence of a wider random effect at different levels (population and individual effect) of unknown degree able to affect the MoM distributions, especially for the early PE cases. In other words, this could mean that early PE has slightly different phenotypes according to individual characteristics that are not intercepted by the statistical model, resulting in a lower DR. Multiple random effects probably affect more multicenter rather than monocentric studies. This makes sense, since the etiopathogenesis of PE is still partially unknown, many factors could affect the marker profile of each patient. Given this speculation, we believe that future studies aimed at improving DR should include the calculation of random effects rather than including new extra markers.

Again, in Goto's paper, the UtA-PI MoM was lower in late PE than in both early PE and unaffected cases. We believe that this unexpected result could be partially related to the high rate of in vitro fertilization (IVF) patients in the groups with PE. In fact, recent major evidence of lower UtA-PI values was observed in in vitro fertilization/intracytoplasmic sperm injection (IVF/ICSI) pregnancies from frozen blastocyst transfer and oocyte donations [6, 7]. Lower UtA-PI is probably the result of a compensatory phenomenon of greater uterine perfusion aimed at balancing dysfunctions both in the mother or in the placenta [6, 7], so the PE phenotypes in these study groups may be different from those in the general population. Adjustment of the UtA-PI values in IVF pregnancies may result in an improvement of the prediction, at least for late PE (where the UtA-PI values are less abnormal than those in early PE), and introduce a notably new PE phenotype. Furthermore, testing the effectiveness of different interventions and treatments based on the different combinations of available biomarkers for different PE phenotypes may lead to patient-specific interventions for each clinical case [8].

Finally, it is the opinion of the writers of this commentary that the validation papers using the FMF algorithm should report, as extra information, the measures of the central tendency and dispersion of the Gaussian distributions

(low and high risk) of their own populations, and how they differ from the original proposed. Knowledge of SD and mean gestational age at delivery with PE could reveal hidden sources of discrepancy, allowing possible improvement of the model performance and new potential custom adjustment for each specific population.

We conclude that the competing risks model proposed by the FMF achieved the best performances in terms of prediction and reproducibility in external studies. Therefore, this approach is eligible for use worldwide soon as the main screening tool in the first trimester of pregnancy.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Rolnik DL, Wright D, Poon LC, O'Gorman N, Syngelaki A, de Paco Matallana C, et al. Aspirin versus Placebo in Pregnancies at high risk for preterm preeclampsia. N Engl J Med. 2017;377: 613–22.
2. O'Gorman N, Wright D, Syngelaki A, Akolekar R, Wright A, Poon LC, et al. Competing risks model in screening for preeclampsia by maternal factors and biomarkers at 11-13 weeks gestation. Am J Obstet Gynecol. 2016;214:103.e1–e12.
3. Wright D, Tan MY, O'Gorman N, Poon LC, Syngelaki A, Wright A, et al. Predictive performance of the competing risk model in screening for preeclampsia. Am J Obstet Gynecol. 2019;220:199. e1–e13.
4. Wright D, Wright A, Nicolaides KH. The competing risk approach for prediction of preeclampsia. Am J Obstet Gynecol. 2020;223:12-23.e7.
5. Chaemsaithong P, Pooh RK, Zheng M, Ma R, Chaiyasit N, Tokunaka M, et al. Prospective evaluation of screening performance of first-trimester prediction models for preterm preeclampsia in an Asian population. Am J Obstet Gynecol. 2019;221:650. e1–e16.
6. Cavoretto PI, Farina A, Gaeta G, Sigismondi C, Spinillo S, Casiero D, et al. Uterine artery Doppler in singleton pregnancies conceived after in-vitro fertilization or intracytoplasmic sperm injection with fresh vs frozen blastocyst transfer: longitudinal cohort study. Ultrasound Obstet Gynecol. 2020;56:603–10. https://doi.org/10.1002/uog.21969.
7. Cavoretto P, Farina A, Miglio R, Zamagni G, Girardelli S, Vanni VS, et al. Prospective longitudinal cohort study of uterine arteries Doppler in singleton pregnancies obtained by IVF/ICSI with oocyte donation or natural conception. Hum Reprod. 2020;35: 2428–38.
8. Lees C, Ferrazzi E. Relevance of haemodynamics in treating preeclampsia. Curr Hypertens Rep. 2017;19:76. https://doi.org/10.1007/s11906-017-0766-6.