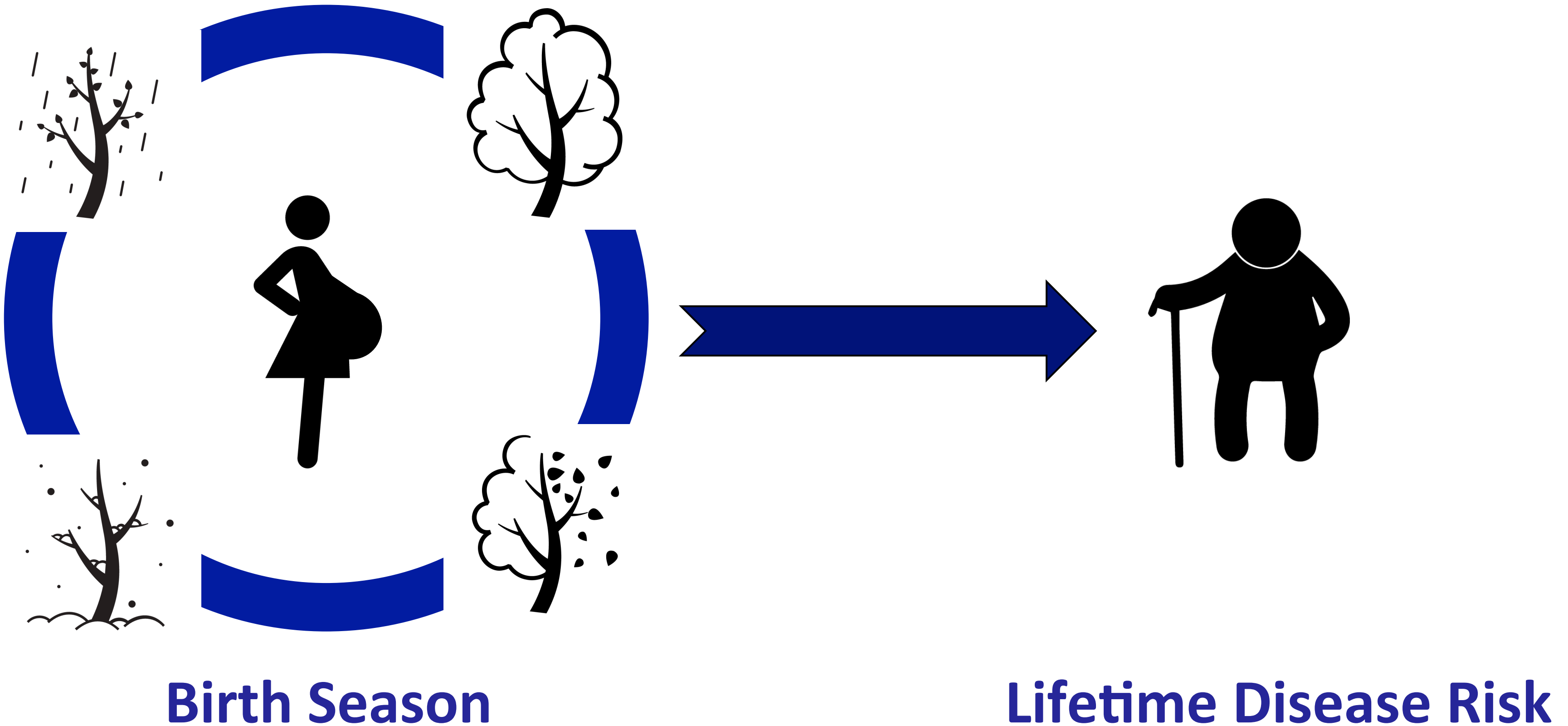
Uncovering Exposures Responsible for Birth Season – Disease Effects: A Global Study

Mary Regina Boland, MA, MPhil, PhD
Assistant Professor of Informatics
Department of Biostatistics, Epidemiology & Informatics,
University of Pennsylvania
Department of Biomedical and Health Informatics,
Children's Hospital of Philadelphia

OHDSI 3rd Annual Symposium
Washington DC, October 18, 2017

Purpose

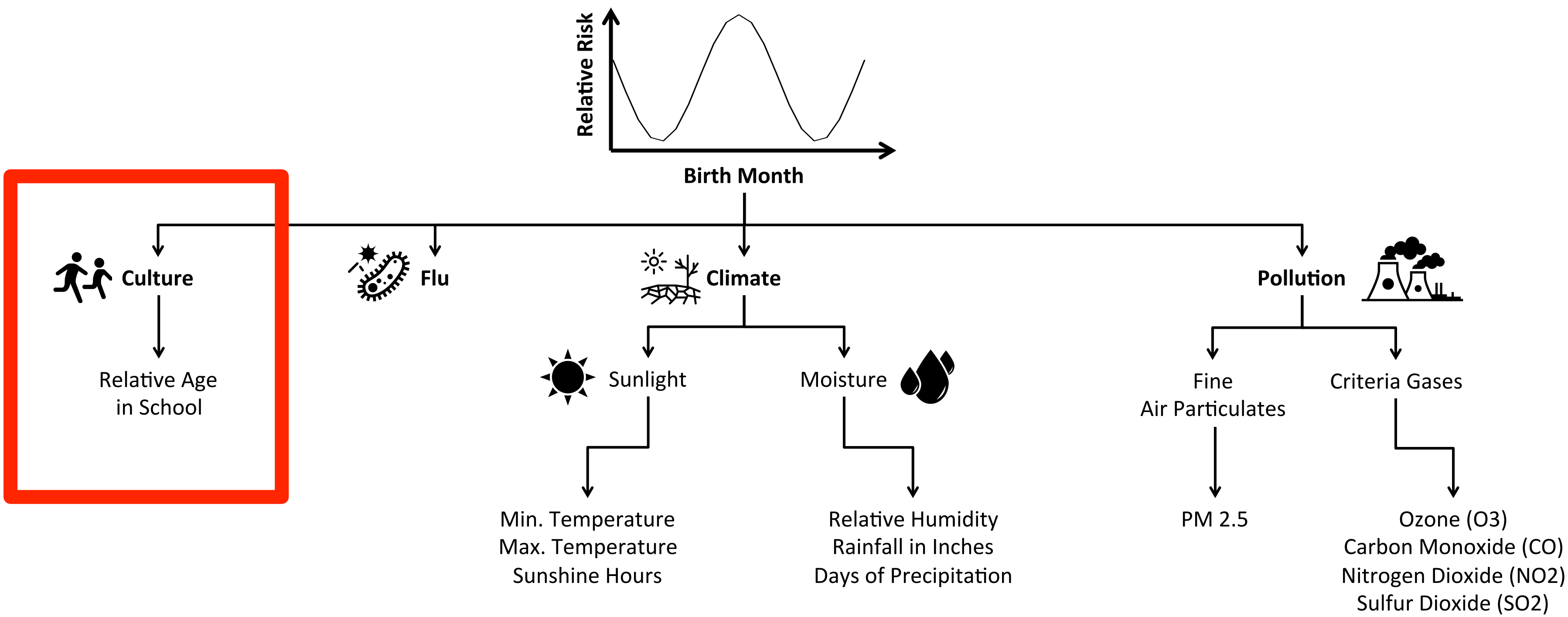
- Two types of effects
- Latent (not readily apparent at birth)
  - Birth season effects
- Overt (readily apparent at birth)
  - Pharmacological drug exposure effects
- Develop informatics methods for using Electronic Health Record data to study these effects


Birth Season
Lifetime Disease Risk

Demographics

Location	Total Number of Patients	% Female	Age	Koppen -Geiger Climate	In- / Out-patient	OMOP CDM Version
New York City: Columbia University Medical Center	1,749,400	54.67%	38 (22, 58)	Cfa	In-patient	V. 4
New York City: Mount Sinai Medical Center	1,169,599	58.03%	53 (36, 66)	Cfa	Both	None
Nashville, Tennessee: Vanderbilt University	3,051,997	51.07%	44 (25, 61)	Cfa	Both	None
Seattle, Washington	1,770,510	50.57%	48 (34, 64)	Csb	Both	None
Taiwan: All areas within Taiwan (99.99% of total population in Taiwan)	909,689	51.07%	35 (20, 50)	Aw	Both	V. 5
Suwon, South Korea: Ajou University School of Medicine	1,848,692	48.26%	42 (28, 57)	Dwa	Both	V.4
	10,499,887					

