National Institute of Environmental Health Sciences
Your Environment. Your Health.

Geospatial Exposures and Tools for Human Cohorts
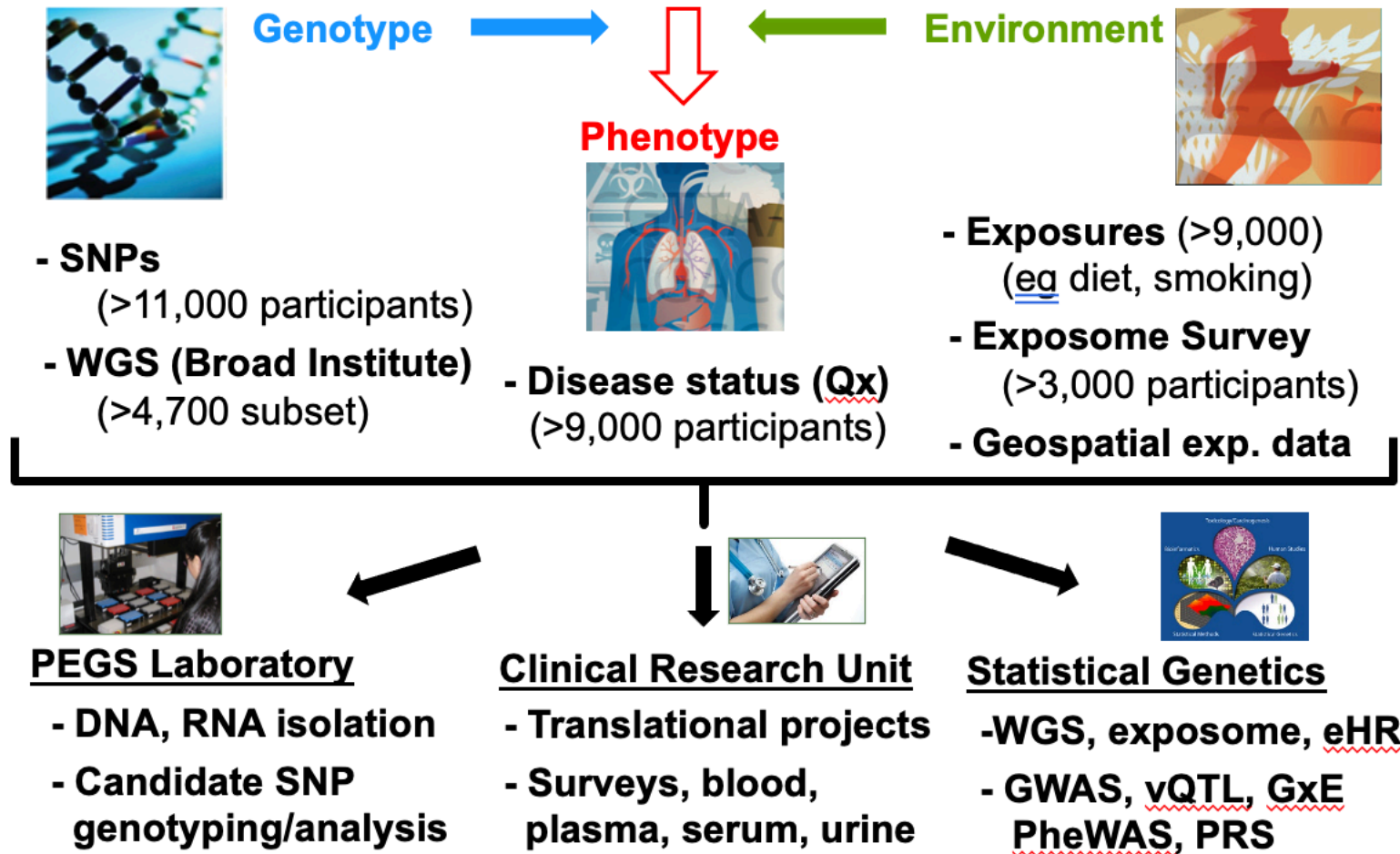
Overview

