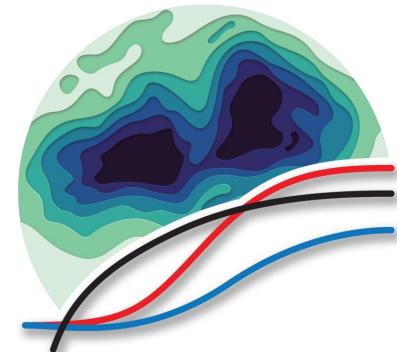


# **GeoTox: Integrating Geospatial Exposure Modeling and Adverse Outcome Pathways Towards Mechanistically-Informed Risk**

Kyle P Messier, PhD

Stadtman Tenure-Track Investigator  
NIEHS, NIMHD



Spatiotemporal Exposures  
and Toxicology (SET) Group

# Spatiotemporal Exposures and Toxicology Group



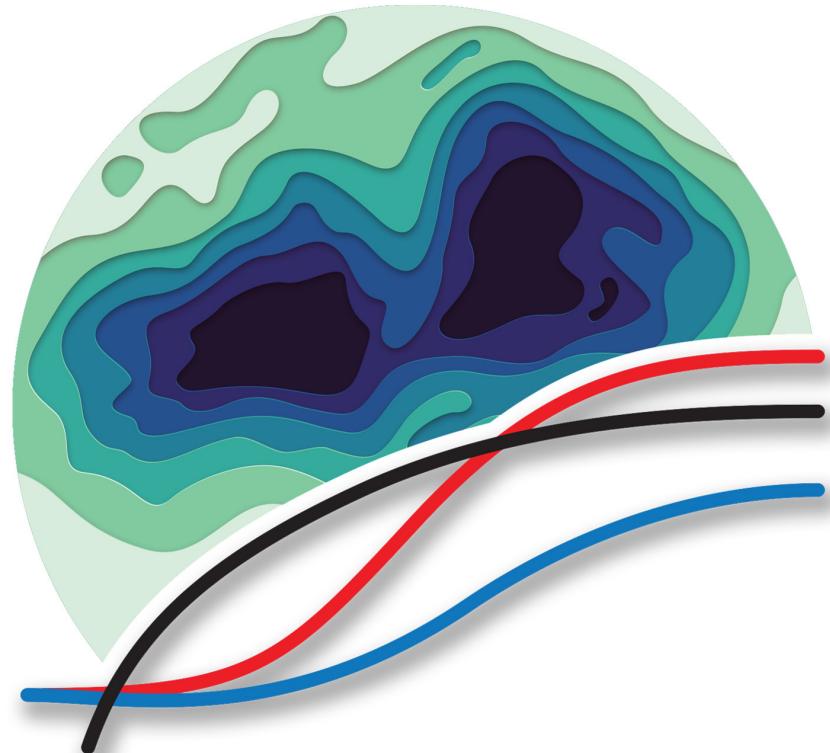
Spatiotemporal Exposure  
Mapping



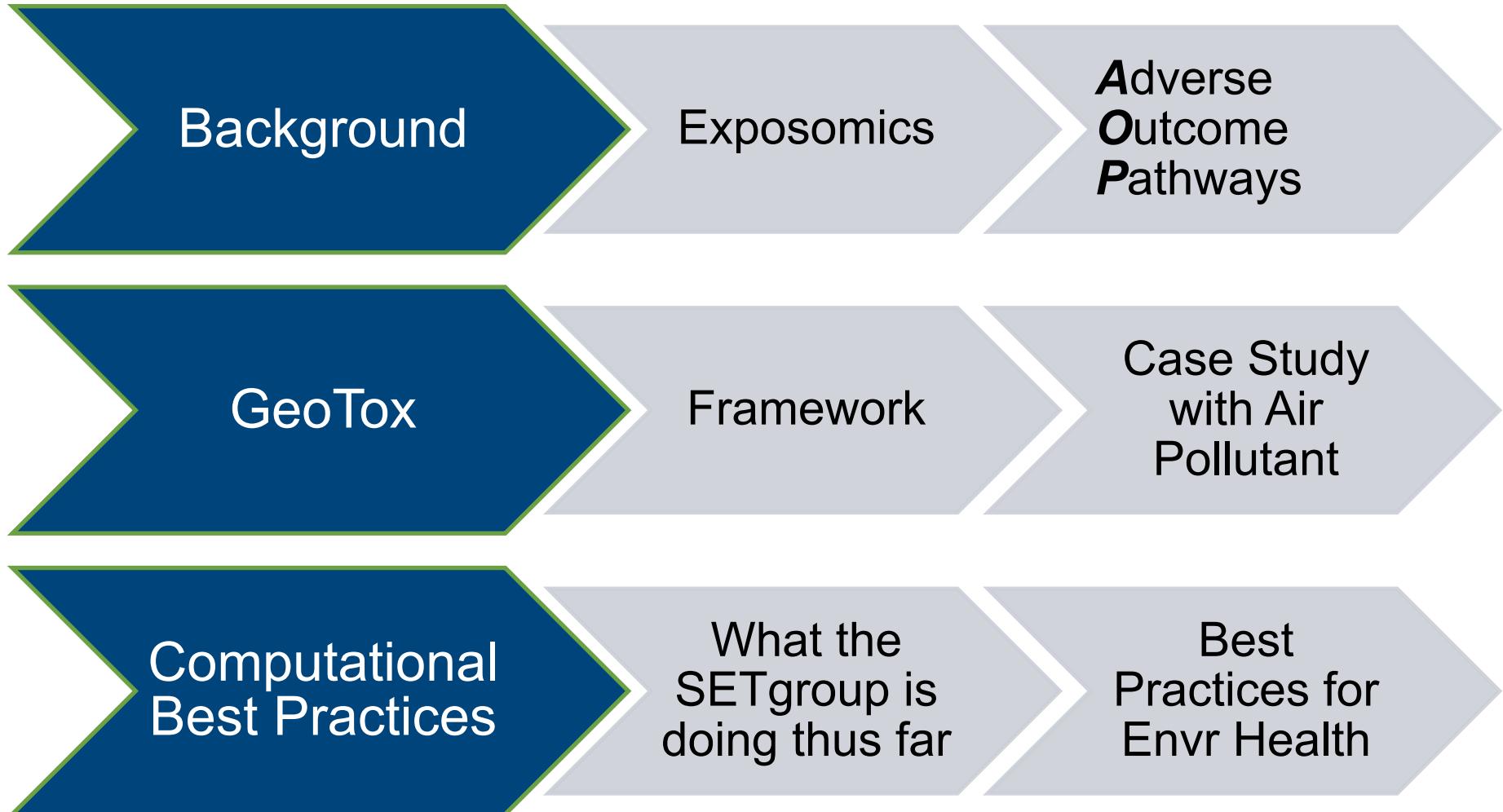
Chemical and Stressor  
Mixtures Prediction



Mechanistically Informed  
Risk Assessment



# Outline



# Background

## **Exposomics, Adverse Outcome Pathways, & GeoTox**

# Exposomics

## Ecosystems

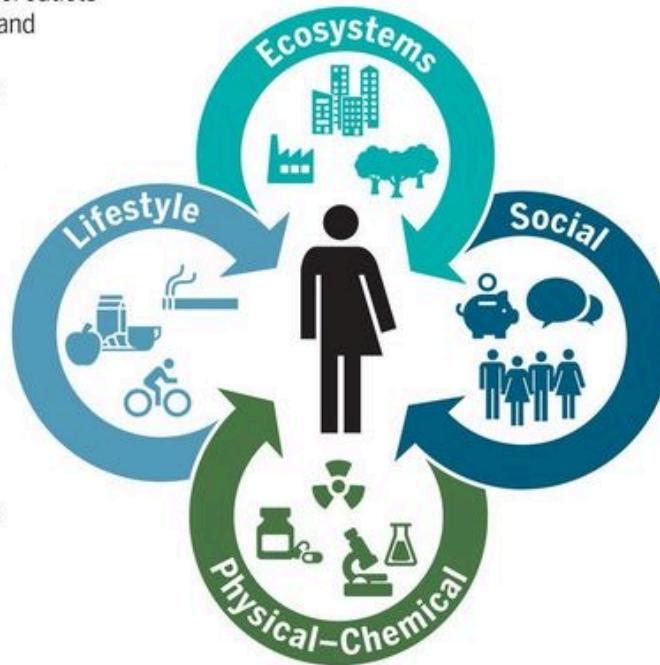
Food outlets, alcohol outlets  
Built environment and urban land uses  
Population density  
Walkability  
Green/blue space