Factors Influencing Birth Month - Disease Relationships

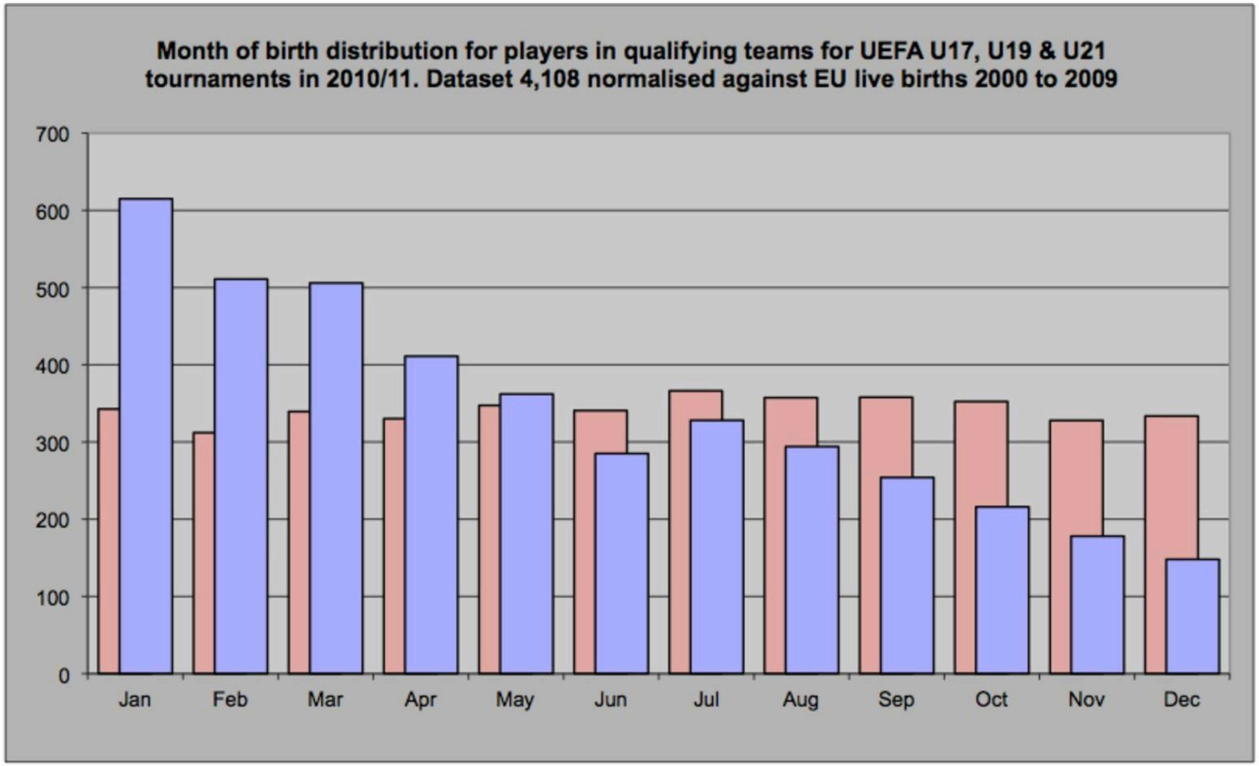


What is Relative Age?

Relative Age Effect

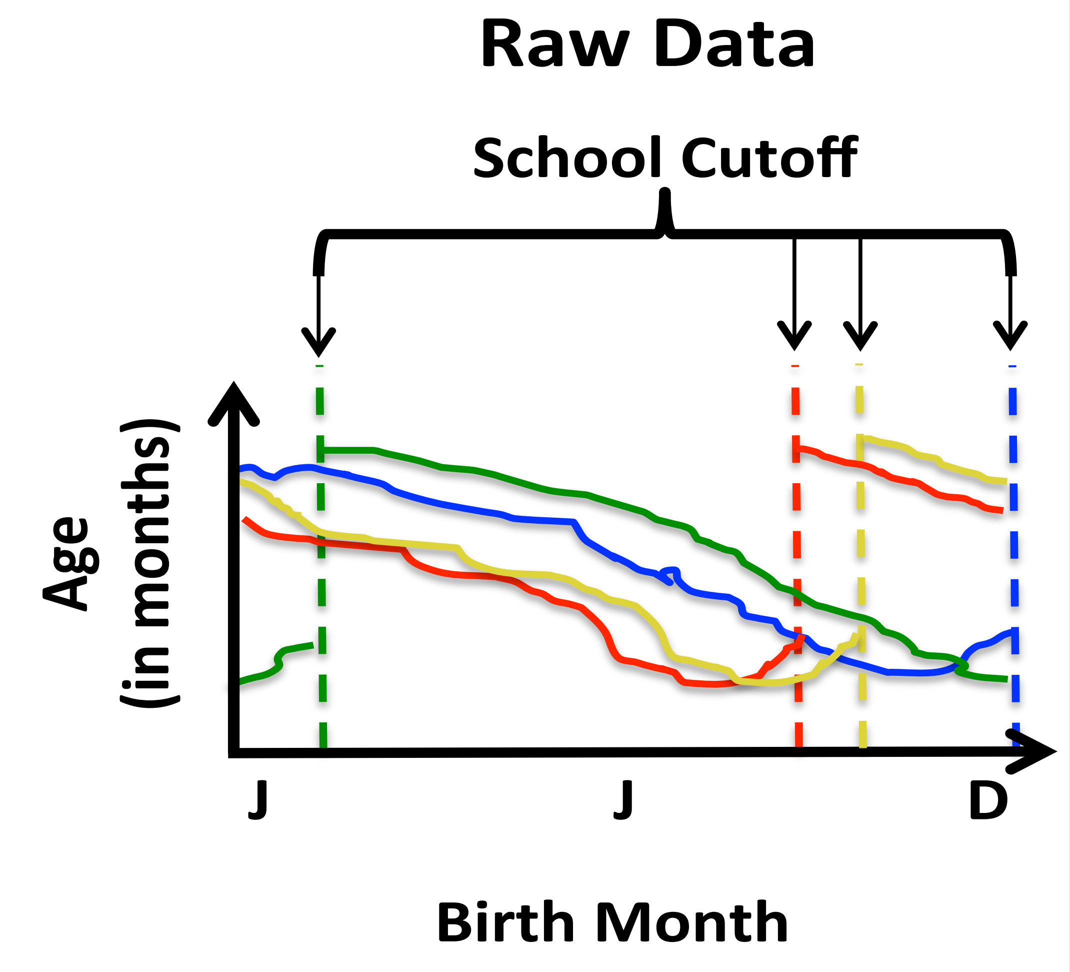
- Age relative to school grade – inadvertently effects the birth month distribution
  - Sports base their cutoff on Jan. 1
    - IOC
    - FIFA
    - AFC
    - CAF
    - CONCACAF
    - CONMEBOL
    - OFC
    - UEFA
    - Older children have the advantage in most competitive sports

Relative Age Effect

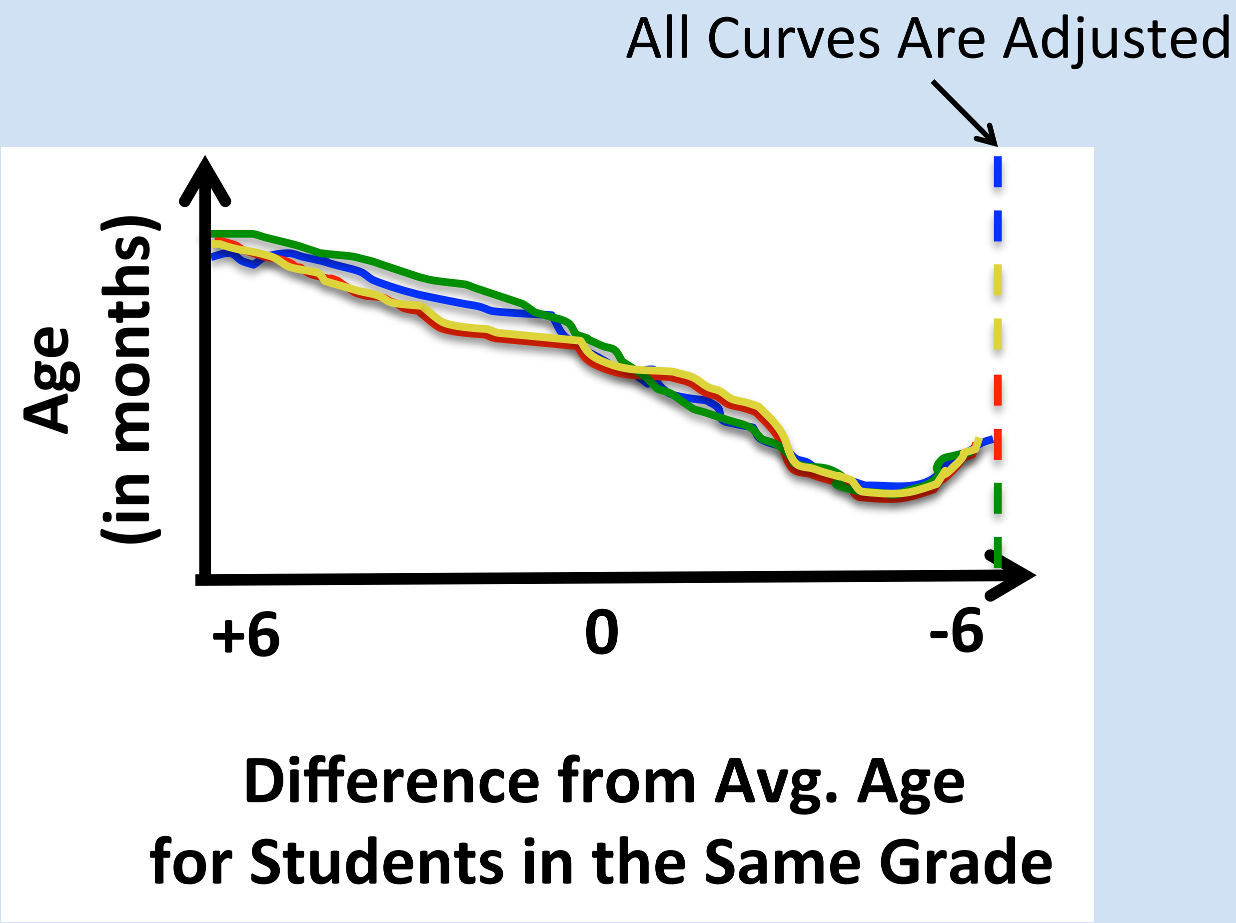


https://en.wikipedia.org/wiki/File:Month_of_birth_distribution_UEFA_youth_tournaments_2010.pdf