Complex traits are the results of genes, environment, and their interactions

Exposome data tries to collect the totality of a person's lifetime exposures (an admittedly impossible goal)

Ongoing NIEHS efforts: PEGS cohort as a motivating example

Personalized Environment and Genes Study

Genotype → ← **Environment**

Phenotype

- SNPs
  (>11,000 participants)
- WGS (Broad Institute)
  (>4,700 subset)

- Disease status (Qx)
  (>9,000 participants)

- Exposures (>9,000)
  (eg diet, smoking)
- Exposome Survey
  (>3,000 participants)
- Geospatial exp. data

PEGS Laboratory
- DNA, RNA isolation
- Candidate SNP
  genotyping/analysis

Clinical Research Unit
- Translational projects
- Surveys, blood,
  plasma, serum, urine

Statistical Genetics
- WGS, exposome, eHR
- GWAS, vQTL, GxE
  PheWAS, PRS

Health and Environment Surveys

Health and Exposure (H&E)
- Health, family history of disease, occupational exposures, socioeconomic status, and lifestyle data

Exposome

- Data on internal and external exposures, including medications, physical activity, stress, sleep, diet, chemical and environmental exposures at home and work

Core medical history

- Provides medical history data on diseases and medical conditions, smoking, alcohol, sleep, and surgeries

Exposome Survey

Part A: 'External Exposome'
approximately 200 questions

- Characteristics of current and past homes
- Workplace characteristics
- Chemical and metal exposures at work
- Hobby exposures
- Ultraviolet light exposures

Part B: 'Internal Exposome'
approximately 240 questions

- Medications
- Vitamins, minerals, dietary supplements
- Chemotherapy/radiation therapy
- Physical activity
- Stress
- Infection
- Sleep
- Dietary behavior
- Dietary intake
- Siblings/twins/birth order
- Genetic history

GIS Exposure Estimates

Growing Number of Data Layers

- Airports
- CAFOs
- Cellular towers
- Drinking water
- Dry cleaners
- Hazardous waste
- Highways
- Nuclear sites
- Wastewater

- Population info
- Power lines
- PR landfills
- Railroads
- Spills
- Sanitary landfills
- Superfund sites
- Toxic release sites,
- Etc.

