Quantifying the Contribution of Environmental Factors in Perinatal Health Disparities: Conceptual Model and Analytical Challenges

Andrea Bellavia, PhD

Department of Environmental Health Harvard T.H. Chan School of Public Health *abellavi@hsph.harvard.edu* 

USDOHaD - October 1, 2018

#### Acknowledgements

- Tamarra James-Todd
- Paige Williams
- Linda Valeri
- Ami Zota
- Susanna Mitro

The talk is based on the following papers:

- Bellavia A, Zota AR, Valeri L, James-Todd T. Multiple mediators approach to study environmental chemicals as determinants of health disparities. Environmental Epidemiology. 2018 Jun 1;2(2):e015.
- Bellavia A, James-Todd T, Williams PL. Approaches for incorporating environmental mixtures as mediators in mediation analysis. Under review

#### Perinatal Health Disparities

- Health disparities are differences in health status that systematically have an adverse effect on the health of less-advantaged populations
- Groups that are considered to be disadvantaged may depend on culture place, sociological, economic or anthropological constructs
- Several disparities in perinatal outcomes have been reported (e.g. racial/ethnic disparities in preterm birth)



To quantify a given disparity we fit a statistical model for testing the association between the social construct X and the outcome Y

 $X \longrightarrow Y$ 

The main characteristic of X is to be non-modifiable. Vulnerable populations can be identified, but interventions and recommendations are more challenging

Evaluating chemicals as a

#### Reducing Health Disparities

- If the construct is un-modifiable, how can we reduce the disparity?
- We would like to develop effective public health recommendations, as well as to implement interventions
- One possibility is to identify proximal risk factors that (partly) explain why these groups show enhanced vulnerability to a specific set of diseases



Specifically, the aim is to identify one or more modifiable factors that 1) are risk factors for Y, and 2) are unequally distributed over levels of X



Statistical methods such as mediation analysis, path analysis, or structural equation modeling (SEM), are the common techniques used to evaluate the contribution of such factors Z in determining the X-Y disparity

Evaluating chemicals as

#### Environmental Health Disparities

A growing body of literature suggests that environmental factors may be important contributors to several health disparities (that is: Z from the previous graph can be an environmental factor E)



#### Example - Race/ethnicity and preterm birth

- In the US, racial/ethnic disparities in the risk of preterm birth have been consistently documented
- Recent studies show a link between high exposure to certain classes of environmental chemicals and preterm birth. (Ferguson et al., 2012)
- Racial/ethnic differences in the exposure to these chemicals have also been observed (James-Todd et al., 2014)
- Does the higher exposure to chemicals (partly) explain the higher risk of preterm birth in specific racial/ethnic subgroups?

- So far, little has been done to quantify the contribution of environmental factors in health disparities
- One of the reasons for this literature gap is the additional number of challenges that evaluating environmental factors involves

# Integrating Sources of Environmental Exposures (Bellavia et al. 2018)

- Health disparity research requires the identification of modifiable risk factors for risk reduction in vulnerable populations
- However, the focus in EH research is often on biomarkers as a proxy of exposures
- A conceptual model for environmental health disparities may require the simultaneous inclusion of biomarkers of exposures and their modifiable sources





Formally, to identify interventions aimed at reducing disparities, we need to (somehow) include modifiable sources of E, named B

#### Why is this challenging?

Two arrows are missing from the previous figure:



Red arrow: the identified source B may not be the only source of E to be unequally distributed over levels of X. In this case, intervening on B will only partially reduce the X-E disparity

Environmental Health Disparities	Integrating sources of exposure	Evaluating chemicals as a mixture	



Blue line: The mechanisms through which B has an impact on Y may be more than just via E (e.g. fast-food and preterm birth). Intervening on B will block all these mechanisms

#### Available methods

- ► We would like to assess: 1) the proportion of X-Y disparity that is due to E, and 2) how much of this proportion would be reduced by intervening on B
- These can be assessed by using recent developments in mediation analysis (i.e. multiple sequential mediators)
- Several caveats need to be taken into account:

- ► 1) The use of the counterfactual approach for mediation is required to identify path-specific effects. This however, requires thinking of effects in terms of interventions
- Is the social construct (like race) a cause? There's no scientific agreement on this point
- 2) A study where all X, B, E, and Y were assessed, possibly at different and sequential time-points, is a clear advantage. Does such study exist?
- 3) Environmental chemicals are likely to present non-linear effects. Non-linearities in mediation analysis can be handled but interpretation is more complicated



 4) If multiple sources are known and available they can be included, but both implementation and interpretation become harder



## Evaluating multiple environmental factors as a mixture (paper under review)

- Humans are generally exposed to a mixture of environmental factors
- Several studies (Taylor et al., 2012) recommend evaluating exposure to environmental mixtures rather than one exposure at the time
- Such reasoning extends to the evaluation of environmental health disparities. A potential contribution to the disparity is more likely given by a mixture of environmental components.

