

Discussion: 'Metabolomic prediction of late-onset preeclampsia,' by Bahado-Singh et al

In the roundtable that follows, clinicians discuss a study published in this issue of the Journal in light of its methodology, relevance to practice, and implications for future research. Article discussed:

Bahado-Singh RO, Akolekar R, Mandal R, et al. First-trimester metabolomic detection of late-onset preeclampsia. *Am J Obstet Gynecol* 2013;208:58.e1-7.

DISCUSSION QUESTIONS

- How can metabolomics be used to identify biomarkers?
- What were the authors' aims?
- What was the study design?
- What were the statistical challenges?
- What information is presented in the tables and figures?
- What additional research is needed?

From the Department of Obstetrics and Gynecology, Washington University in St. Louis, St. Louis, MO:

Moderators

Heather A. Frey, MD
Fellow

George A. Macones, MD, MSCE
Mitchell and Elaine Yanow Professor and Head

Discussants

Jenifer Allsworth, PhD
Assistant Professor

Lauren Theilen, MD
Second-Year Resident

Janine Spain, MD
Second-Year Resident

The authors report no conflicts of interest.

0002-9378/\$36.00

© 2013 Mosby, Inc. All rights reserved.

<http://dx.doi.org/10.1016/j.ajog.2012.11.021>



See related article, page 58



For a summary and analysis of this discussion, see page 87

INTRODUCTION

Preeclampsia occurs in 5-8% of pregnancies, most commonly presenting at term as late-onset disease.¹ Primary prevention is an enticing goal for researchers and clinicians, but any strategy would require accurate identification of high-risk women early in pregnancy. In a new study, Bahado-Singh and colleagues demonstrated that metabolomics—the comprehensive examination of metabolites—might help distinguish these patients during the first-trimester. In short, metabolomics considers the structure of these small molecules, their interactions, and reactions to modifications in their surroundings. Journal Club members agreed that while the researchers' results must be validated in other populations, their work was well-designed and provocative.

Heather A. Frey, MD and George A. Macones, MD, MSCE, Associate Editor

BACKGROUND

Frey: What are biomarkers? How can metabolomics be used to identify biomarkers for diseases like preeclampsia?

Spain: Biomarkers are indicators of a biologic state. They can be objectively measured and evaluated as a sign of a physiologic or pathologic process. Metabolomics can be used to identify metabolites that are present in patients who develop preeclampsia. If some prove to have good sensitivity and specificity for predicting the disorder, these may then be used as a biomarker to determine which patients are more likely to develop preeclampsia.

Frey: How is research that uses biomarkers to predict disease clinically relevant?

Spain: Identification of biomarkers that can predict or diagnose a disease—particularly one in which early

intervention might make a significant difference in outcome—or assess a therapeutic response with both a high sensitivity and specificity can be extremely useful clinically.

Frey: What were the authors' aims?

Theilen: The authors wanted to identify first-trimester maternal serum biomarkers for late-onset preeclampsia. They also wanted to determine the diagnostic accuracy of these biomarkers for predicting late-onset preeclampsia and to assess the ability of metabolomic analysis to distinguish late-onset from early-onset preeclampsia.

STUDY DESIGN

Frey: What was the study design? What are the pros/cons of this type of design?

Spain: This was a nested case-control study. In a case-control study, predictor variables are compared among a group of individuals with a particular outcome and a group without the outcome. Case-control studies are useful as they are relatively quick, inexpensive, and capable of showing associations that can then be used as a basis for a larger randomized controlled trial. The cons of this study design are that an association can be implied, but causality cannot be proven, and as an observational study, it is subject to confounding and bias. A nested case-control study is performed within an existing study cohort, which reduces selection bias.

Frey: What are some of the statistical challenges when analyzing metabolomic data?

Allsworth: As demonstrated by this study of late-onset preeclampsia, metabolomic methods may generate large numbers of possible predictive variables for analysis. The analyst may also find that there are more dependent variables than individ-

ual study participants. These complexities make visualization and summarization difficult when using standard statistical approaches. Moreover, the presence of a large number of dependent variables with varying extents of correlation raise other methodologic concerns, particularly, multicollinearity. Multicollinearity is a common concern when analyzing metabolomic data, as it is likely that multiple first-trimester metabolites observed may be correlated with one another; this correlation may adversely influence variance estimates of regression parameters.