Address at the time of survey completion and longest lived childhood address
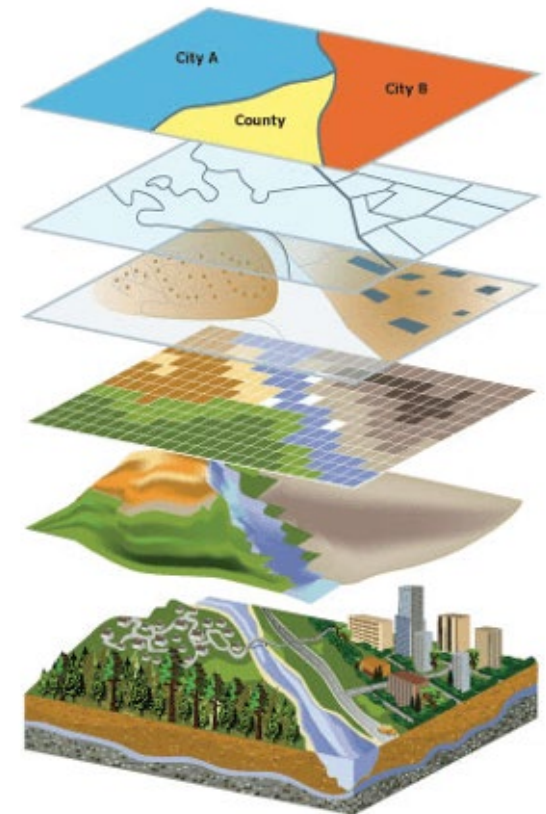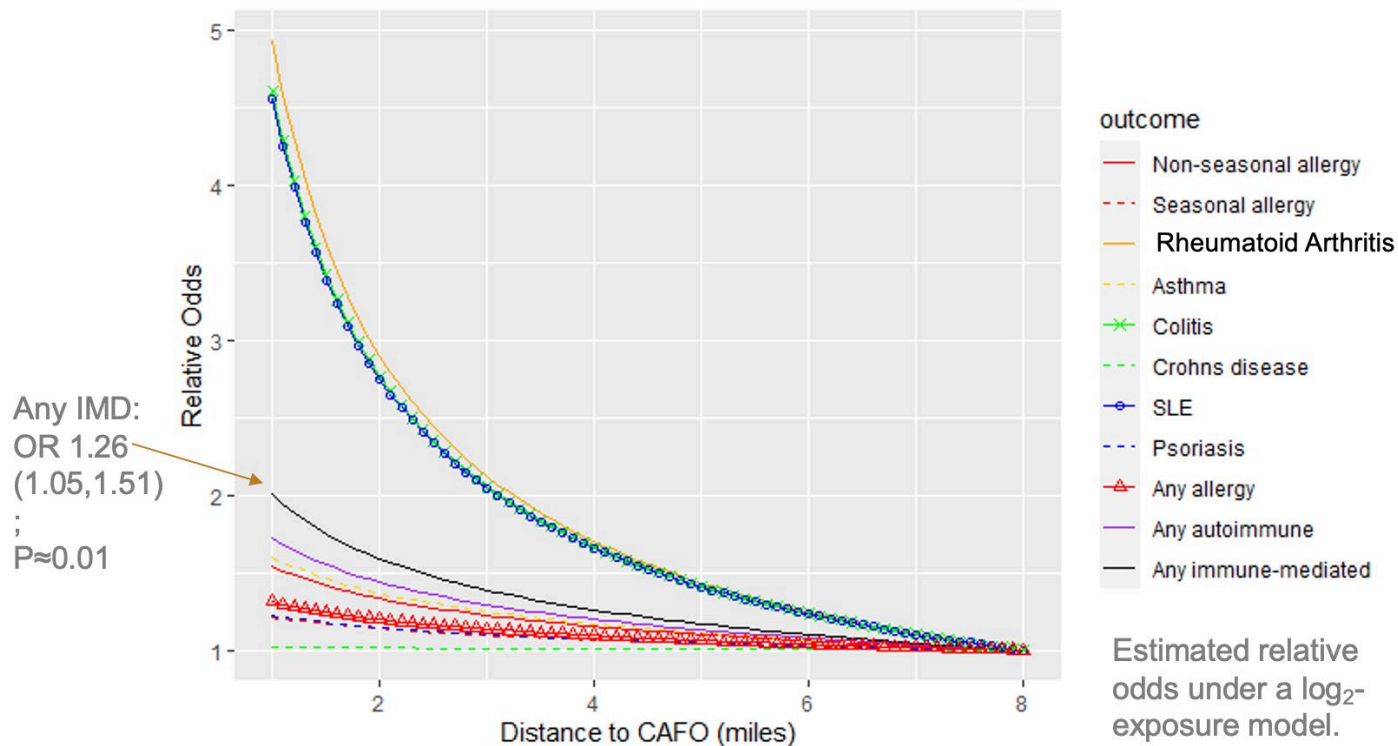
Immune Mediated Diseases

We hypothesize that proximity to swine CAFOs and IMD risk SNPs are associated with increased occurrence of immune-mediated diseases (IMD).

For array of IMD outcomes and mechanistic outcome groups we assessed:

1. The association between CAFO proximity and IMD

2. The association between IMD SNPs and IMD

3. Gene-environment interaction in IMD (GxE)[1]

[1]Interaction modeled according to Aliev et al (2014)