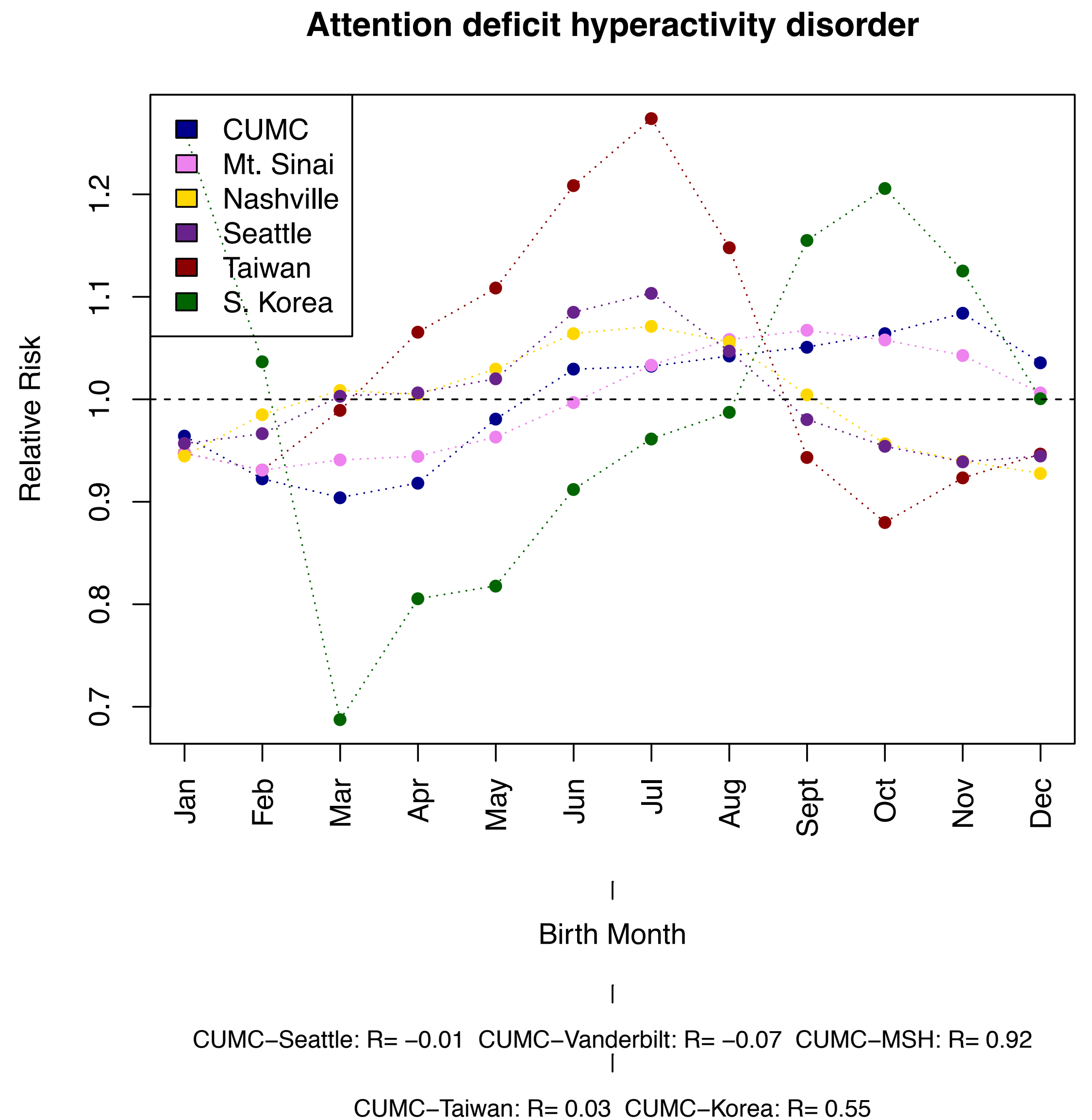
Adjusting for Relative Age



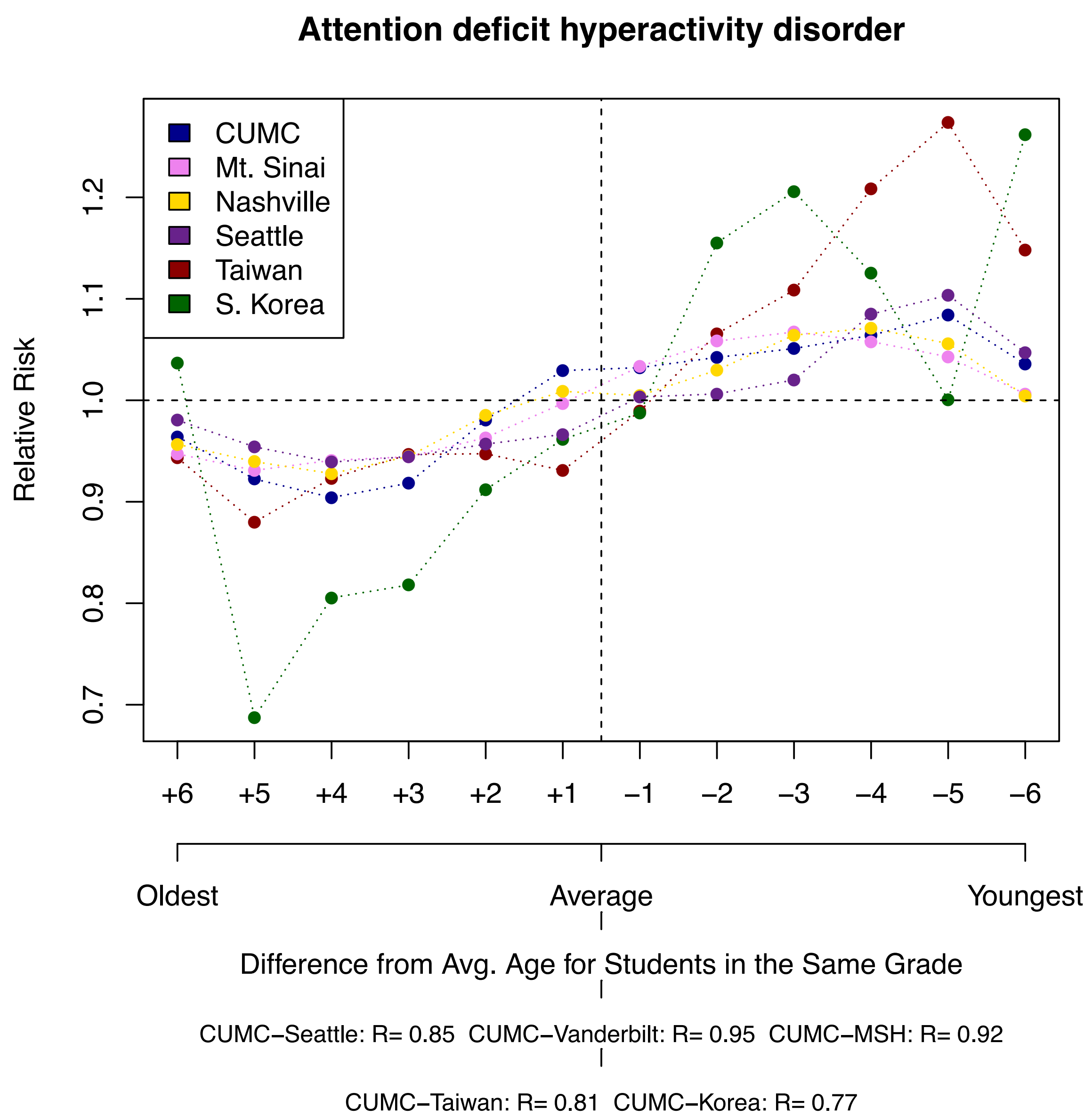
Adjusting for Relative Age



Un-Adjusted


Attention deficit hyperactivity disorder
CUMC
Mt. Sinai
Nashville
Seattle
Taiwan
S. Korea
Relative Risk
0.7
0.8
0.9
1.0
1.1
1.2
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sept
Oct
Nov
Dec
Birth Month
CUMC–Seattle: R= –0.01 CUMC–Vanderbilt: R= –0.07 CUMC–MSH: R= 0.92
CUMC–Taiwan: R= 0.03 CUMC–Korea: R= 0.55

Adjusted



Attention deficit hyperactivity disorder
CUMC
Mt. Sinai
Nashville
Seattle
Taiwan
S. Korea
Relative Risk
0.7
0.8
0.9
1.0
1.1
1.2
+6 +5 +4 +3 +2 +1 −1 −2 −3 −4 −5 −6
Oldest
Average
Youngest
Difference from Avg. Age for Students in the Same Grade
CUMC−Seattle: R= 0.85 CUMC−Vanderbilt: R= 0.95 CUMC−MSH: R= 0.92
CUMC−Taiwan: R= 0.81 CUMC−Korea: R= 0.77

Relative Age and ADHD - Literature

- "**Relative age, as an indicator of neurocognitive maturity, is crucial in the risk of being diagnosed with ADHD** and receiving ADHD medication among children and adolescents. Our findings emphasize the importance of considering the age of a child within a grade when diagnosing ADHD and prescribing medication for treating ADHD" –Chen 2016

- "**Relative age among classmates affects academic performance** among boys and girls into puberty, as well as children' s risk of being prescribed stimulants for ADHD. **This should be taken into account when evaluating children' s performance and behavior in school to prevent unnecessary stimulant prescribing**." – Zoega 2012