## Lifestyle

Physical activity  
Sleep behavior  
Diet  
Drug use  
Smoking  
Alcohol use

## Social

Household income  
Inequality  
Social capital  
Social networks  
Cultural norms  
Cultural capital  
Psychological and mental stress



## Physical-Chemical

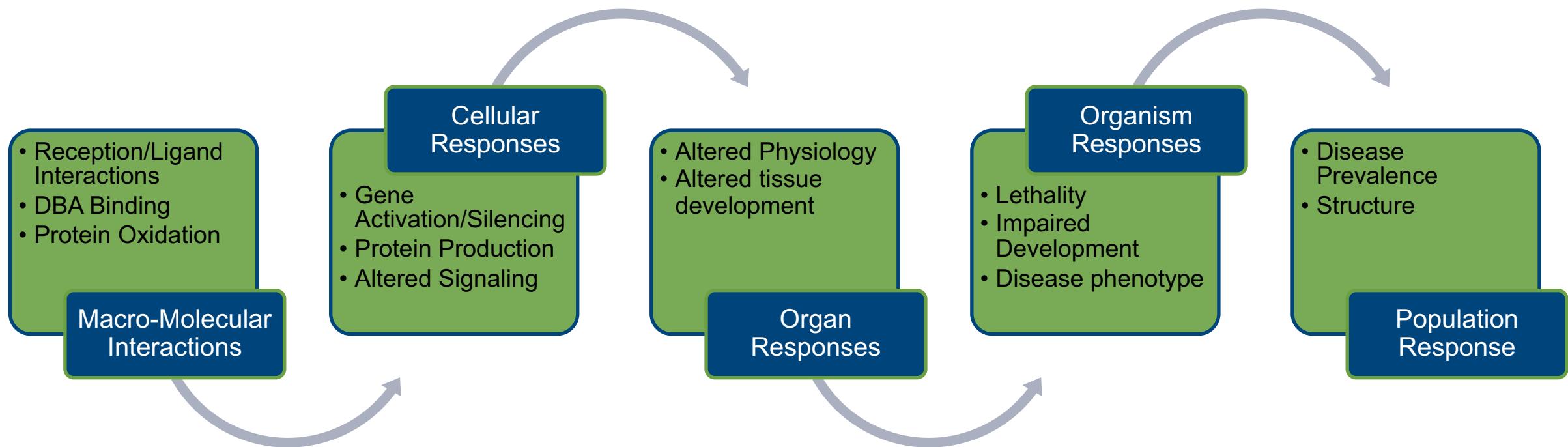
Temperature/humidity  
Electromagnetic fields  
Ambient light  
Odor and noise  
Point, line sources, e.g., factories, ports  
Outdoor and indoor air pollution  
Agricultural activities, livestock  
Pollen/mold/fungus  
Pesticides  
Fragrance products  
Flame retardants (PBDEs)  
Persistent organic pollutants  
Plastic and plasticizers  
Food contaminants  
Soil contaminants  
Drinking water contamination  
Groundwater contamination  
Surface water contamination  
Occupational exposures

The exposome and health: Where chemistry meets biology, Volume: 367, Issue: 6476, Pages: 392-396, DOI: (10.1126/science.aay3164)

