Abstract Details

Session title: Endometriosis, endometrium and fallopian tube, and benign disorders of the endometrium and fallopian tube

Session type: Poster viewing

Presentation number: P-298

Abstract title:

Biomarker discovery associated with endometrial-related conditions is biased by the menstrual cycle effect in the transcriptomic analysis

A. Devesa Peiro[1,2], P. Sebastian-Leon[1,3], A. Pellicer[2,3,4], P. Diaz-Gimeno[1,3].
[1]IVI Foundation- IVI-RMA Global, Genomic & Systems Reproductive Medicine, Valencia, Spain.
[2]University of Valencia, Pediatrics- Obstetrics & Gynaecology, Valencia, Spain.
[3]Health Research Institute la Fe, Reproductive Medicine, Valencia, Spain.
[4]IVI-RMA Rome, Reproductive Medicine, Rome, Italy.

Study question:

Does menstrual cycle affect the results of case versus control gene expression studies searching for endometrial-related-condition biomarkers?

Summary answer:
Menstrual cycle effect is a confounding variable that must be considered at case versus control gene expression studies to identify reliable biomarkers of endometrial-related conditions.

What is known already:
The endometrium is a dynamic tissue involving many genes whose expression changes throughout the menstrual cycle. Consequently, it is important for differential expression analyses to accurately define at molecular level the cycle phase of the endometrial biopsies. Transcriptomic approaches are increasingly used in Reproductive Medicine to identify genes associated to endometrial-related conditions. However, the menstrual cycle has not been properly considered in these studies and it is an important issue to ensure that biomarker discovery is not biased by the menstrual cycle phase, masking the genes associated to the endometrial condition.

Study design, size, duration:
This *in-silico* analysis involves case vs control studies evaluating endometrial-related conditions: 3 of endometriosis (n=37, n=19, n=111), 1 of recurrent implantation failure (n=115), 1 of recurrent pregnancy loss (n=20) and 1 of uterine pelvic pathology (n=71). The original raw data was re-analysed using the same procedure; and the menstrual cycle effect was evaluated comparing the differentially expressed genes with and without removing the registered menstrual cycle phase in the analysis.

Participants/materials, setting, methods:
Keywords of endometrial transcriptomics and related conditions were searched at Gene Expression Omnibus (GEO). For each included study, raw data were pre-processed, normalized using quantile method and explored through Principal Component Analysis (PCA) to estimate the menstrual cycle effect on the data. Significant differentially expressed genes (False Discovery Rates (FDR) <0.05) were compared (Fisher's exact test, FDR) for each individual study before and after removing the menstrual cycle effect using linear models (Limma R-package).

Main results and the role of chance:
A total of 42 gene expression individual studies evaluating endometrial-related conditions were retrieved from GEO. Of these, 14.3% had not registered the menstrual cycle phase at the time of endometrial biopsy collection. From the remaining 85.7%, the studies with cycle phase registered for all samples, raw data available, and n > 3 per group were included. In all of them, the three first PCA's components grouped the samples according to the cycle phase rather than to the condition. This menstrual cycle bias was demonstrated in all studies, as a significant increment of an average of 4.6 times more differentially expressed genes were detected when this effect was removed (Fisher's exact tests FDR < 0.0006, 4.40E-

16 < FDR < 5.06E-04). For these studies, on average 46.78% new genes associated with the endometrial-related conditions were highlighted as new condition-specific biomarkers that would not have been detected without removing the menstrual cycle bias in the analysis. Consequently, the cycle phase of endometrial biopsies should not only be registered, but also controlled at a molecular level in the data analysis, to optimize the detection of reliable biomarkers of endometrial-related conditions.

Limitations, reasons for caution:
The bias of the menstrual cycle effect in biomarker discovery has only been evaluated on the available GEO endometrial-related-condition studies that met the criteria. Since this effect has been demonstrated in all of them, it could be extrapolated to case vs control studies evaluating other endometrial-related conditions.

Wider implications of the findings:
This research demonstrates the menstrual cycle bias in biomarker discovery of endometrial-related conditions, introducing a new biomarker taxonomy that could distinguish between genes affected by the menstrual cycle and/or the researched condition. These results establish new guidelines to accurately detect endometrial biomarkers, improving reproducibility and gaining statistical power.

Trial registration number:
not applicable. Research supported by IVI Foundation, IVI-RMA Global. A.Devesa-Peiro is granted by the Ministry of Science, Innovation and Universities (FPU/15/01398).

Keywords:
menstrual cycle bias
endometrial gene expression
transcriptomics
biomarker discovery
endometrial-related conditions