Environmental Health Disparities	Integrating sources of exposure	Evaluating chemicals as a mixture 000000	

- Nevertheless, the different components of the mixture may have different chemical/biological properties and therefore different effects on Y.
- Ideally, we would like to both identify the overall contribution of the mixture, as well as the specific contribution of its components
- Furthermore, interactions between mixture components should be taken into account



Andrea Bellavia, PhD Environmental Disparities

#### Methods

Several methods for mixtures have been developed. A broad review can be found in Taylor et al., 2012:

Single chemical analysis	Classic linear regression
	(ordinary least squares)
Multiple regression	Classic linear regression
	(ordinary least squares)
Visualization, structural equation modeling (SEM), and principal component analysis (PCA)	Classification and prediction
Informed sparse PCA and segmented regression	Classification and prediction
Bayesian g-formula	Classification and prediction
PCA	Classification and prediction
Classification and regression trees (CART)	Classification and prediction
Bayesian profile regression	Classification and prediction
Random forest	Classification and prediction
Multivariate adaptive regression splines (MARS)	Classification and prediction
Bayesian non-parametric regression	Classification and prediction
Bayesian additive regression trees (BART) and negative sparse PCA (NSPCA)	Classification and prediction
Conformal predictions	Classification and prediction
Bayesian kernel machine regression (BKMR)	Exposure-response surface estimation
Building Bayesian networks	Exposure-response surface
Exposure surface smoothing (ESS)	Exposure-response surface estimation
Modes of action (results presented for $Z=0$ strata)	Other
easible solution algorithm (FSA)	Other
exploratory data analysis (EDA)	Other
Novel approach and least-angle regression (LARS)	Variable selection
Machine learning	Variable selection
Two-step variable selection and least absolute shrinkage and selection operator (LASSO)	Variable selection
Two-step shrinkage-based regression	Variable selection
Factor mixture models	Variable selection
Subset and bootstrap	Variable selection
Variable selection regression (VSR)	Variable selection
Bayesian estimation of weighted sum	Variable shrinkage strategie
Shrinkage methods (LASSO/LARS)	Variable shrinkage strategie
Weighted quantile sum regression (WQS)	Variable shrinkage strategie
LASSO	Variable shrinkage strategie

Andrea Bellavia, PhD Environmental Disparities We identified and described two main approaches to integrate such methods in a disparity (mediation) model:

 a) Use scoring (e.g. weighted quantile sum) or classification (e.g. principal component analysis) methods and reduce the mixture to one or few independent (and uncorrelated) mediators.

**Advantage:** approaches based on the correlation structure will likely identify the "real-world" exposure patterns. **Limitation:** characterizing the specific contribution of each component, as well as identifying interactions, is not straightforward.

b) Use selection methods (e.g. LASSO, regression tree) in a two-stage procedure, first identifying important components and later building a model for the disparity that only includes those selected exposures.

**Advantage:** any other method can be used, including recent advanced approaches **Limitation:** selection will only be based on the E - Y association and not on the full disparity model

### Summary

- Environmental factors are potential contributors to the development of perinatal health disparities, and represent potential target for public health interventions and recommendations
- Few studies have focused on (or are even designed for studying) environmental disparities. A methodological framework for their investigation is not established
- Mediation analysis can be useful for evaluating environmental health disparities, but this framework should take into account specific features such as integrating sources of exposures and evaluating environmental factors as a mixture

## Summary (cont.)

- The most suitable method may depend on the question of interest and on the data structure
- While work remains to be done, methods are available and should be used.
- Both integrating sources of exposures, and moving beyond the analysis of environmental factors one by one, provide a better understanding of environmental effects on perinatal outcomes and their contribution to disparities.

#### References

- Bellavia A et al. Multiple mediators approach to study environmental chemicals as determinants of health disparities. Environmental Epidemiology. 2018 Jun 1;2(2):e015.
- Bobb JF et al. Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. Biostatistics. 2014 Dec 22;16(3):493-508.
- Carrico C et al. Characterization of weighted quantile sum regression for highly correlated data in a risk analysis setting. Journal of Agricultural, Biological, and Environmental Statistics. 2015 Mar 1;20(1):100-20.
- Correia K, Williams PL. A hierarchical modeling approach for assessing the safety of exposure to complex antiretroviral drug regimens during pregnancy. Statistical methods in medical research. 2017 Jan.
- Naimi Al et al. Mediation analysis for health disparities research. American journal of epidemiology. 2016 Aug 3;184(4):315-24.
- Taylor KW et al. Statistical approaches for assessing health effects of environmental chemical mixtures in epidemiology: lessons from an innovative workshop. Environmental health perspectives. 2016 Dec;124(12):A227.