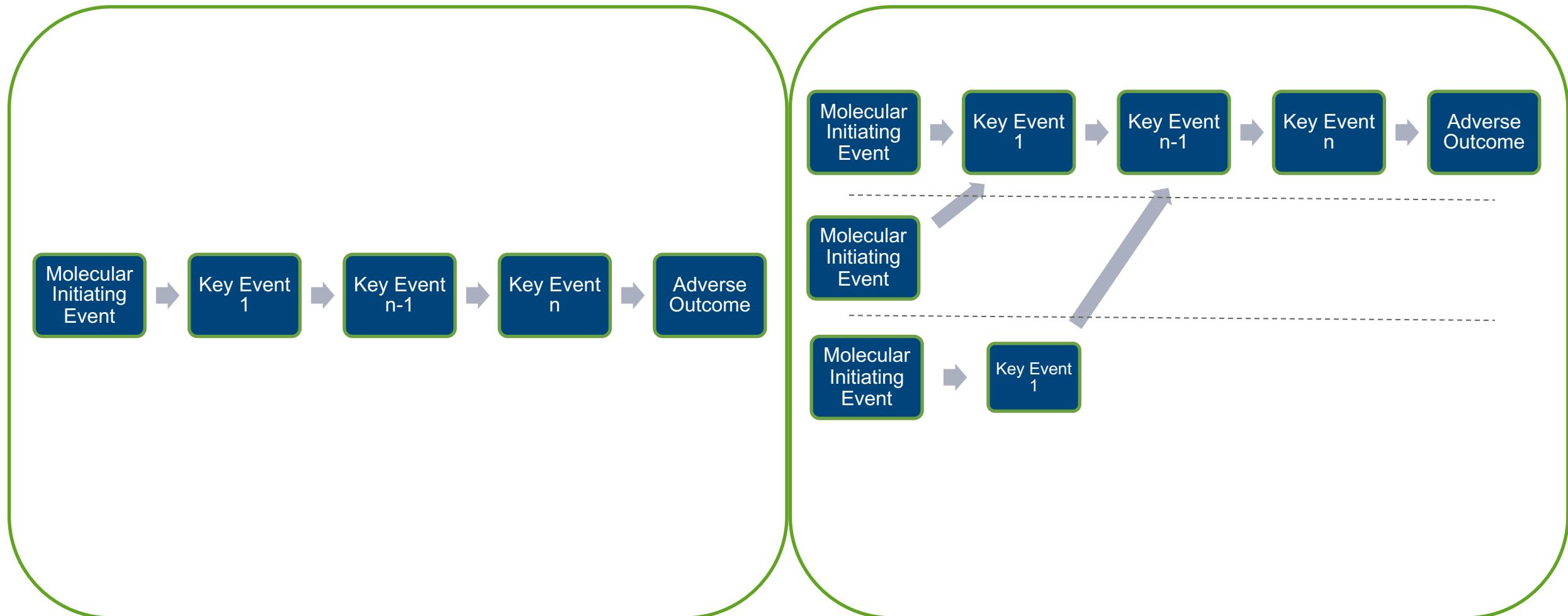
Science

AAAS

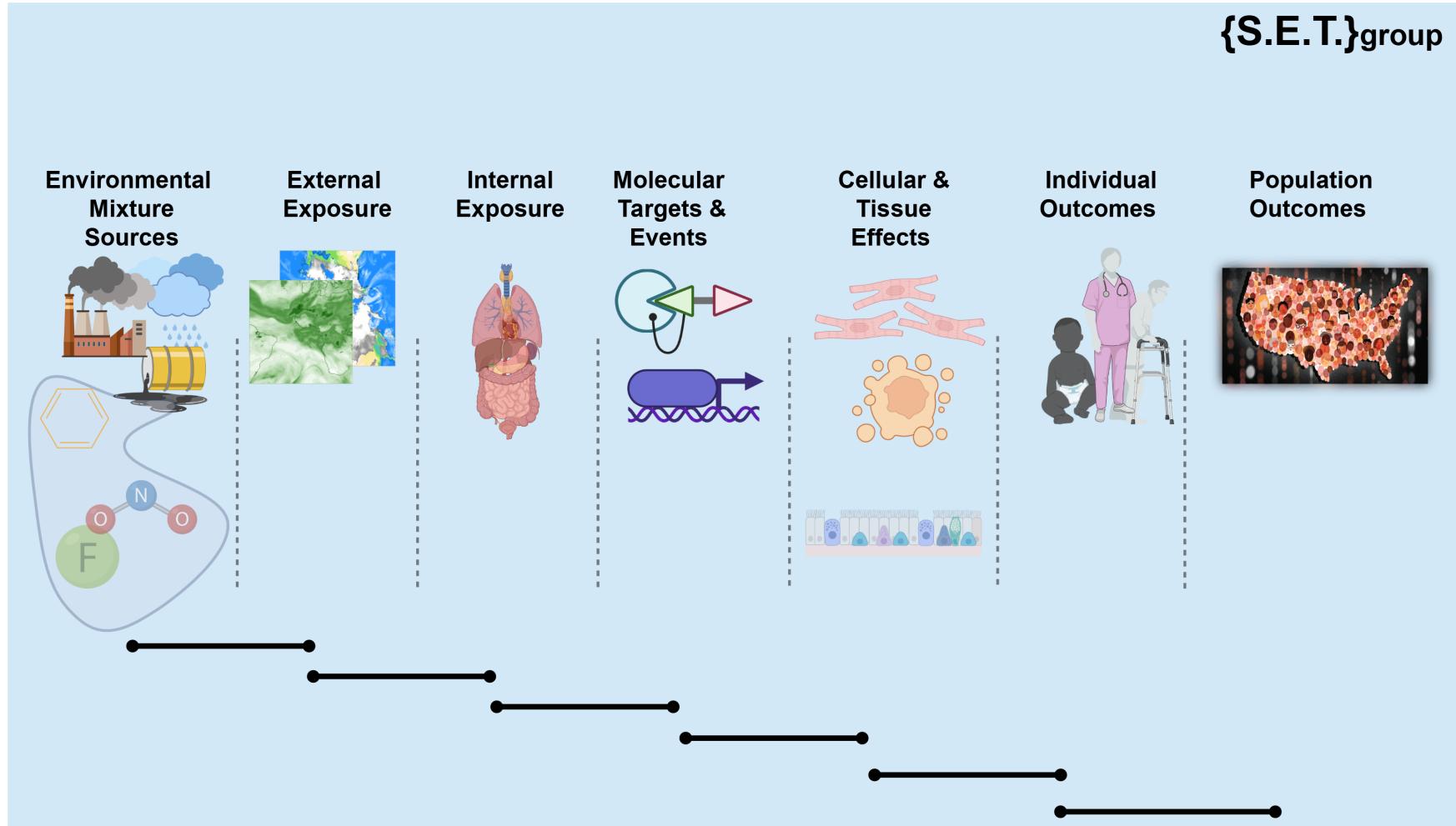
# Adverse Outcome Pathway



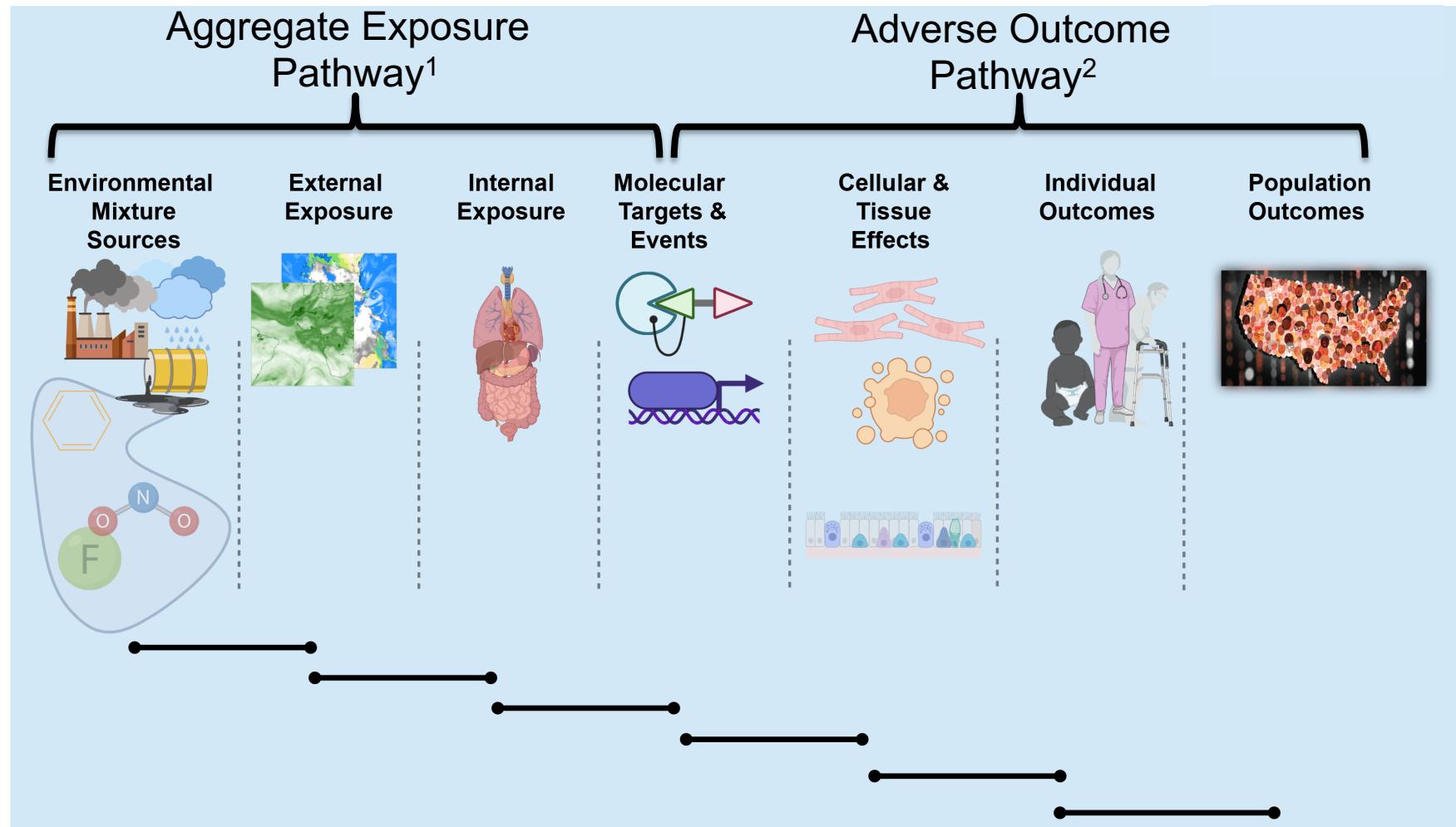
# Adverse Outcome Pathways



{S.E.T.}group



A Cascade of Events: The Events MUST Occur In This Order



1. Teeguarden JG, Tan YM, Edwards SW, Leonard JA, Anderson KA, Corley RA, Kile ML, Simonich SM, Stone D, Tanguay RL, Waters KM. Completing the link between exposure science and toxicology for improved environmental health decision making: the aggregate exposure pathway framework.