Frey: What are principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA)?

Allsworth: PCA is a nonparametric technique that reduces the number of predictive variables through the estimation of a few linear combinations—called principal components. The first estimated component accounts for the maximum amount of total variance in the observed variables. The second component is not correlated with the first component and accounts for the maximum amount of variance not accounted for by the first component; and so on. The total number of components may equal the number of variables included, but typically, the first few components account for a significant extent of the variance. Once estimated, components may then serve as predictive variables in analysis.

PLS-DA, an approach based on multiple linear regression, was used in this study for classification and biomarker selection. It reduces data through the estimation of latent variables. If a statistically significant discrimination between classes—in this instance, cases and controls—is detected, metabolic biomarkers could be found.

Frey: How does permutation testing minimize the possibility that the observed data separation in PLS-DA is due to chance?

Allsworth: One of the potential limitations of PLS-DA is the accuracy of estimation of class assignment; researchers have proved that this approach may attempt to separate classes when no difference exists. Therefore, validation of

PLS-DA models is essential. One approach, permutation testing, evaluates whether the specific class assigned to an individual study participant is significantly better than random classification into arbitrary groups. In this instance, the permutation tests were repeated 2000 times.

RESULTS

Frey: What information is presented in Table 2 and Figures 2 and 3?

Theilen: Table 2 presents the mean serum metabolite concentrations for 41 metabolites as determined by nuclear magnetic resonance (NMR) in the late-onset preeclampsia and control groups. The investigators identified 17 metabolites where a statistically significant difference in levels was noted between women with late-onset preeclampsia and controls ($P < .05$). Figure 2 represents the PLS-DA in 2 and 3 dimensions for the NMR-based metabolomic measurements of serum from women with late-onset preeclampsia compared to controls. This type of analysis demonstrates segregation of the 2 groups, indicating that a significant difference in metabolite concentration exists between the 2 groups. Figure 3 illustrates the variable importance in projection (VIP) ranking, which ranks the metabolites based on their importance in women with late-onset preeclampsia vs controls. It also indicates whether their concentrations are increased or decreased. The most powerful group discriminators were glycerol, carnitine, and methylhistidine, all of which were increased in the late-onset preeclampsia group compared to the control group.

Frey: How are prediction models created and assessed? How do you interpret the test characteristics of the prediction models presented in Tables 3 and 4?

Spain: Prediction models use a limited number of independent variables to create probability equations for development of an outcome; in this study, biochemical and demographic variables. The sensitivity and specificity of the prediction models are assessed. Receiver operating characteristics (ROC) curves are

plotted for continuous or ordinal data based on the sensitivity and on the false positive rate, computed as 1-specificity. The calculated area under the curve (AUC) for the graph indicates the ability of the model to predict the outcome (1.0 represents a perfect test and 0.5 a test that is no more predictive of an outcome than chance). The prediction models presented in Tables 3 and 4 show that the use of metabolites in combination with demographic variables was able to predict late-onset preeclampsia with high specificity and an area under the curve that suggested a fair prediction model.

CONCLUSIONS

Frey: How does the heterogeneity of late-onset PE affect the results of this study?

Theilen: As described in the article, late-onset preeclampsia is a disease with a more diverse underlying pathophysiology than its early-onset counterpart. Metabolomics may allow for the identification of a unifying factor in the pathogenesis of late-onset preeclampsia.

Frey: What additional research is needed?

Theilen: Given the fairly sophisticated technical nature of this study, it will be important to validate the authors' findings at other institutions. Additionally, it would be interesting to see whether differences in maternal serum metabolites within the late-onset preeclampsia group correlate with disease phenotypes, as the diagnosis of late-onset preeclampsia encompasses a broad range of morbidities for mother and fetus. Lastly, in order to use the authors' findings in a clinical setting, there needs to be a benefit to early detection of preeclampsia in terms of prevention and treatment. Further research may be warranted for targeting the serum metabolite differences noted in this study as a means for therapeutic intervention. ■

REFERENCE

1. American Congress of Obstetricians and Gynecologists Committee on Practice Bulletins—Obstetrics. ACOG practice bulletin no. 33, January 2002. Diagnosis and management of preeclampsia and eclampsia. *Obstet Gynecol* 2002;99:159-67.