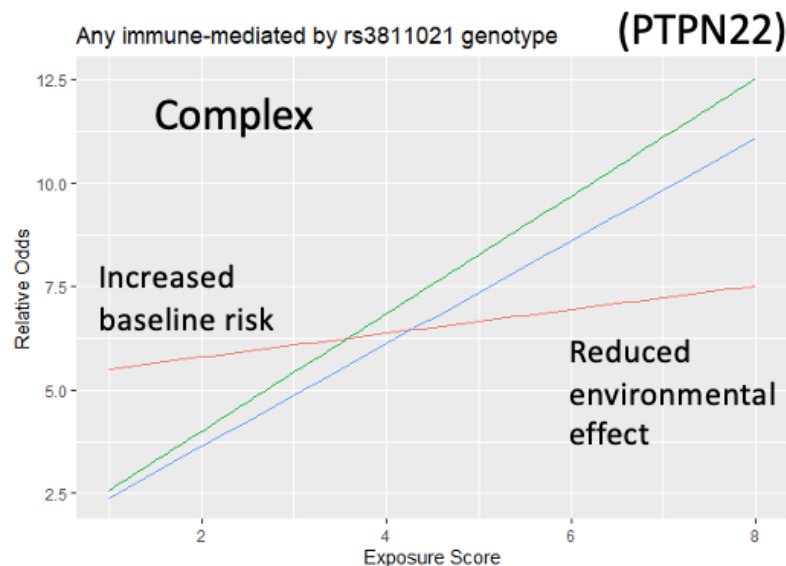# Association between CAFO Proximity and IMD

Any IMD:
OR 1.26
(1.05,1.51)
;
P≈0.01

outcome
- —— Non-seasonal allergy
- - - - Seasonal allergy
- —— Rheumatoid Arthritis
- - - - Asthma
- —✕— Colitis
- - - - Crohns disease
- —○— SLE
- - - - Psoriasis
- —△— Any allergy
- —— Any autoimmune
- —— Any immune-mediated

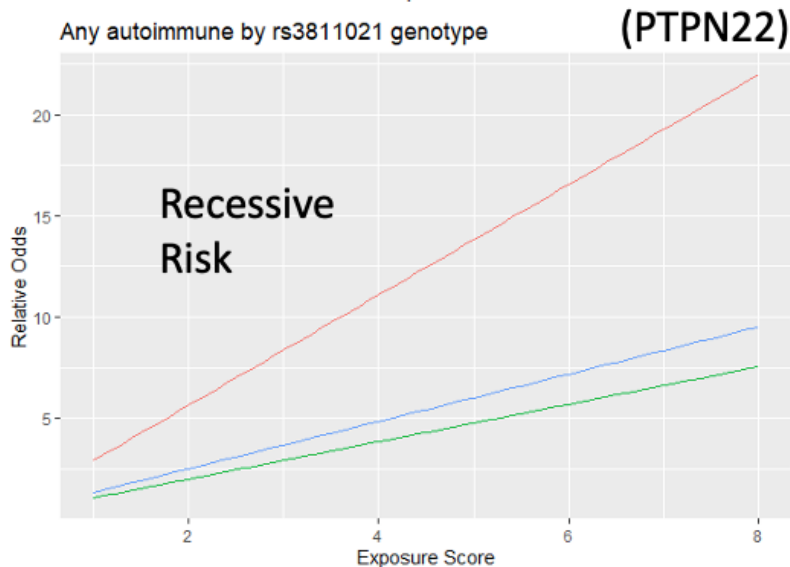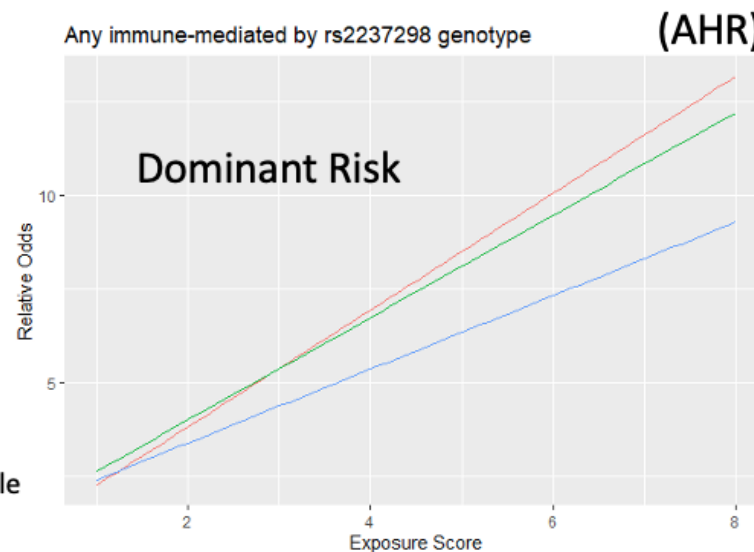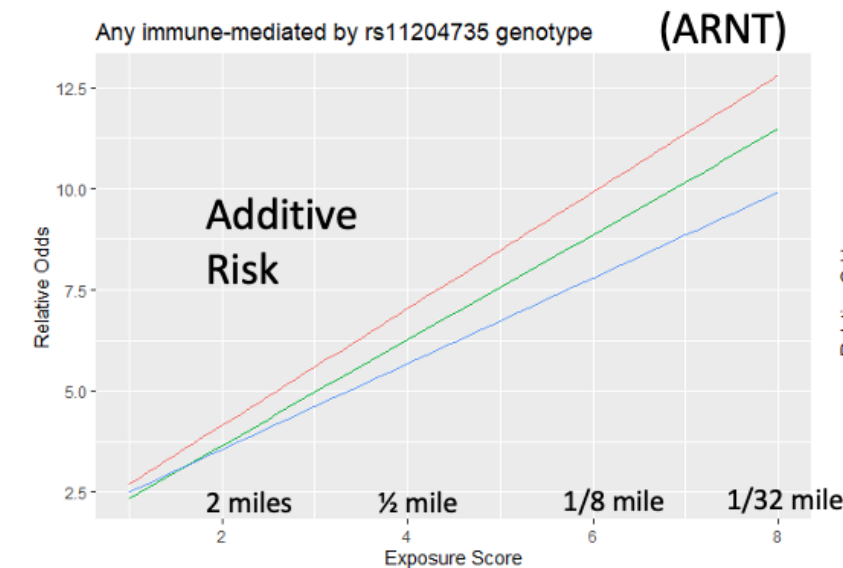Estimated relative odds under a $\log_2$-exposure model.