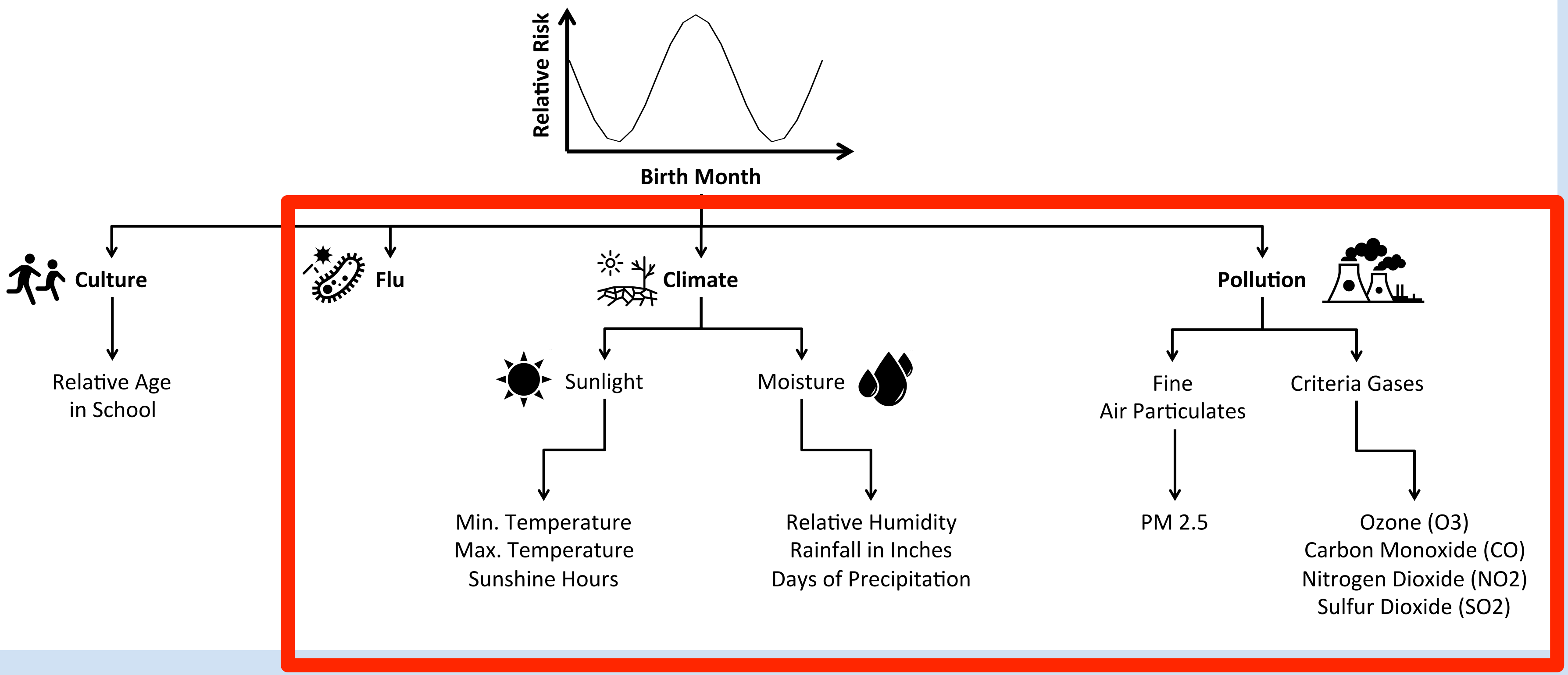
- "Since ADHD is an underlying neurological problem where incidence rates should not change dramatically from one birth date to the next, these **results suggest that age relative to peers in class, and the resulting differences in behavior, directly affects a child's probability of being diagnosed with and treated for ADHD**" –Evans 2006

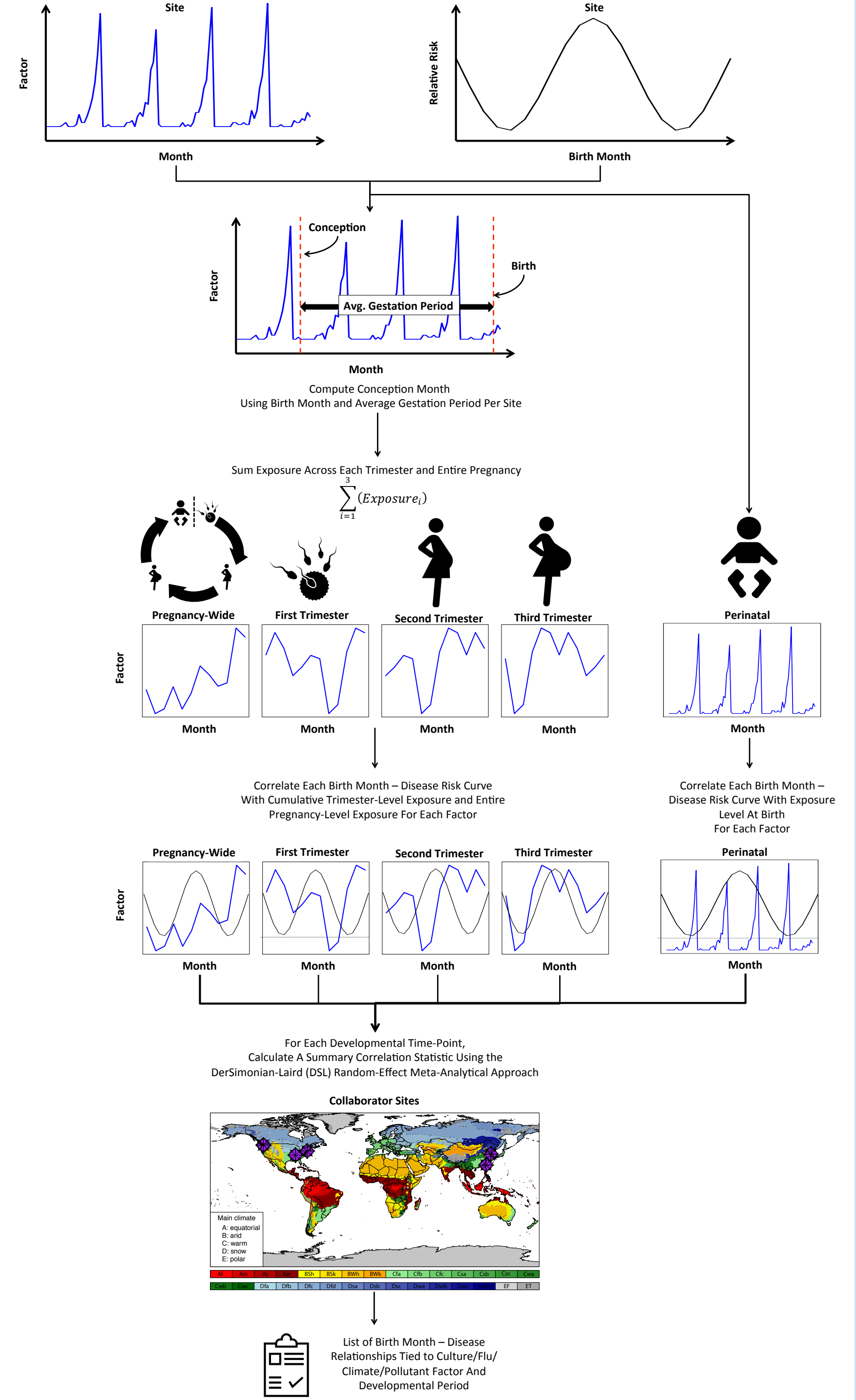
- "CONCLUSIONS. ADHD diagnosis is likely to be influenced by a child's social and school environment as well as exogenous child characteristics. **Concerns that increased pressures for school performance are associated with increased ADHD diagnoses may be justified**." – Schneider 2006

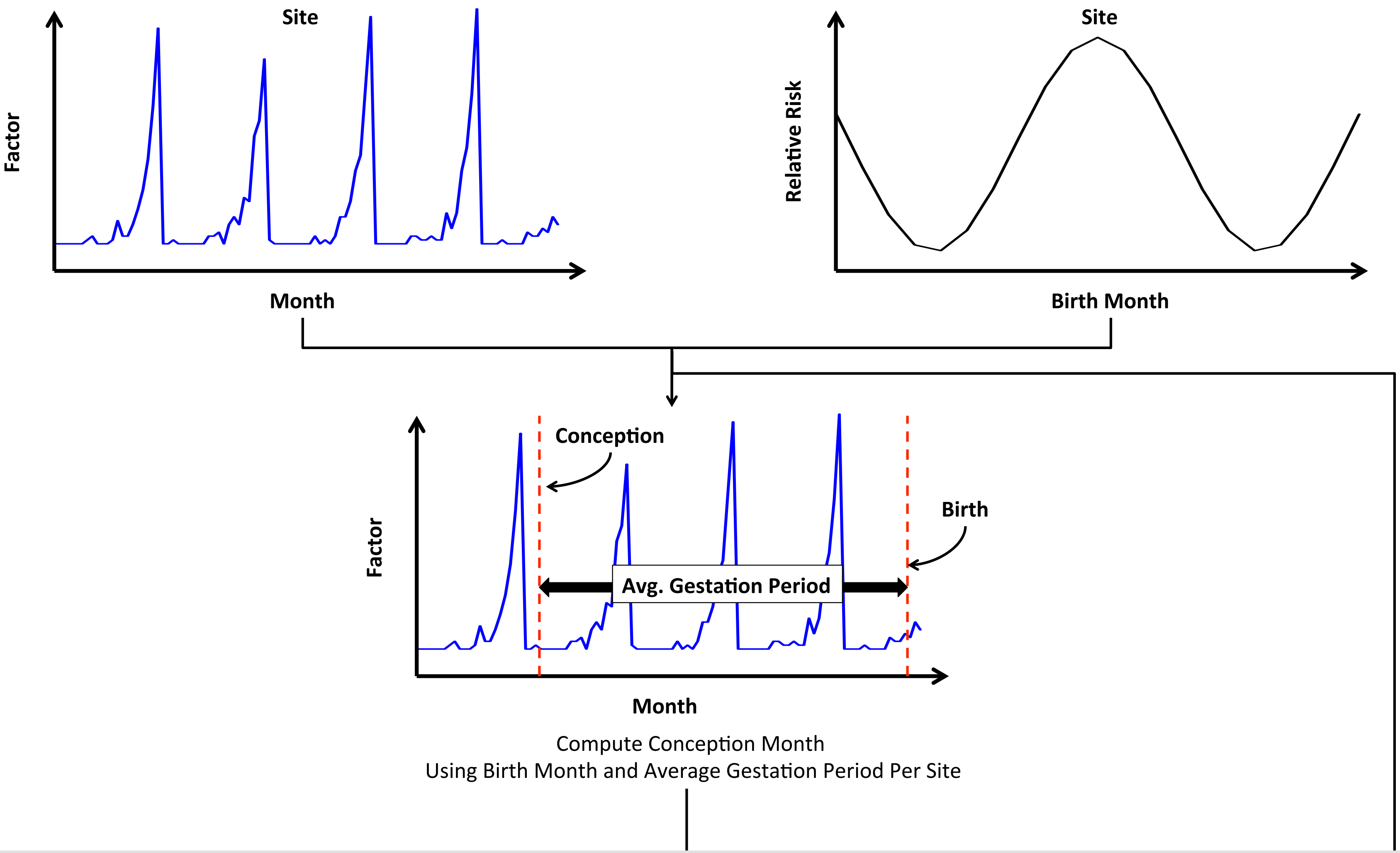
Attention Problems Are Due To School Cutoff Periods and Relative Age Effect