2. <http://aop.wiki.org>; Society for the Advancement of Adverse Outcome Pathways

## AEP + AOP = **GeoTox**

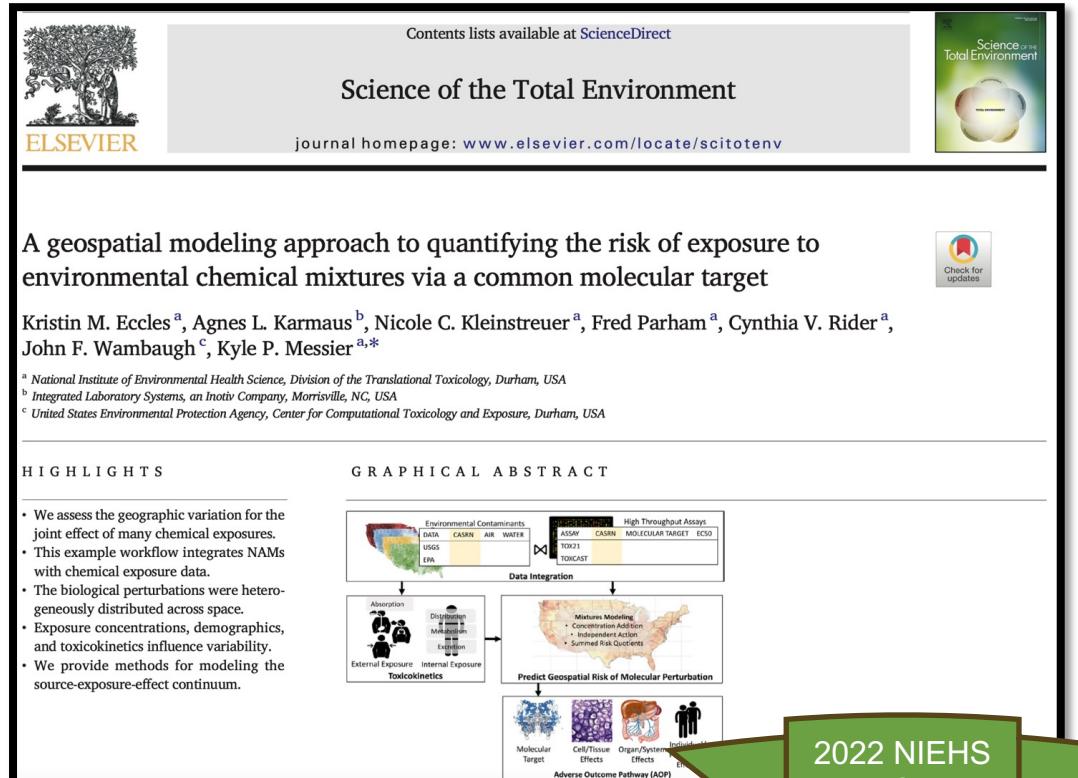
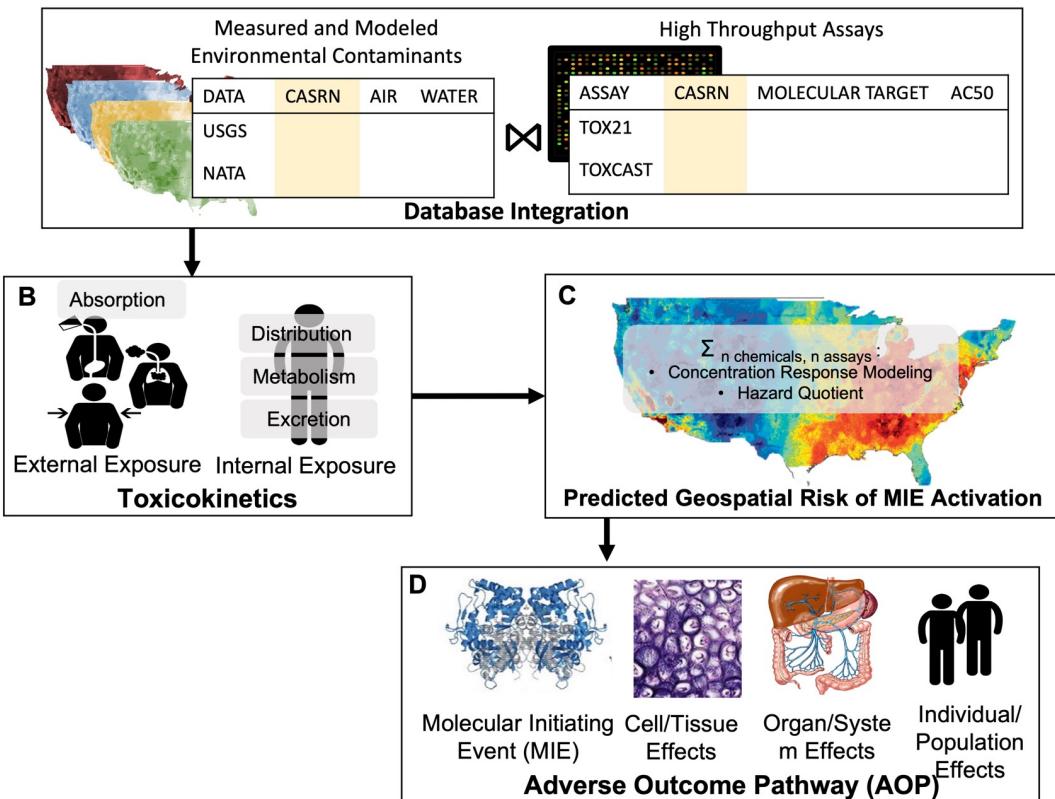
AEP is a comprehensive external analysis of source, media, and transformations

AOPs provide a linkage specific biological target, pathway or process by a stressor and an adverse outcome(s) considered relevant to risk assessment

# GeoTox

**Integrating geospatial exposure modeling with mechanistic toxicology data**

# GeoTox Proof of Concept



2022 NIEHS  
POY

# Key Steps in GeoTox Risk Mapping

## External Exposure

- Geospatial models
- Gaussian process, land-use regression, chemical transport

## Internal Exposure

- Inhalation, Ingestion, Dermal absorption

## In vitro equivalent dose

- Physiologically Based Toxicokinetic Models
- In-vitro to In-vivo Extrapolation (IVIVE)