- Proximity to CAFOs was consistently associated with increased risk across IMD phenotypes.
- Significant associations were observed with rheumatoid arthritis, colitis, and any IMD (c=269).

Genetic associations with IMD

- Test for associations between SNPs previously associated with an IMD
  - AHRR, ARNT, PTPN22 and RSBN1

- rs2066853 (AHR) significantly associated with decreased odds of IMD, any allergy, and seasonal allergy

- Minor alleles increase disease, consistent across multiple IMD phenotypes
  - Suggests common mechanisms

Evidence for GXE Interactions

$$ES = -\log_2(d)+3$$

Any immune-mediated by rs11204735 genotype — (ARNT)

Additive Risk

2 miles   ½ mile   1/8 mile   1/32 mile

Any immune-mediated by rs2237298 genotype — (AHR)

Dominant Risk

Any autoimmune by rs3811021 genotype — (PTPN22)

Recessive Risk

Any immune-mediated by rs3811021 genotype — (PTPN22)

Complex

Increased baseline risk

Reduced environmental effect

Ongoing Analysis

Many projects to understand the etiology of disease, and find gene-environment interactions

- Questionnaire based Exposome-wide Associations Studies

- Polygenic risk score development and GXE

- Methods development for prioritizing SNPs for GXE analysis
  - Genelist.niehs.nih.gov