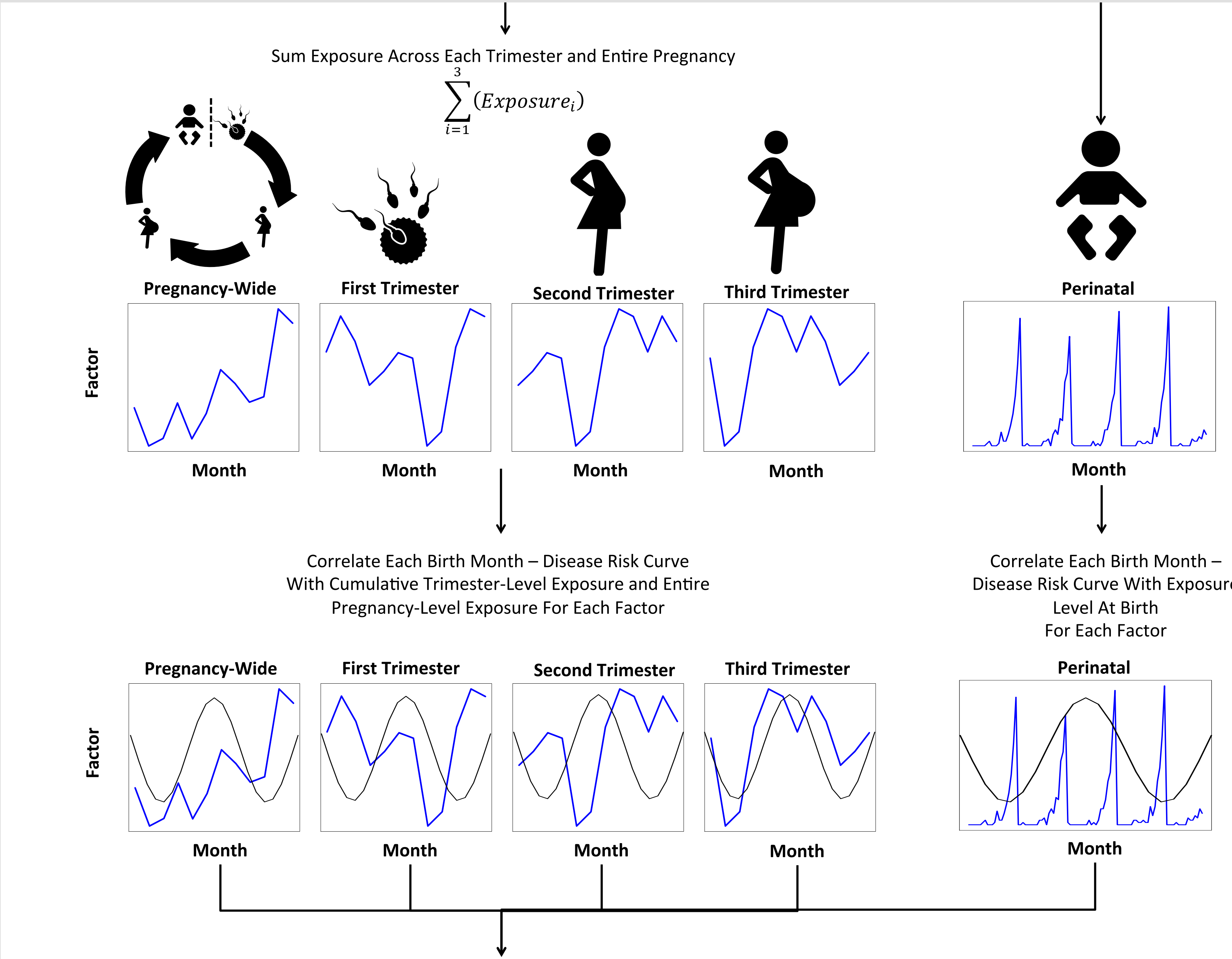
The Average Difference was 17.97% Between Peak High and Low Months!

Factors Influencing Birth Month - Disease Relationships

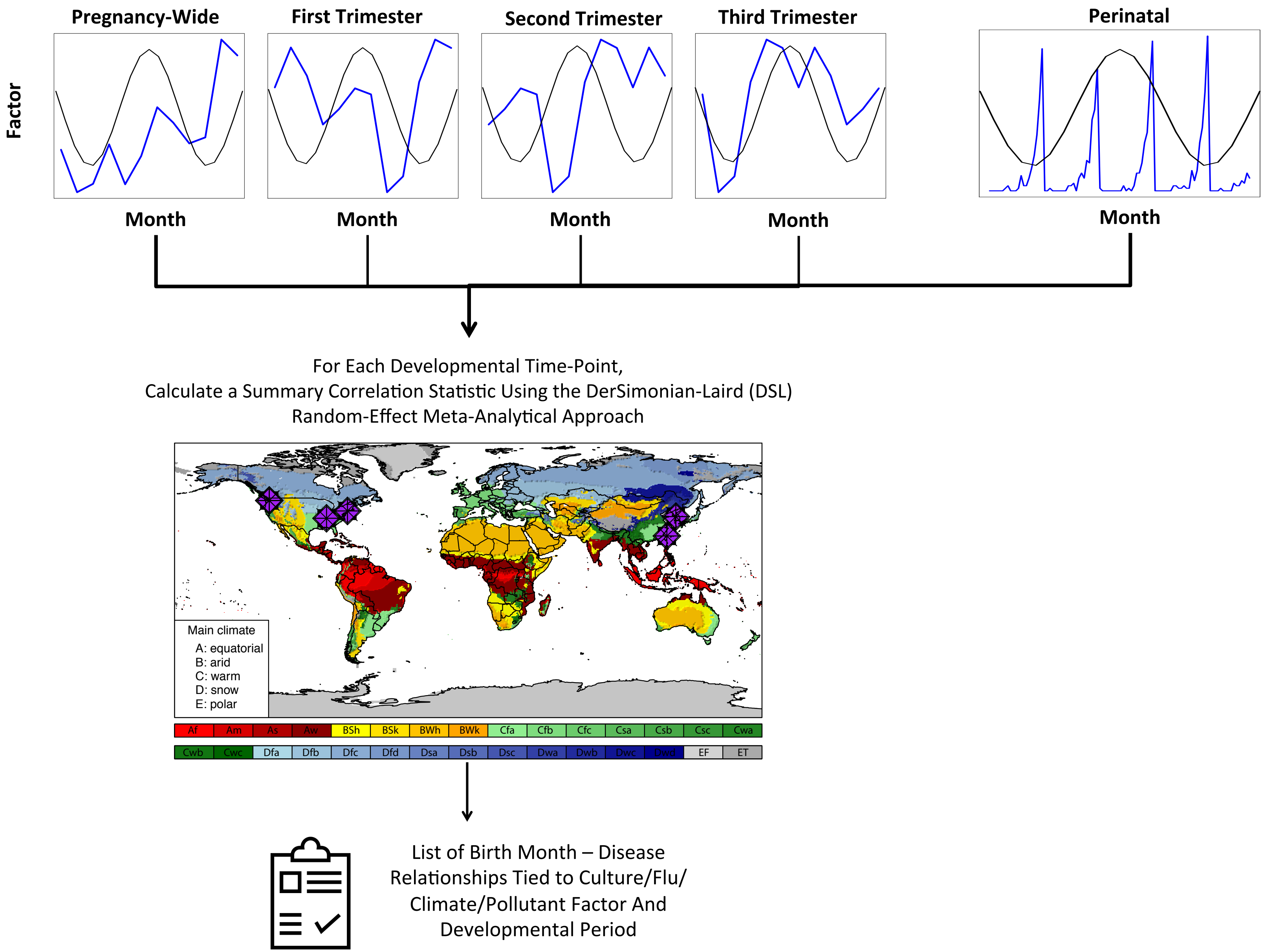



Site
Factor
Month
Site
Relative Risk
Birth Month
Conception
Birth
Factor
Avg. Gestation Period
Month
Compute Conception Month
Using Birth Month and Average Gestation Period Per Site
Sum Exposure Across Each Trimester and Entire Pregnancy
$\sum_{i=1}^{3}(Exposure_i)$
Pregnancy-Wide
First Trimester
Second Trimester
Third Trimester
Perinatal
Factor
Month
Month
Month
Month
Month
Correlate Each Birth Month – Disease Risk Curve
With Cumulative Trimester-Level Exposure and Entire
Pregnancy-Level Exposure For Each Factor
Correlate Each Birth Month –
Disease Risk Curve With Exposure
Level At Birth
For Each Factor
Pregnancy-Wide
First Trimester
Second Trimester
Third Trimester
Perinatal
Factor
Month
Month
Month
Month
Month
For Each Developmental Time-Point,
Calculate A Summary Correlation Statistic Using the
DerSimonian-Laird (DSL) Random-Effect Meta-Analytical Approach
Collaborator Sites
Main climate
A: equatorial
B: arid
C: warm
D: snow
E: polar
List of Birth Month – Disease
Relationships Tied to Culture/Flu/
Climate/Pollutant Factor And
Developmental Period


Site
Factor
Month
Site
Relative Risk
Birth Month
Conception
Birth
Factor
Avg. Gestation Period
Month
Compute Conception Month
Using Birth Month and Average Gestation Period Per Site