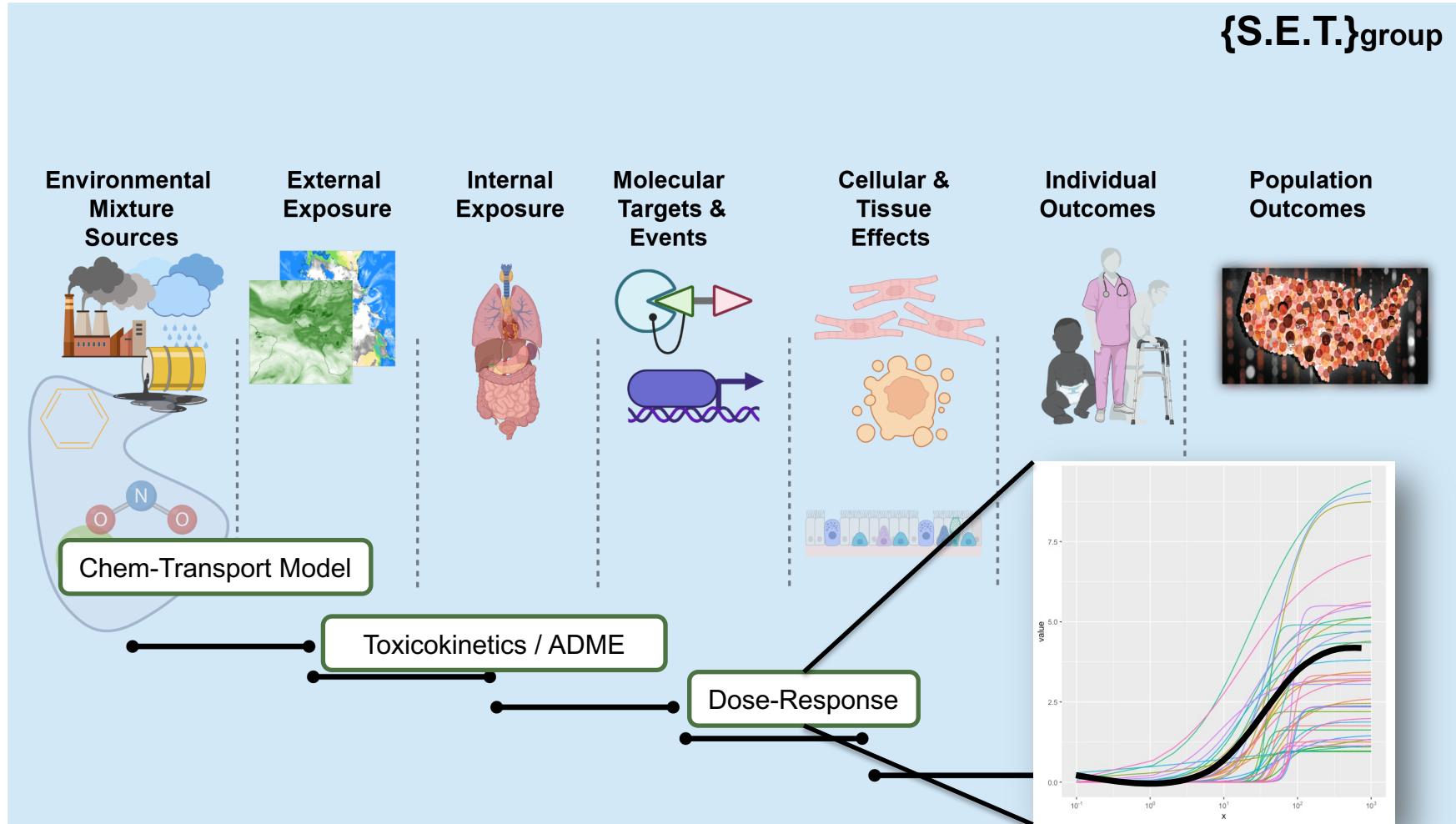
## Dose Response Modeling

- Toxicological Mixtures Predictions
- Generalized Concentration Addition, Independent Action

## Adverse Outcome Pathways (AOP)

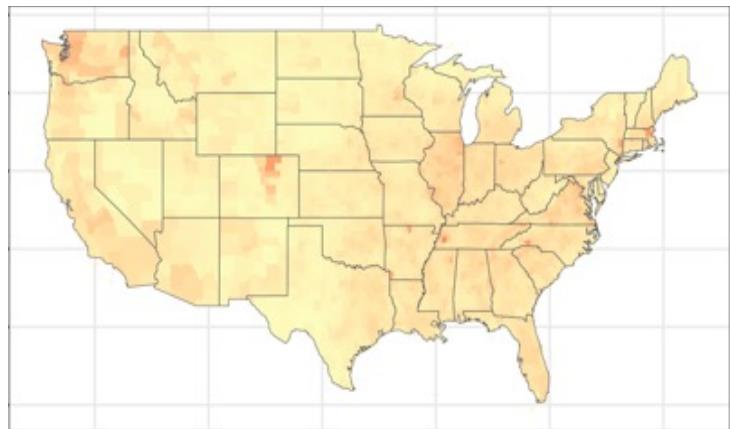
- Integration into AOP(s)

## {S.E.T.}group

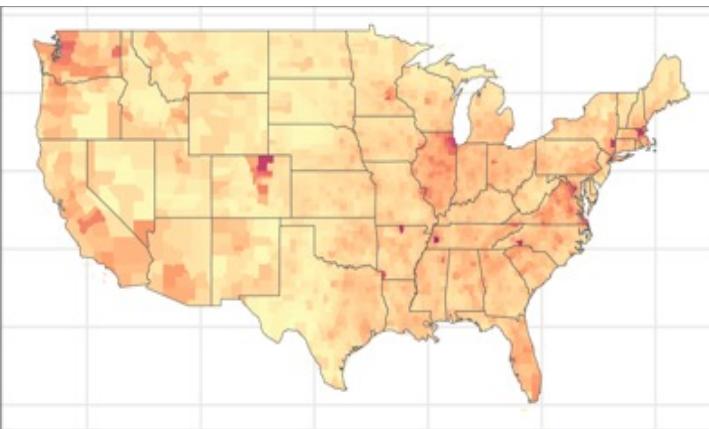


# Mapped Risk of Molecular Perturbation

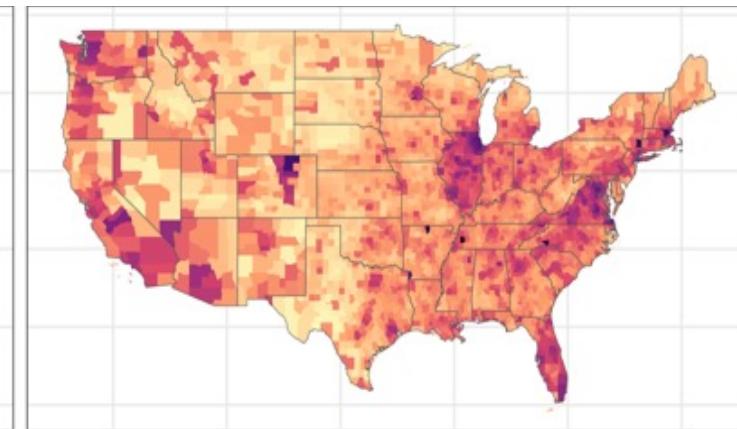
5th Percentile



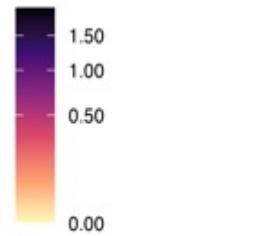
Median



95th Percentile



Predicted Response  
Log2 Fold Change  
mRNA Expression



1.50
1.00
0.50
0.00

## ***GeoTox and Climate Risk: Near Term Possibilities***

### Mapping of AOP key events

- Eccles et al. 2022

### Increased precision in exposure assessment

- Geospatial approximation of molecular epidemiological methods

### Bringing precision environmental medicine to a larger scale

- Spatial epidemiology exposures for cohort studies can be more sensitive if information is available (e.g. WGS, RNA-seq) to differentiate ADME within individuals

## Long-Term Possibilities

Geospatial risk assessment for chemicals or stressors without epidemiology data

- # Chemical D-R via NAM >> # Chemical D-R via epidemiology

Combining multiple AOPs and multiple routes of exposures

- Comprehensive exposomic risk assessment

## Some of the Current Limitations of GeoTox



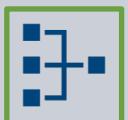
Infinite Mixtures Problem



Geospatial chemical exposure models have a “lamp-post” problem



Uncertainty in each step of the framework



Pathways are complex!

# Software and Computational Best-Practices

**Towards Open, Available, Reproducible, Extensible, and Tested  
Geospatial Models**

# Tidymodels as motivation

**Tidymodels packages**

**INSTALLATION AND USE**

- Install many of the packages in the tidymodels ecosystem by running `install.packages("tidymodels")`.
- Run `library(tidymodels)` to load the core packages and make them available in your current R session.