- Genome-wide association analysis

- Association of geospatial exposures with diseases

- Linkage to health records from Duke and UNC Chapel Hill

Potential Sources of Exposome Data

Biomarker measurements in biospecimens

- Blood

- Urine

- Plasma

- Saliva

- Feces

- Placenta

- Umbilical cord blood

- Breast milk

- Hair, nails

Potential Sources of Exposome Data
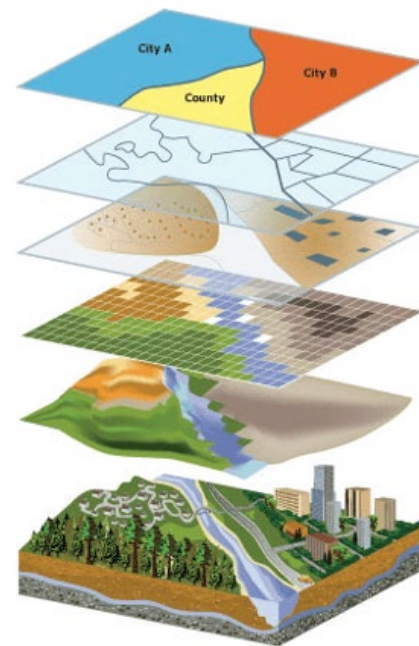
Questionnaire data

- Residential exposures

- Occupational exposures

- Socioecomic status

- Educational

- Diet

- Lifestyle

- Medications

Potential Sources of Exposome Data

Geospatial Exposures – linked to addresses

- Residential address

- Address history

- Occupational address(es)

- Childhood address(es)

Potential Sources of Exposome Data

Data resources with exposure models

Residential address

- Address history

- Occupational address(es)

- Childhood address(es)

Wearables and smart phones

Numerous exposure models and databases

- Weather and atmosphere

- Air pollution

- Groundwater contamination

- Distance to hazards

- Health disparities information

- Built environment

Challenges in data "munging"

- Different geographic extents

- Different data sources

- Different special resolutions

- Different temporal resolutions

- Not consistently structured

Challenges in data "munging"

- Different geographic extents

- Different data sources

- Different special resolutions

- Different temporal resolutions

- Not consistently structured

Data efforts at NIEHS: building software resources to empower assembling data resources

