Identifying patterns in environmental mixtures: a Bayesian approach & application to endocrine disrupting chemicals

E A Gibson[1]*, J Goldsmith, F Perera, P Factor-Litvak, J Paisley, J B Herbstman, M-A Kioumourtzoglou

[1]Environmental Health Science Department, Mailman School of Public Health, Columbia University, New York, NY

*e.a.gibson@columbia.edu

COLUMBIA
MAILMAN SCHOOL
OF PUBLIC HEALTH

Pattern Recognition

- Chemical exposure patterns can identify
  - Sources of exposure
  - Behaviors leading to exposure
- Link patterns to adverse health outcomes
  - Efficient regulations
  - Targeted interventions

Non-parametric Bayesian Non-negative Matrix Factorization

- Non-negative continuous priors on pattern loadings and scores
  - Increases interpretability
- Non-parametric prior estimates number of patterns
  - User does not define number of patterns
- No orthogonality constraint
  - Correlated patterns more realistic in EH
- Full posterior distribution of pattern loadings and scores
  - Accounts for uncertainty in estimation

1. Distinct patterns & independent scores
2. Overlapping patterns & independent scores
3. Overlapping patterns & correlated scores
- 100 simulations each
- 4 underlying patterns
- 50 chemicals

Comparing **npBNMF** with other pattern recognition methods in simulations:

npBNMF

PCA

Factor Analysis

NMF
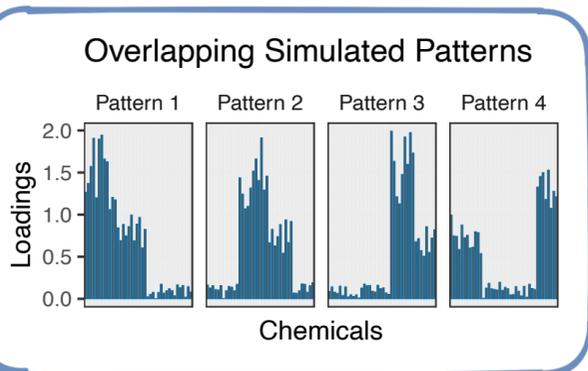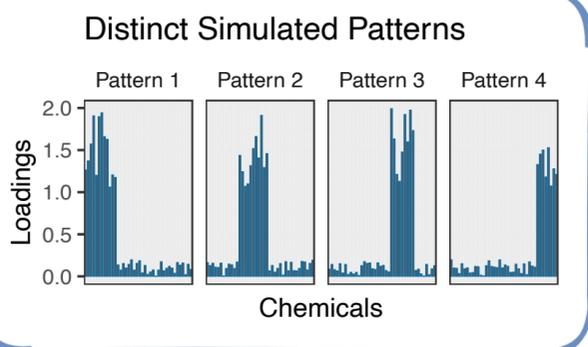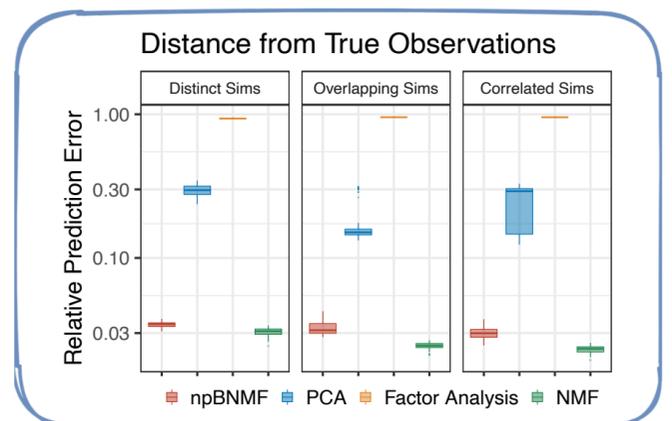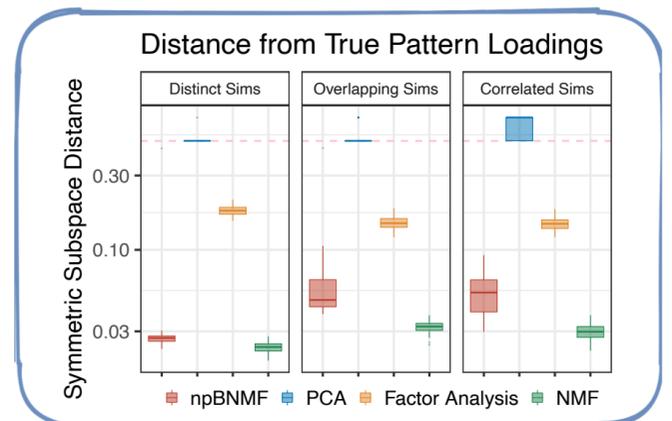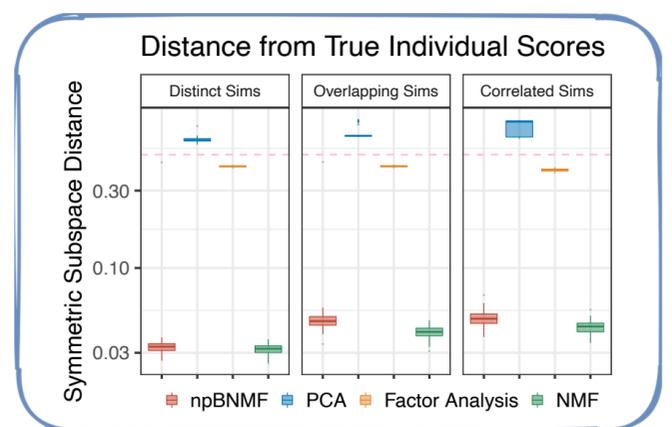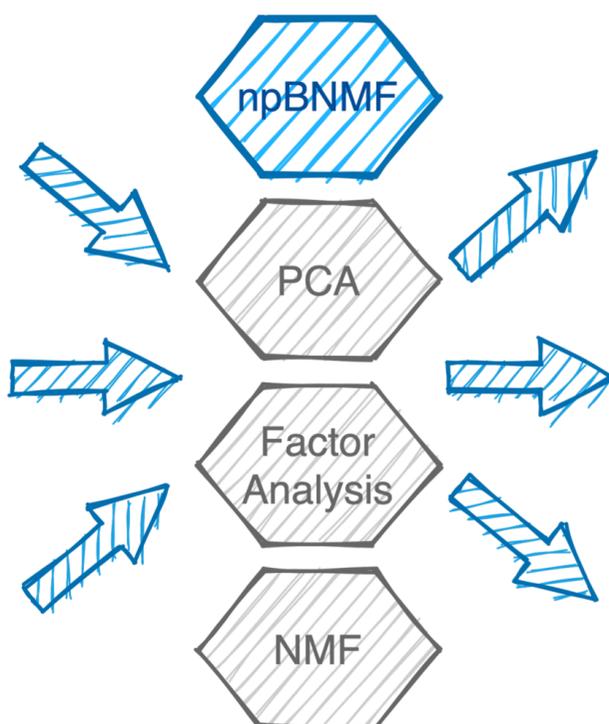
* Ask me about simulation specifications!
* Ask me about distance metrics!
* Ask me about pattern derivation!