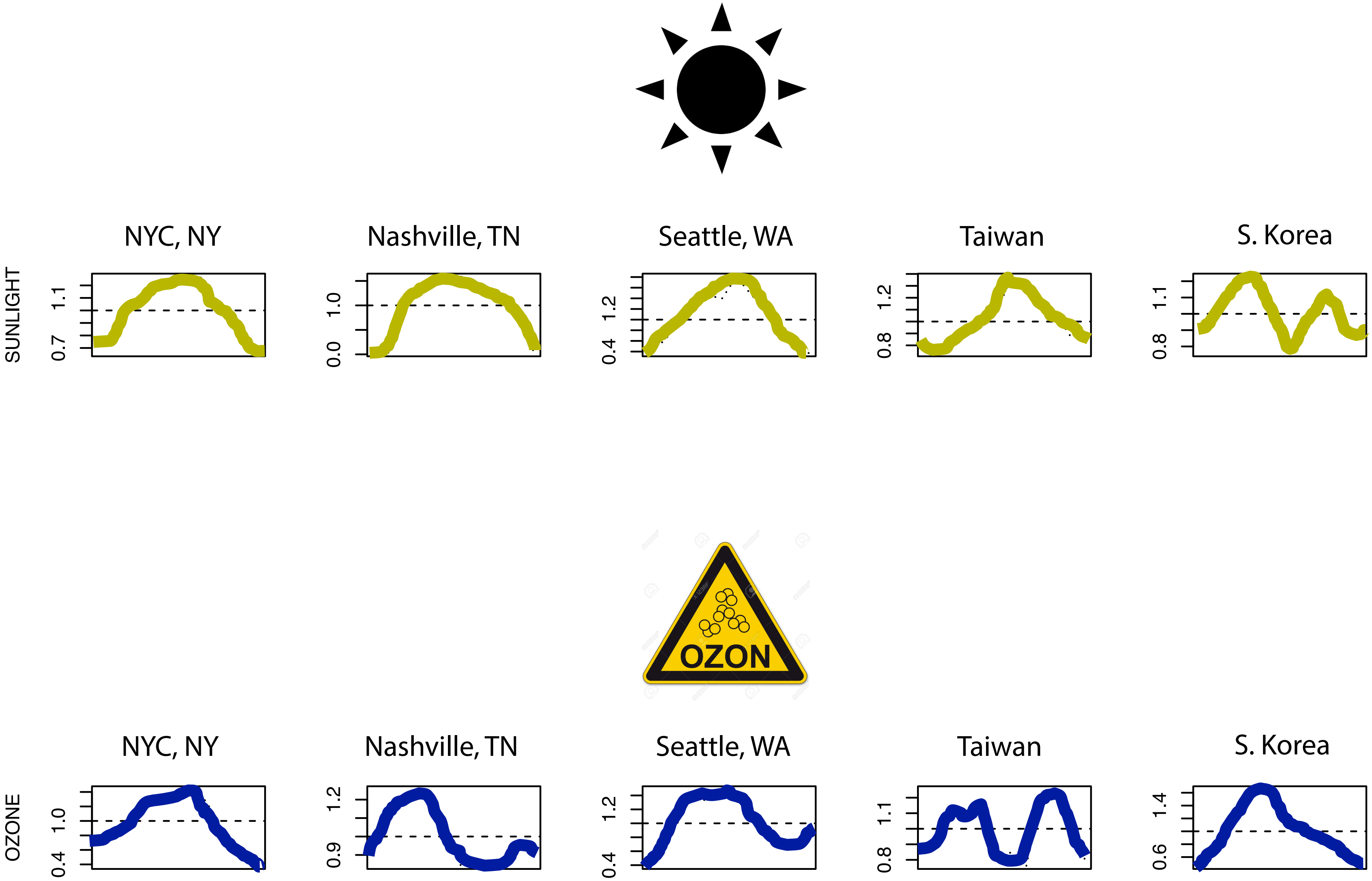
Sum Exposure Across Each Trimester and Entire Pregnancy
$$\sum_{i=1}^{3}(Exposure_i)$$
Pregnancy-Wide
First Trimester
Second Trimester
Third Trimester
Perinatal
Factor
Month
Correlate Each Birth Month – Disease Risk Curve With Cumulative Trimester-Level Exposure and Entire Pregnancy-Level Exposure For Each Factor
Correlate Each Birth Month – Disease Risk Curve With Exposure Level At Birth For Each Factor
Pregnancy-Wide
First Trimester
Second Trimester
Third Trimester
Perinatal
Factor
Month


Pregnancy-Wide
First Trimester
Second Trimester
Third Trimester
Perinatal
Factor
Month
Month
Month
Month
Month
For Each Developmental Time-Point,
Calculate a Summary Correlation Statistic Using the DerSimonian-Laird (DSL)
Random-Effect Meta-Analytical Approach
Main climate
A: equatorial
B: arid
C: warm
D: snow
E: polar
Af Am As Aw BSh BSk BWh BWk Cfa Cfb Cfc Csa Csb Csc Cwa
Cwb Cwc Dfa Dfb Dfc Dfd Dsa Dsb Dsc Dwa Dwb Dwc Dwd EF ET
List of Birth Month – Disease
Relationships Tied to Culture/Flu/
Climate/Pollutant Factor And
Developmental Period

DerSimonian-Laird (DL or DSL) Meta-Analysis Statistic

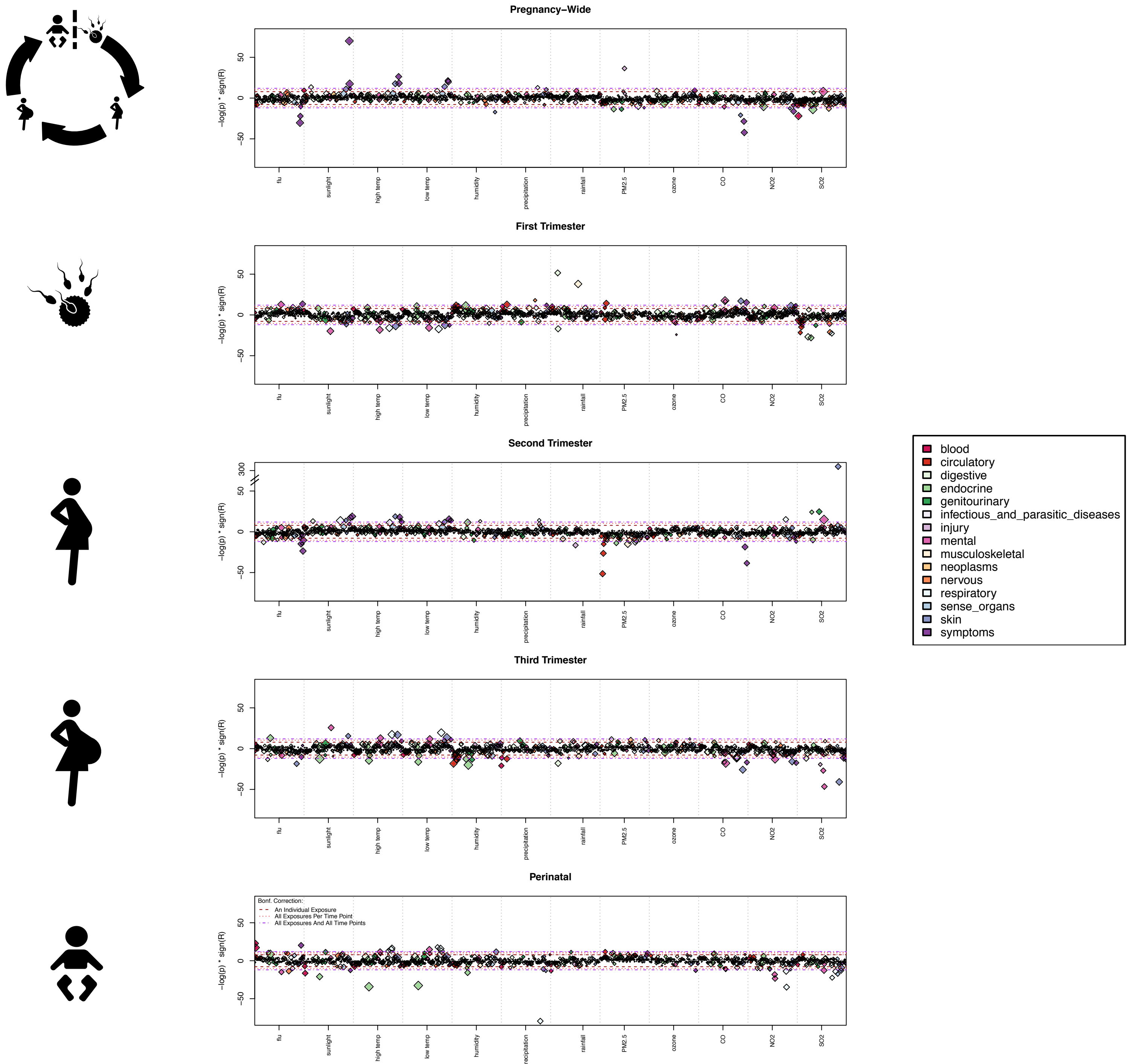
- Approach for Estimating Variance Between Studies
  - Data from >4 sites required for accurate estimation of inter-study heterogeneity
  - DSL works well with large sample sizes (Jackson et al. 2010)
- Effect statistic: Pearson's Correlation
- Effect sizes are weighted by sample size (i.e., N)
- "Random Effects" Approach
  - *Variance of synthesized effect statistic based on idea that studies included in the analysis are a random sample of all possible studies that could have been included*
- Implemented in R using the Schulze 2004 method