Pollutant	Geographic Extent	Type	Data Source	Spatial Resolution	Temporal Resolution	Link to data documentation
Aerosol Index	Worldwide	Orbital spectroscopy	Aura	2 km	Daily	https://earthdata.nasa.gov/earth-observation-data/near-real-time/hazards-and-disasters/air-quality
UV Aerosol Index	Worldwide	Orbital spectroscopy	Suomi	2 km	Daily	https://earthdata.nasa.gov/earth-observation-data/near-real-time/hazards-and-disasters/air-quality
Aerosol Optical Depth	Worldwide	Orbital spectroscopy	MODIS (Aqua/Terra)	1 km	Daily	https://earthdata.nasa.gov/earth-observation-data/near-real-time/hazards-and-disasters/air-quality
Carbon Monoxide	Worldwide	Orbital spectroscopy	Aqua/Aura/Terra	2 km	Day/Night	https://earthdata.nasa.gov/earth-observation-data/near-real-time/hazards-and-disasters/air-quality
Fires	Worldwide	Orbital spectroscopy	Aqua/Tetta/Suomi	1 km	Daily	https://earthdata.nasa.gov/earth-observation-data/near-real-time/hazards-and-disasters/air-quality
Nitrous Oxide	Worldwide	Orbital spectroscopy	Aura	5 km	Day/Night	https://earthdata.nasa.gov/earth-observation-data/near-real-time/hazards-and-disasters/air-quality
Ozone	Worldwide	Orbital spectroscopy	Aura / Suomi	2 km	Day/Night	https://earthdata.nasa.gov/earth-observation-data/near-real-time/hazards-and-disasters/air-quality
Sulfur Dioxide	Worldwide	Orbital spectroscopy	Aqua/Aura/Suomi	2km	Day/Night	https://earthdata.nasa.gov/earth-observation-data/near-real-time/hazards-and-disasters/air-quality
Land use	Worldwide	Orbital spectroscopy	MODIS (Aqua/Terra)	500m	6-month	https://modis.gsfc.nasa.gov/data/dataprod/mod12.php
Roadways	Worldwide	Survey/Administrative	Compilation	Long/Lat	~Decade	https://sedac.ciesin.columbia.edu/data/set/groads-global-roads-open-access-v1
Temperature	Worldwide	Meterology	Weather station network	5 degree grid	Monthly	https://www.ncdc.noaa.gov/data-access/marineocean-data/noaa-global-surface-temperature-noaaglobaltemp
Sulfur Dioxide	United States	Point emissions estimates	National Emissions Inventory	Long/Lat	3 Years	https://www.epa.gov/sites/production/files/2020-04/documents/nei2017_tsd_full_30apr2020.pdf
Volatile organic componds	United States	Point emissions estimates	National Emissions Inventory	Long/Lat	3 Years	https://www.epa.gov/sites/production/files/2020-04/documents/nei2017_tsd_full_30apr2020.pdf
Nitrogen oxides	United States	Point emissions estimates	National Emissions Inventory	Long/Lat	3 Years	https://www.epa.gov/sites/production/files/2020-04/documents/nei2017_tsd_full_30apr2020.pdf
Carbon dioxide	United States	Point emissions estimates	National Emissions Inventory	Long/Lat	3 Years	https://www.epa.gov/sites/production/files/2020-04/documents/nei2017_tsd_full_30apr2020.pdf
Lead	United States	Point emissions estimates	National Emissions Inventory	Long/Lat	3 Years	https://www.epa.gov/sites/production/files/2020-04/documents/nei2017_tsd_full_30apr2020.pdf
PM$_{2.5}$	United States	Point emissions estimates	National Emissions Inventory	Long/Lat	3 Years	https://www.epa.gov/sites/production/files/2020-04/documents/nei2017_tsd_full_30apr2020.pdf
PM$_{10}$	United States	Point emissions estimates	National Emissions Inventory	Long/Lat	3 Years	https://www.epa.gov/sites/production/files/2020-04/documents/nei2017_tsd_full_30apr2020.pdf
Ammonia	United States	Point emissions estimates	National Emissions Inventory	Long/Lat	3 Years	https://www.epa.gov/sites/production/files/2020-04/documents/nei2017_tsd_full_30apr2020.pdf
Carcinogens	United States	Point emissions estimates	Toxics Release Inventory	Long/Lat	Yearly	https://www.epa.gov/toxics-release-inventory-tri-program/tri-listed-chemicals
Acute respiratory toxicants	United States	Point emissions estimates	Toxics Release Inventory	Long/Lat	Yearly	https://www.epa.gov/toxics-release-inventory-tri-program/tri-listed-chemicals
Persistent organic pollutants	United States	Point emissions estimates	Toxics Release Inventory	Long/Lat	Yearly	https://www.epa.gov/toxics-release-inventory-tri-program/tri-listed-chemicals
Criteria air pollutants	United States	Fused Emissions Model	Synthesis of multiple datasets	Census Tract / BG	Yearly	https://www.caces.us/data
PM$_{2.5}$	United States	Downscaler model	EPA Air Quality System	Metro area	Daily	https://healthdata.gov/dataset/daily-census-tract-level-pm25-concentrations-2011-2014