Distinct Simulated Patterns

Pattern 1   Pattern 2   Pattern 3   Pattern 4
Loadings
Chemicals

Overlapping Simulated Patterns

Pattern 1   Pattern 2   Pattern 3   Pattern 4
Loadings
Chemicals

Distance from True Individual Scores

Distinct Sims   Overlapping Sims   Correlated Sims
Symmetric Subspace Distance
npBNMF   PCA   Factor Analysis   NMF

Distance from True Pattern Loadings

Distinct Sims   Overlapping Sims   Correlated Sims
Symmetric Subspace Distance
npBNMF   PCA   Factor Analysis   NMF

Distance from True Observations

Distinct Sims   Overlapping Sims   Correlated Sims
Relative Prediction Error
npBNMF   PCA   Factor Analysis   NMF

Research Question: *Can npBNMF identify exposure patterns of phenols, parabens, & phthalates in pregnant women?*

Study Population
- Mothers & Newborns Cohort
  - 342 pregnant women 18-35

Exposure Assessment
- Third trimester spot urine samples
- 5 phenols, 3 parabens, & 9 phthalates

Identified Patterns of EDC Exposure during Pregnancy

Pattern 1   Pattern 2   Pattern 3
Loadings
Class
Parabens
Phenols
Phthalates
Chemicals

Results & Discussion
- 3 exposure patterns:
  - BP-3 alone—found in sunscreen
    - Potential unique exposure route
  - Majority phenols & parabens + MEP
    - Potential shared route through personal care products with fragrance
  - Majority phthalates + BPA
    - Potential shared route through diet
- npBNMF is useful tool to identify environmental exposures patterns
- Identified exposure patterns may be linked to multiple health outcomes