Relative Exposures
Vary Across Study Sites



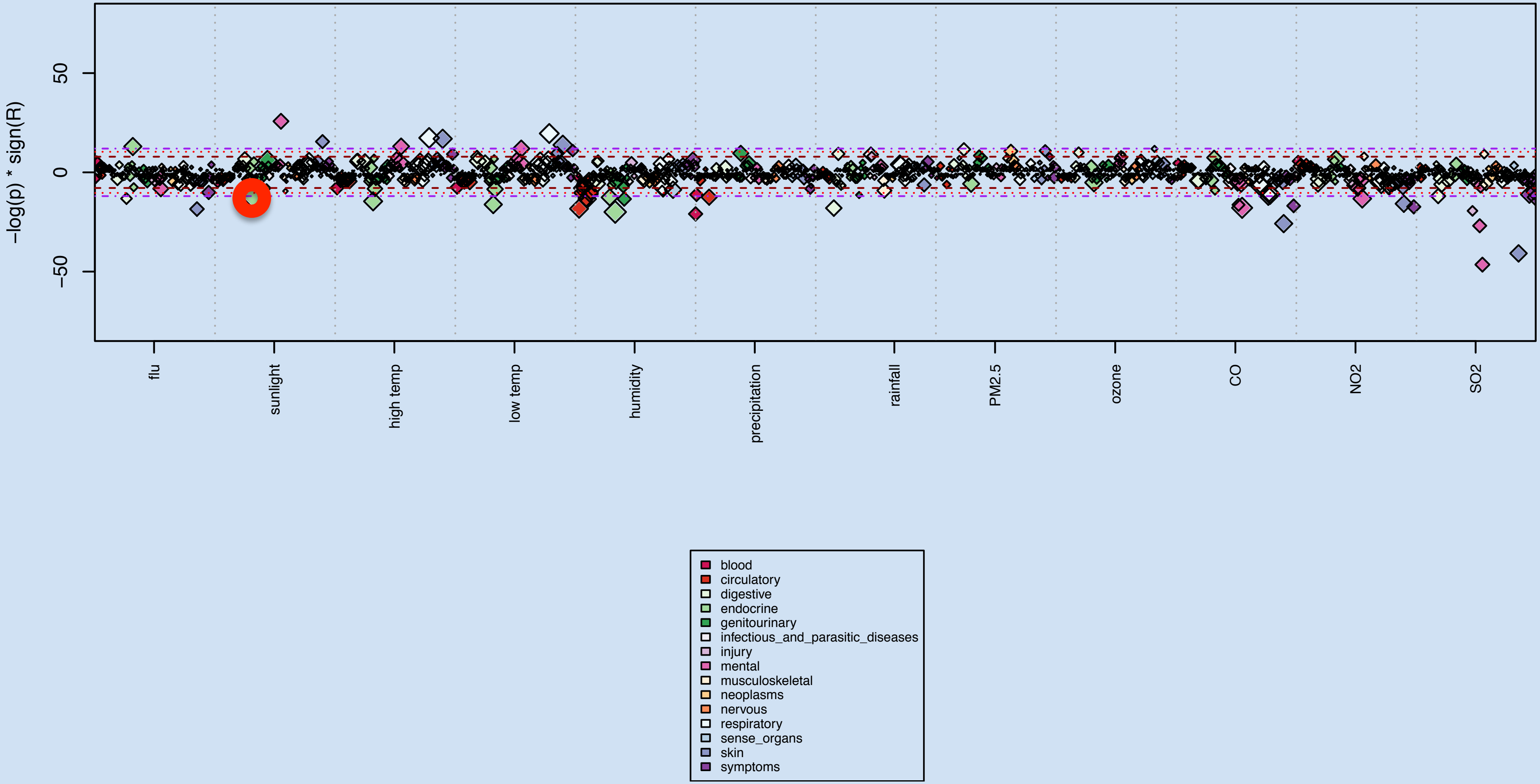
SUNLIGHT
NYC, NY
0.7
1.1
Nashville, TN
0.0
1.0
Seattle, WA
0.4
1.2
Taiwan
0.8
1.2
S. Korea
0.8
1.1
OZON
OZONE
NYC, NY
0.4
1.0
Nashville, TN
0.9
1.2
Seattle, WA
0.4
1.2
Taiwan
0.8
1.1
S. Korea
0.6
1.4

Results: Pooled Site-Wide Correlation



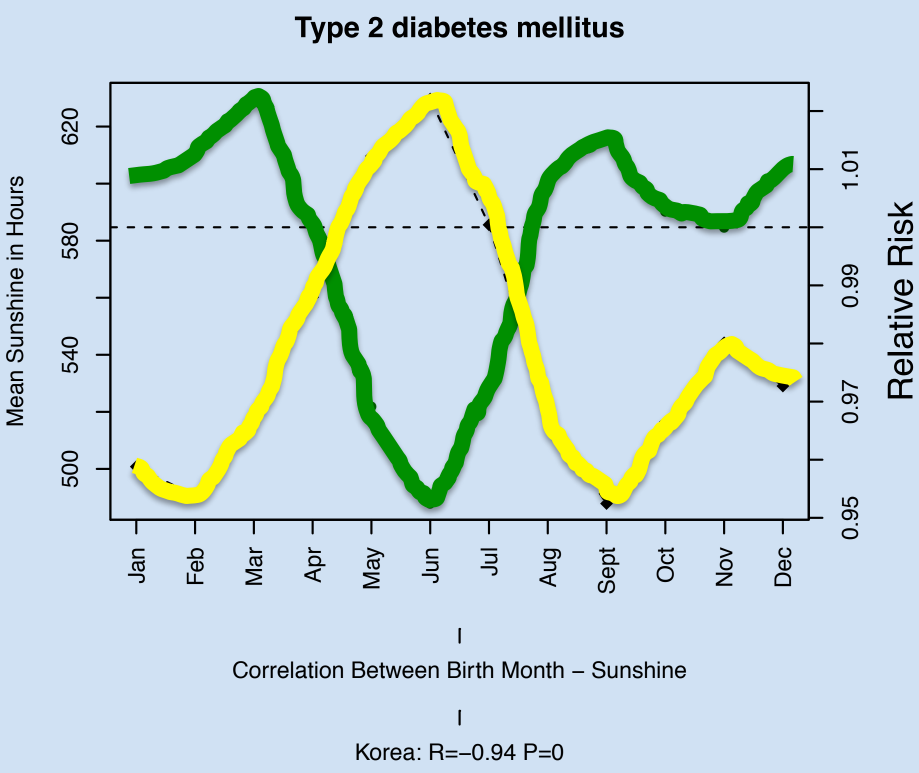
Pregnancy–Wide
First Trimester
Second Trimester
Third Trimester
Perinatal
−log(p) * sign(R)
50
0
−50
300
flu
sunlight
high temp
low temp
humidity
precipitation
rainfall
PM2.5
ozone
CO
NO2
SO2
Bonf. Correction:
An Individual Exposure
All Exposures Per Time Point
All Exposures And All Time Points
blood
circulatory
digestive
endocrine
genitourinary
infectious_and_parasitic_diseases
injury
mental
musculoskeletal
neoplasms
nervous
respiratory
sense_organs
skin
symptoms