**CORE TIDYMODELS**

The core tidymodels packages work together to enable a wide variety of modeling approaches:



- tidymodels**  
tidymodels is a meta-package that installs and loads the core packages listed below that you need for modeling and machine learning. [Go to package ...](#)
- rsample**  
rsample provides infrastructure for efficient data splitting and resampling. [Go to package ...](#)
- parsnip**  
parsnip is a tidy, unified interface to models that can be used to try a range of models without getting bogged down in the syntactical minutiae of the underlying packages. [Go to package ...](#)
- recipes**  
recipes is a tidy interface to data pre-processing tools for feature engineering. [Go to package ...](#)
- workflows**  
workflows bundle your pre-processing, modeling, and post-processing together. [Go to package ...](#)
- tune**  
tune helps you optimize the hyperparameters of your model and pre-processing steps. [Go to package ...](#)
- yardstick**  
yardstick measures the effectiveness of models using performance metrics. [Go to package ...](#)

Daniel Zilber

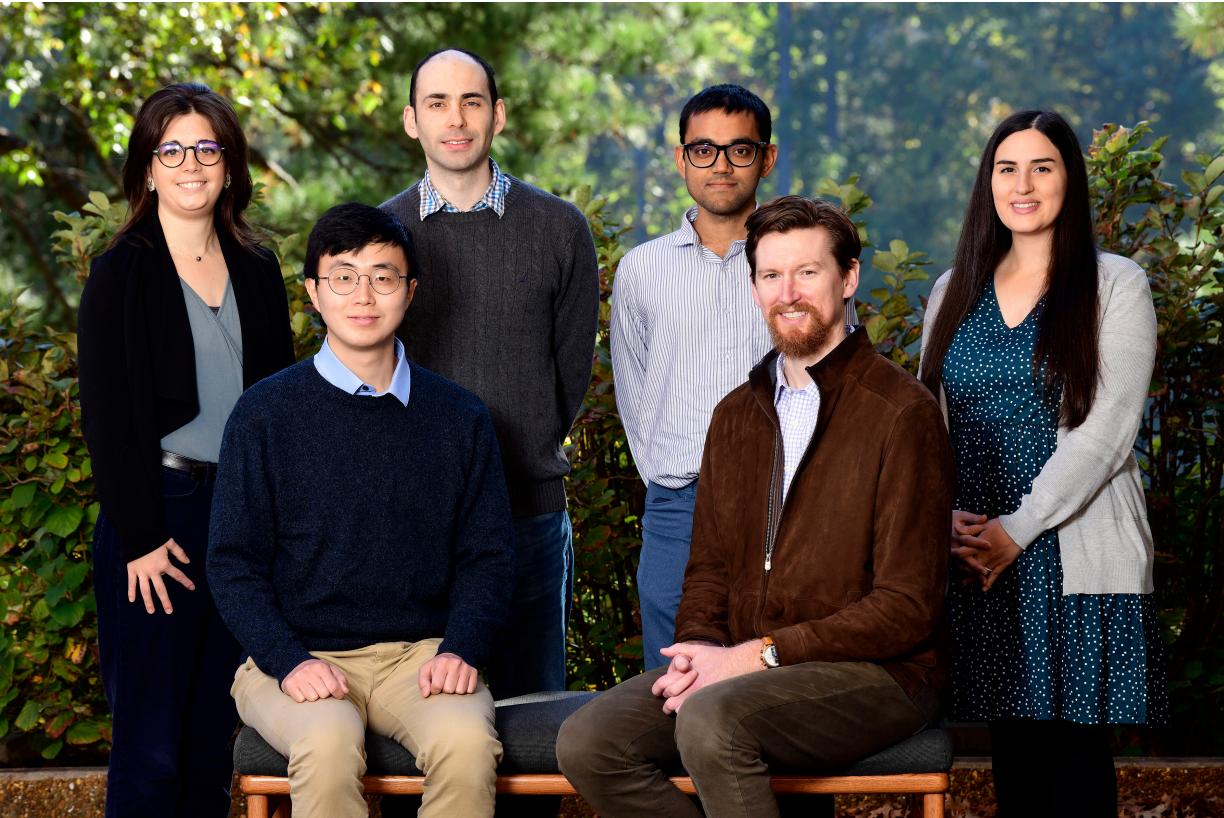
Ranadeep Daw

Eva Marques

Mariana Alifa

Insang Song

Kyle Messier



Mitchell Manware  
(Data Analyst)

## Software and Computational Best-Practices

Test Driven Development

Continuous Integration

Build Checks

Style / Linting

Workflows / Pipelines

# Test Driven Development: Tests and Coverage



Coverage Flags Commits Pulls Settings

NRTAPmodel / R / calculate\_covariates.R

```

28 #' @export
29 calc_koppen_geiger <-
30   function(
31     path_koppen = "./input/koppen_geiger/raw/Beck_KG_V1_present_0p0083.tif",
32     sites,
33     id_col = "site_id") {
34   ## You will get "sites" in memory after sourcing the file above
35   kg_rast <- terra::rast(path_koppen)
36   sites_tr <- sites
37
38   if (!methods::is(sites, "SpatVector")) {
39     sites_tr <- terra::vect(sites)
40   }
41   sites_kg <- terra::project(sites_tr, terra::crs(kg_rast))
42   sites_kg_extract <- terra::extract(kg_rast, sites_kg)
43
44   # The starting value is NA as the color table has 0 value in it
45   kg_class <-
46   c(
47     NA, "Af", "Am", "Aw", "BWh", "BWk", "BSh", "BSk", "Csa", "Csb",
48     "Csc", "Cwa", "Cwb", "Cwc", "Cfa", "Cfb", "Cfc",
49     "Dsa", "Dsb", "Dsc", "Dsd", "Dws", "Dwb", "Dwc", "Dwd",
50     "Dfa", "Dfb", "Dfc", "Dfd", "Ett", "Ef"
51   )
52   kg_coltab <- terra::coltab(kg_rast)
53   kg_coltab <- kg_coltab[1:1][seq(1, 31), ]
54   kg_colclass <- data.frame(
55     value = kg_coltab$value,
56     class_kg = kg_class
57   )
58
59   sites_kg_extract[[id_col]] <- unlist(sites_kg[[id_col]])
60   colnames(sites_kg_extract)[2] <- "value"
61   sites_kg_extract_e <- merge(sites_kg_extract, kg_colclass, by = "value")
62 
```

- **Unit tests:** Test a function or very specific piece of code e.g.
  - Does my function return an error when the input is wrong?
  - Does function return the expected value given a set of inputs?
  - `expect_*`
- **Integration tests:** Tests the input from function X works with function Y
- See *testthat* R package



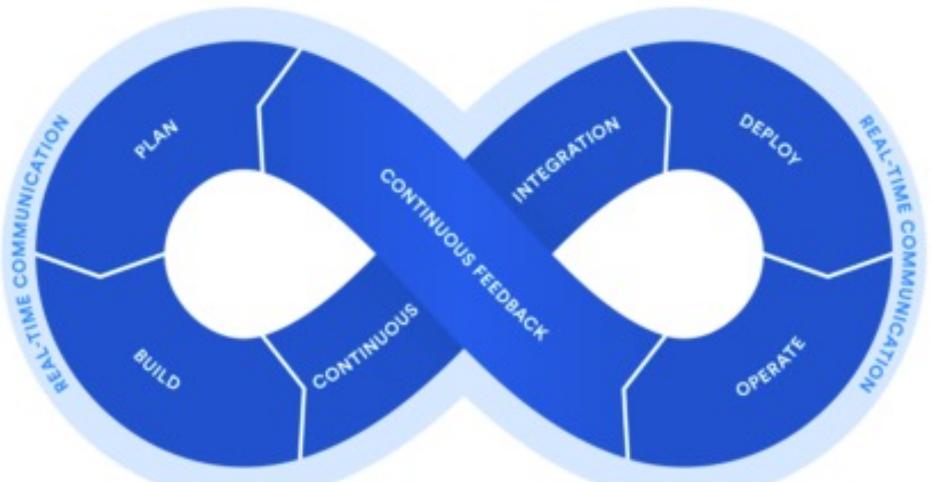
# Test Driven Development: The Philosophy



School of Athens, Raphael (1509–1511)

- The **Philosophy** of Test-Driven Development (TDD)
  1. Write a test
  2. Run the test → Expect it to **fail**
  3. Write the minimum code
  4. Run the test → Till it **passes**
  5. Refactor → Repeat → Complete
- Forces you to think about code requirements
- Forces explicit testing of code
- Minimizes bugs

# Continuous Integration



Frequent code commits



Detects errors sooner



Reduces complexity of  
debugging



Run code tests automatically



Makes working in groups  
easier

## Build Check / R CMD check

Checks for major errors, minor errors, and recommendations

- **ERROR:** Severe problems that you should fix regardless of whether or not you're submitting to CRAN.
- **WARNING:** Likely problems that you must fix if you're planning to submit to CRAN (and a good idea to look into even if you're not)<sup>1</sup>.
- **NOTE:** Mild problems or, in a few cases, just an observation.