PM$_{10}$	United States	Air monitor data	EPA Air Quality System	Metro area	Daily	https://www.epa.gov/aqs
Pesticides	United States	Survey/Land Use	USGS	County	Yearly	https://water.usgs.gov/nawqa/pnsp/usage/maps/
Roadway Density	United States	Survey/Land Use	USDOT	Census Block	~Decade	https://www.transportation.gov/mission/health/proximity-major-roadways#aboutthedata
Race/ethnicity	United States	Population survey	American Community Survey	Census Tract or ZIP	Yearly (rolling sample)	https://www.census.gov/programs-surveys/acs
Age/Sex	United States	Population survey	American Community Survey	Census Tract or ZIP	Yearly (rolling sample)	https://www.census.gov/programs-surveys/acs
Education/Income/Employment	United States	Population survey	American Community Survey	Census Tract or ZIP	Yearly (rolling sample)	https://www.census.gov/programs-surveys/acs
Poverty/public assistance	United States	Population survey	American Community Survey	Census Tract or ZIP	Yearly (rolling sample)	https://www.census.gov/programs-surveys/acs
Housing Indicators	United States	Population survey	American Community Survey	Census Tract or ZIP	Yearly (rolling sample)	https://www.census.gov/programs-surveys/acs
Social Vulnerability Index	United States	Derived Rank Measure	Derived from ACS Data	Census Tract	2 Years	https://svi.cdc.gov/data-and-tools-download.html
Quality of care	United States	Process indicator	Medicare/Medicaid Data	County	Yearly	https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityMeasures
Airports	United States	Point emissions sources	FAA	Long/lat	~Decade	https://catalog.data.gov/dataset/airports
Cellular Broadcast Towers	United States	Point emissions sources	FCC	Long/lat	~Decade	https://hifld-geoplatform.opendata.arcgis.com/datasets/0835ba2ed38f494196c14af8407454fb_0
Hazardous waste disposal sites	United States	Point emissions sources	EPA	Long/lat	~Decade	https://www.epa.gov/enviro/envirofacts-data-service-api
Nuclear	United States	Point emissions sources	EIA	Long/lat	~Decade	https://www.eia.gov/maps/layer_info-m.php
Power lines	United States	Point emissions sources	NGA	Long/lat	~Decade	https://hifld-geoplatform.opendata.arcgis.com/datasets/electric-power-transmission-lines
Railroads	United States	Point emissions sources	USDOT	Long/lat	~Decade	https://railroads.dot.gov/maps-and-data/maps-geographic-information-system/maps-geographic-information-system
Subtitle D landfills	United States	Point emissions sources	NGA	Long/lat	~Decade	https://hifld-geoplatform.opendata.arcgis.com/datasets/solid-waste-landfill-facilities
Superfund sites	United States	Point emissions sources	EPA	Long/lat	~Decade	https://catalog.data.gov/dataset/superfund-sites1e8f4
Wastewater treatment plants	United States	Point emissions sources	EPA	Long/lat	~Decade	https://hifld-geoplatform.opendata.arcgis.com/datasets/4b9bac25263047c19e617d7bd7b30701_0
Animal Feeding Operations	North Carolina	Point emissions sources	NCDEQ	Long/lat	Yearly	https://deq.nc.gov/about/divisions/water-resources/water-quality-regional-operations/afo
Dry cleaning solvent	North Carolina	Point emissions sources	NCDEQ	Long/lat	~Decade	https://deq.nc.gov/about/divisions/waste-management/dry-cleaning-solvent-cleanup-act-program
Pre-regulatory landfills	North Carolina	Point emissions sources	NCDEQ	Long/lat	~Decade	https://deq.nc.gov/about/divisions/waste-management/superfund-section/pre-regulatory-landfill-program
Hazardous waste spills	North Carolina	Point emissions sources	NCDEQ	Long/lat	~Decade	

What types of exposure would be the most relevant?

- Lifestyle

- Built environment

- Weather/atmosphere

- Stress

- Toxic or chemical exposures

- Diet and medications

- Social determinants of health

- Others?

NIEHS Acknowledgments

- PEGS Leadership
  - Janet Hall (co-PI)
  - David Fargo
  - Charles Schmitt

- GIS Modeling
  - Kyle Messier
  - Melissa Lowe
  - Kristin Eccles

- CAFOs Study
  - Fred Miller
  - Nat McNell
  - Shepherd Sherman

- Statistical Genetics
  - Farida Akhtari
  - John House
  - Eunice Lee
  - Xiaron Tong
  - Dillon Lloyd
  - Jasmine Mack

- Bioinformatics
  - Adam Burkholder
  - Jianying Li

- Data Scholar
  - Lara Clark

- Common Disease Working Group
  - Kelly Chandler
  - Kim McAllister
  - Aubrey Miller
  - Alex Merrick
  - Alison Harrill
  - David Balshaw
  - Yuxia Cui
  - Michelle Bennett
  - Cindy Lawler

Questions?

alison.motsinger-reif@nih.gov