Third Trimester

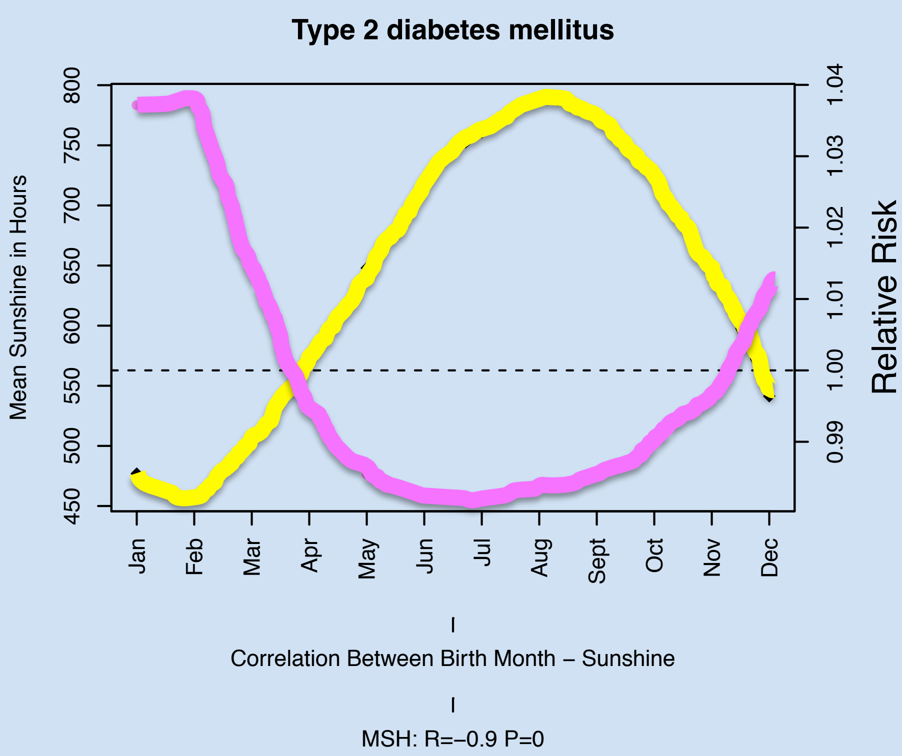


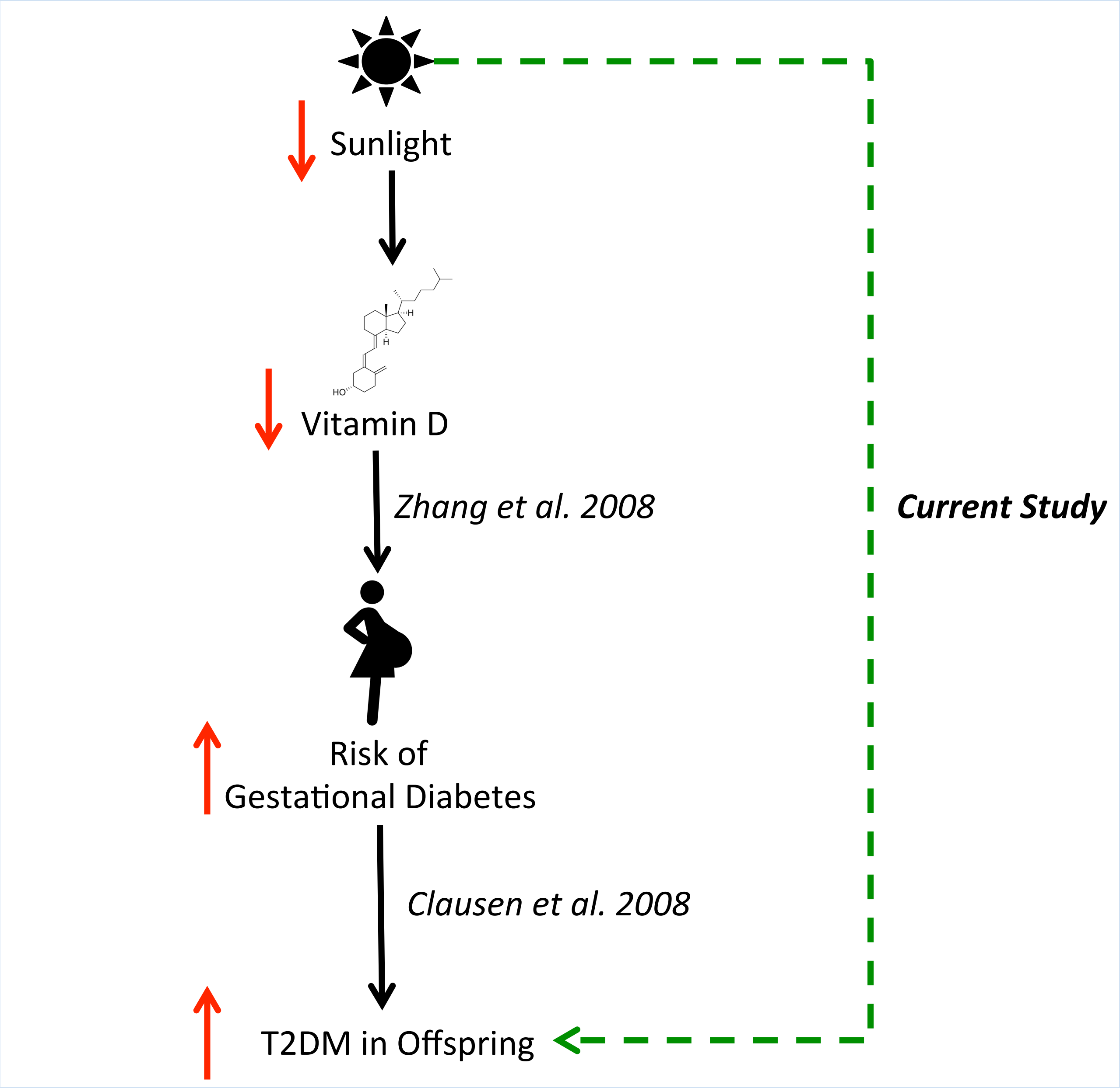
Third Trimester Sunlight and Diabetes

South Korea

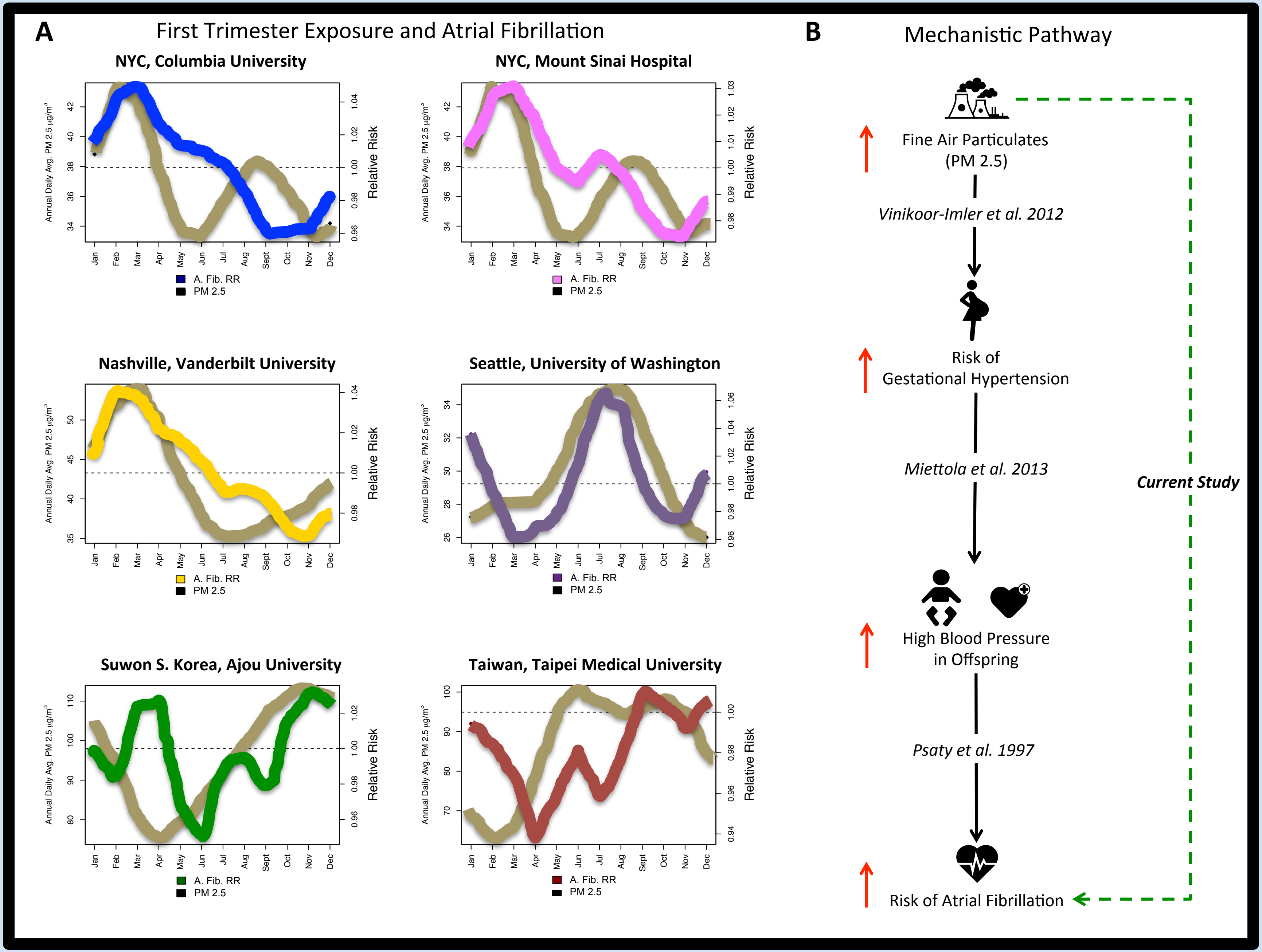


NYC, Mt Sinai





Sunlight
HO
H
Vitamin D
Zhang et al. 2008
Current Study
Risk of
Gestational Diabetes
Clausen et al. 2008
T2DM in Offspring


A
First Trimester Exposure and Atrial Fibrillation
NYC, Columbia University
NYC, Mount Sinai Hospital
Nashville, Vanderbilt University
Seattle, University of Washington
Suwon S. Korea, Ajou University
Taiwan, Taipei Medical University
Annual Daily Avg. PM 2.5 µg/m³
Relative Risk
Jan Feb Mar Apr May Jun Jul Aug Sept Oct Nov Dec
A. Fib. RR
PM 2.5
B
Mechanistic Pathway
Fine Air Particulates
(PM 2.5)
Vinikoor-Imler et al. 2012
Risk of
Gestational Hypertension
Miettola et al. 2013
High Blood Pressure
in Offspring
Psaty et al. 1997
Risk of Atrial Fibrillation
Current Study

Summary of Studies

- Climate is important in understanding disease susceptibility

- Birth Month – Disease Relationships Are Correlated with Certain Exposures During Specific Trimesters Across All Sites

- Results Fit With Known Biological Developmental Pathways

MR Boland, P Parhi, R Miotto, R Carroll, U Iqbal, P-A Nguyen, M Schuemie, SC You, D Smith, S Mooney, P Ryan, Y-C Li, RW Park, J Denny, JT Dudley, G Hripcsak, P Gentine, NP Tatonetti. Uncovering Exposures Responsible for Birth Season – Disease Effects: A Global Study. *J Am Med Inform Assoc.* **2017**; *In press.*
https://academic.oup.com/jamia/article/doi/10.1093/jamia/ocx105/4265709/Uncovering-exposures-responsible-for-birth-season

MR Boland, P Parhi, P Gentine, NP Tatonetti. Climate Classification is an Important Factor in Assessing Quality-of-Care Across Hospitals. *Scientific Reports.* **2017**; 7:4948.
https://www.nature.com/articles/s41598-017-04708-3

Acknowledgments

Dissertation Committee
- Nicholas P Tatonetti (Adviser)
- Andrew Gelman
- George Hripcsak
- Dennis Vitkup
- Pierre Gentine

DBMI Students

Precision Medicine Folks: Sarah Oldham, Sophia Li Ferry, Alex Fedotov, Wendy Chung, Siqin Ye, Daichi Shimbo, Henry Ginsberg

Tuition Funding
- Dean's Fund (Sept. 2013-Jun. 2014)
- T15 LM00707 (Jul. 2014-Jun. 2016)
- NCATS, NIH, TL1 TR000082, formerly NCRR TL1 RR024158 (Jul. 2016 – Jun. 2017)

Other Funding Sources
- R01 LM006910
- R01 GM107145

Colleagues and Collaborators on Key Dissertation Work

Columbia University: Andrew Gelman, Pradipta Parhi, Fernanda Polubriaginof, Zachary Shahn, David Madigan, George Hripcsak, Pierre Gentine, Nicholas P Tatonetti

Mt Sinai: Li Li, Riccardo Miotto, Brian Kidd, Joel Dudley

Vanderbilt University: Robert Carroll, Josh Denny

University of Washington: Don Smith, Sean Mooney

Taiwan: Usman Iqbal, Alex Nguyen, Martijn Schuemie, Jack Li

South Korea: Seng Chan You, Rae Woong Park

OHDSI: Patrick Ryan

Broad Institute: Konrad Karczewski

Acknowledgments

Perelman
School of Medicine
UNIVERSITY *of* PENNSYLVANIA

Questions?

Mary Regina Boland
bolandm@upenn.edu
Twitter: **@MaryRBoland**