Otherwise, you can learn  
about errors in the bad way!

<sup>1</sup><https://r-pkgs.org/>

## Style and Linting

Consistent style → readability and\_maintainability

Identifying syn.tax errors;

Semantic Errors

# Pipelines and Workflows

A pipeline is a computational workflow that does statistics, analytics, or data science

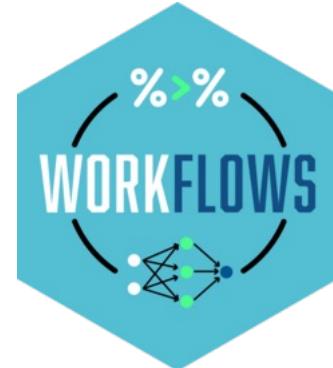
Coordinate analysis with a DAG like structure

Skip steps that are not needed with updates

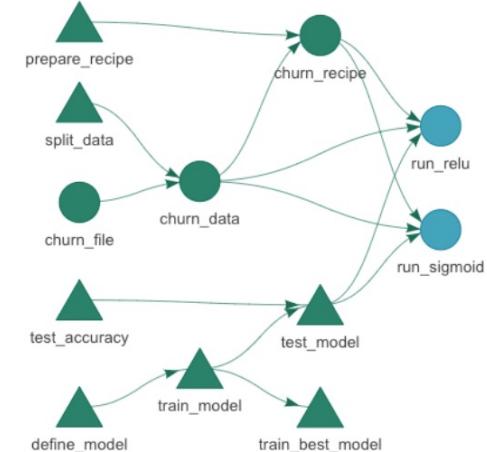
Facilitates code and package integration



`tar_visnetwork()`



Up to date  
Outdated  
Stem  
Object  
Function



# Best Practices are Needed for Complex Environmental Health Pipelines



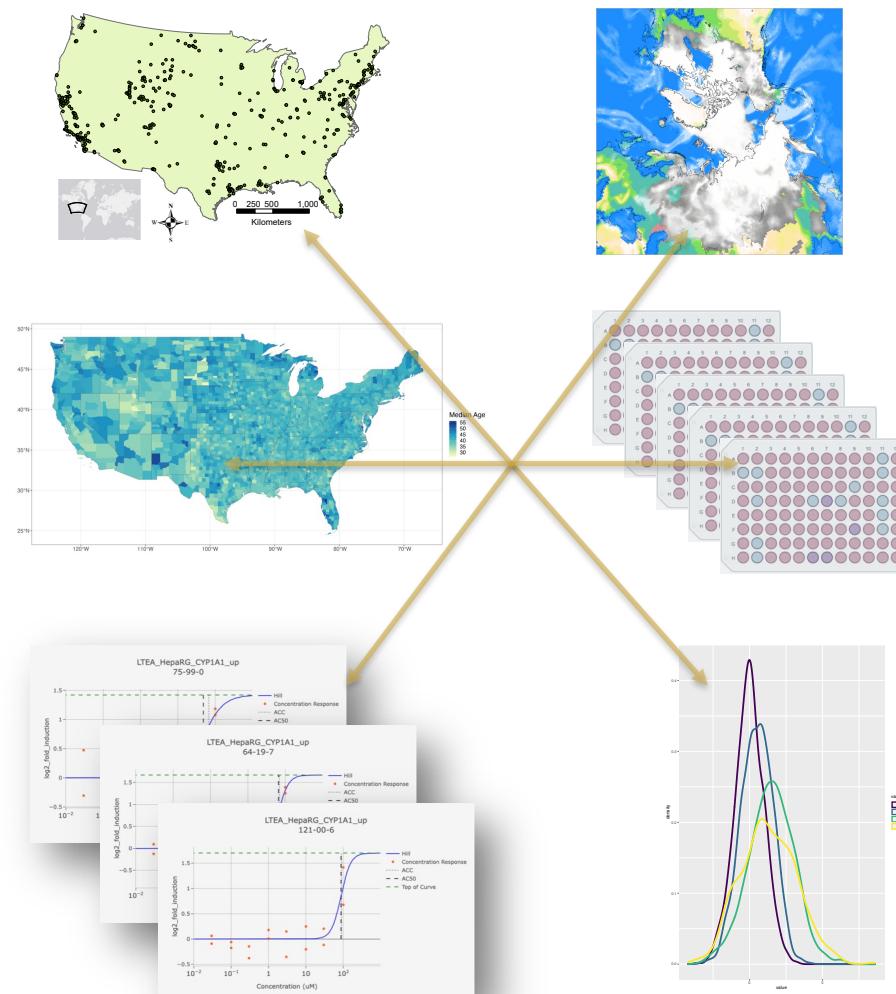
Geospatial modeling and GeoTox offer a potential tractable approach for quantifying the exposome health impacts



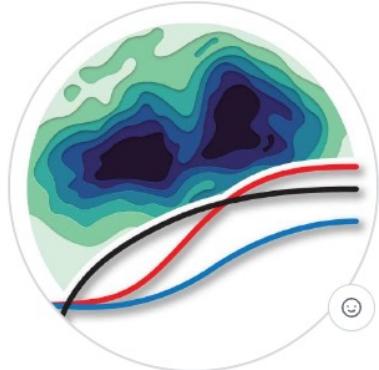
Need best practices to build towards a very complex analysis and understanding



(1) In-Situ Monitoring Data (2) Atmospheric / Geophysical Data (3) Census Data (4) In-Vitro Screening Data (5) Concentration-Response Modeling (6) Probabilistic Models (7) ...



# Documented, Tested, and Open



## {SET}group

Spatiotemporal-Exposures-and-Toxicology

Kyle P Messier, PhD Stadtman  
Investigator -- Geospatial exposure and  
risk assessment methods with tox data  
integration. He/Him @NIEHS

[Edit profile](#)

5 followers · 2 following

National Institute of Environmental Health Sciences  
Research Triangle Park, North Carolina  
13:08 (UTC -04:00)  
<https://www.niehs.nih.gov/research/atniehs/labs/ptb/spatiotemporal/index.cfm>

Spatiotemporal-Exposures-and-Toxicology / README.md

## { Spatiotemporal Exposures and Toxicology }

Github for open-source code and projects from [{SET}](#).

### Methods Used

- Spatial and Spatiotemporal Statistics
  - Gaussian processes
  - Penalized Regression
- Geographic Information Systems
- Land Use Regression
- Artificial Neural Networks

### Software We Use

- R, RMarkdown, RShiny
- Julia

[README](#) [License](#) [MIT license](#)

test-coverage passing codecov 91% R-CMD-check passing lint passing lifecycle experimental

## Air Pollution Data for the Masses: An Open-Access, Test-Driven, and Reproducible Pipeline PM<sub>2.5</sub> Hybrid Model

Group Project for the Spatiotemporal Exposures and Toxicology group with help from friends 😊 🎉 🌎

# GeoTox R package

- R package expected by Society of Toxicology meeting (March 2024)
- Increase accessibility and extensibility of GeoTox
- Improve computational speed

[README](#) [License](#) [MIT license](#)

---

**GeoTox**

The GeoToxPackage can , as introduced in [Eccles KM, Karmaus AL, Kleinstreuer NC, Parham F, Rider CV, Wambaugh JF, Messier KP. A geospatial modeling approach to quantifying the risk of exposure to environmental chemical mixtures via a common molecular target. Sci Total Environ. 2023 Jan 10;855:158905. doi: 10.1016/j.scitotenv.2022.158905. Epub 2022 Sep 21. PMID: 36152849; PMCID: PMC9979101.](#)

**GeoToxPackage** 0.0.0.9000 Reference Articles ▾

**GeoToxPackage**

The GeoToxPackage can , as introduced in [Eccles KM, Karmaus AL, Kleinstreuer NC, Parham F, Rider CV, Wambaugh JF, Messier KP. A geospatial modeling approach to quantifying the risk of exposure to environmental chemical mixtures via a common molecular target. Sci Total Environ. 2023 Jan 10;855:158905. doi: 10.1016/j.scitotenv.2022.158905. Epub 2022 Sep 21. PMID: 36152849; PMCID: PMC9979101.](#)

**Installation**  
 You can install the development version of GeoToxPackage from [GitHub](#) with:

```
# install.packages("devtools")
devtools::install_github("Spatiotemporal-Exposures-and-Toxicology/GeoToxPackage")
```

**Example**  

```
library(GeoToxPackage)
library(dplyr, warn.conflicts = FALSE)
```

**Estimate chemical concentration-response curves**  

```
conc <- 10^rep(-2:2, each = 3)
tp <- 100 # top asymptote
ga <- 1.6 # AC50
gw <- 1.2 # slope
resp <- tp / (1 + (ga / conc)^gw) + rnorm(length(conc), sd = 5)

fit_2param <- fit_hill(log10(conc), resp) # slope fixed at 1
fit_3param <- fit_hill(log10(conc), resp, fixed_slope = FALSE)

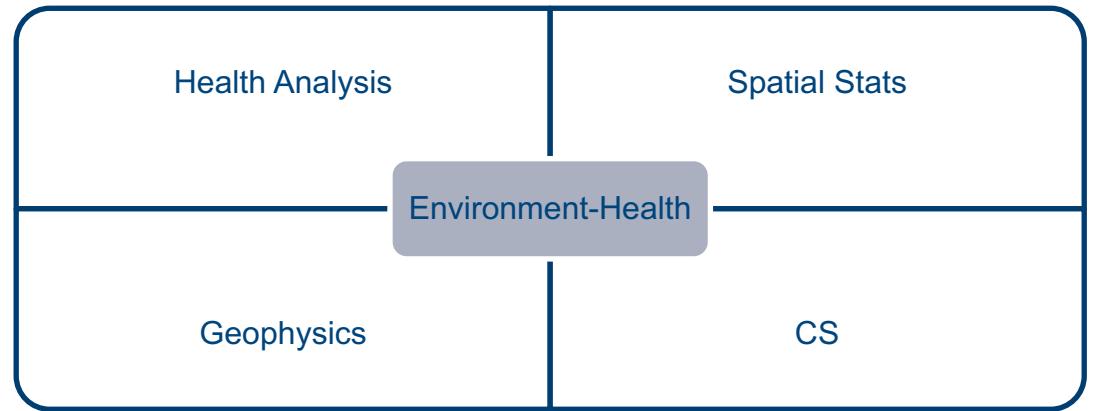
rbind(
  "inputs" = c(tp, log10(ga), gw, NA),
  "3-param" = c(fit_3param$par),
  "2-param" = c(fit_2param$par[1:2], 1, fit_2param$par[3])
)
#>          tp logAC50 slope t-error
#> inputs 100.00000 0.2041200 1.2000000      NA
#> 3-param 97.63578 0.1516374 1.485236 1.177891
#> 2-param 101.53494 0.2360917 1.000000 1.417567
```

**License**  
[Full license](#)  
[MIT + file LICENSE](#)

**Citation**  
[Citing GeoToxPackage](#)

**Developers**  
 Skylar Marvel  
 Author, contributor  
 David Reif  
 Author, contributor  
 Kyle Messier  
 Maintainer, author, contributor

**Dev status**  
   
   

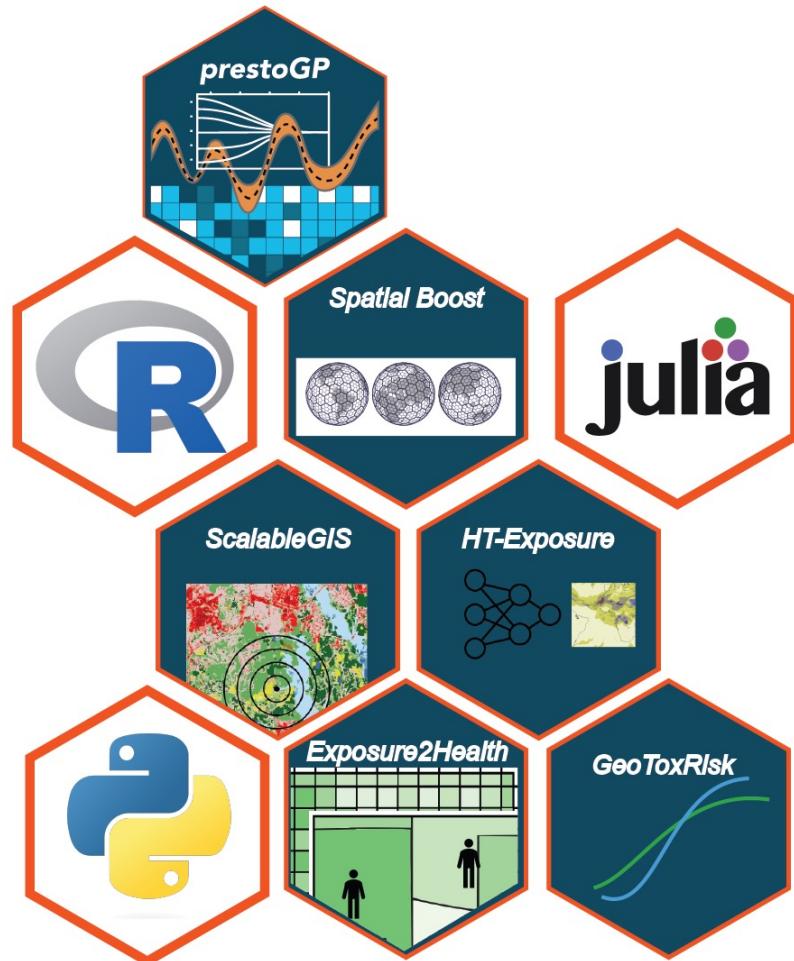



Tentative or Early Plans to Form a Working Group on Software Best Practices in Environmental and Climate Health

Reach out if you are interested

CF conventions like process

NIEHS is moving to the NIH GitHub Enterprise



# Acknowledgements

## SET group

- Daniel Zilber, PhD
- Insang Song, PhD
- Mariana Alifa, PhD
- Ranadeep Daw, PhD
- Eva Marques, PhD
- Mitchell Manware, MS
- Former members:
  - Kristin Eccles, PhD
  - Melissa Lowe, MS
  - Taylor Potter, BA

## DTT

- David Reif, PhD
- Skylar Marvel, PhD

## Biostatistics and Comp Biology

- Alison Motsinger-Reif, PhD

## Funding



## Contact

[kyle.messier@nih.gov](mailto:kyle.messier@nih.gov)

<https://github.com/Spatiotemporal-Exposures-and-Toxicology>