The HHEAR Data Repository: Facilitating environmental health and exposome research through data science, harmonization, and accessibility

COLUMBIA | MAILMAN SCHOOL OF PUBLIC HEALTH

Jeanette A Stingone, James Masters, Sofia Bengoa, Paulo Pinheiro, Liyang Yu, Jamie P. McCusker, Deborah L. McGuinness, Susan L. Teitelbaum

Mount Sinai   Rensselaer

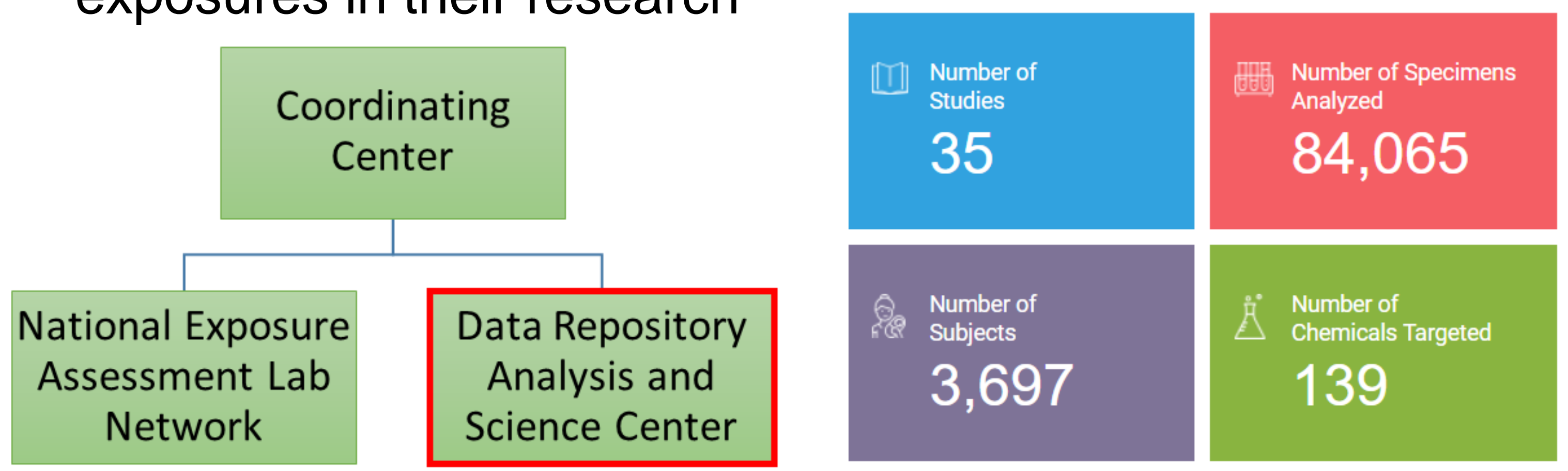https://hheardatacenter.mssm.edu/

The HHEAR Data Repository uses semantic technology and data standards to harmonize across disparate epidemiologic studies. This enables secondary data analysis of pooled data, allowing for investigations with larger sample sizes and greater exposure variability.
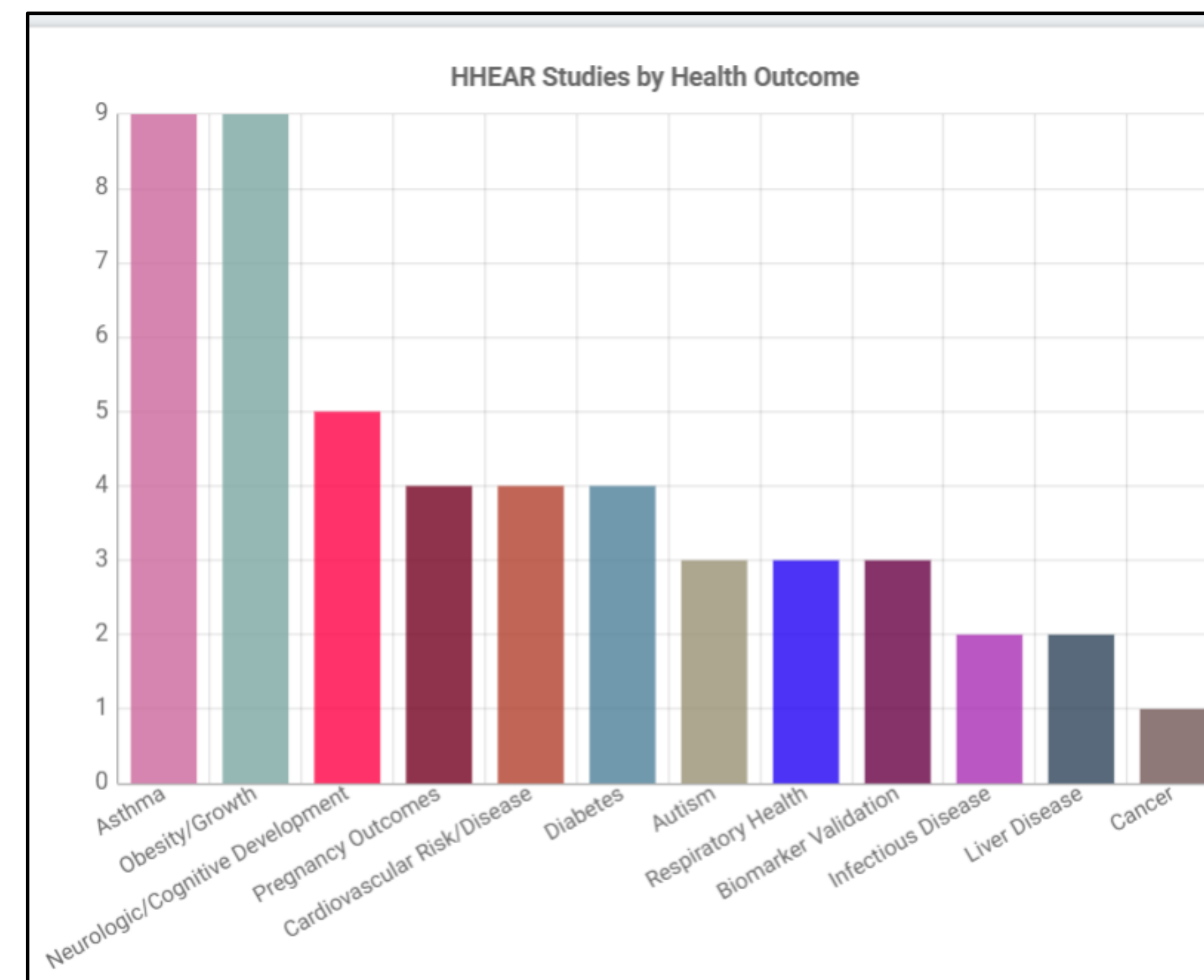
What is HHEAR?
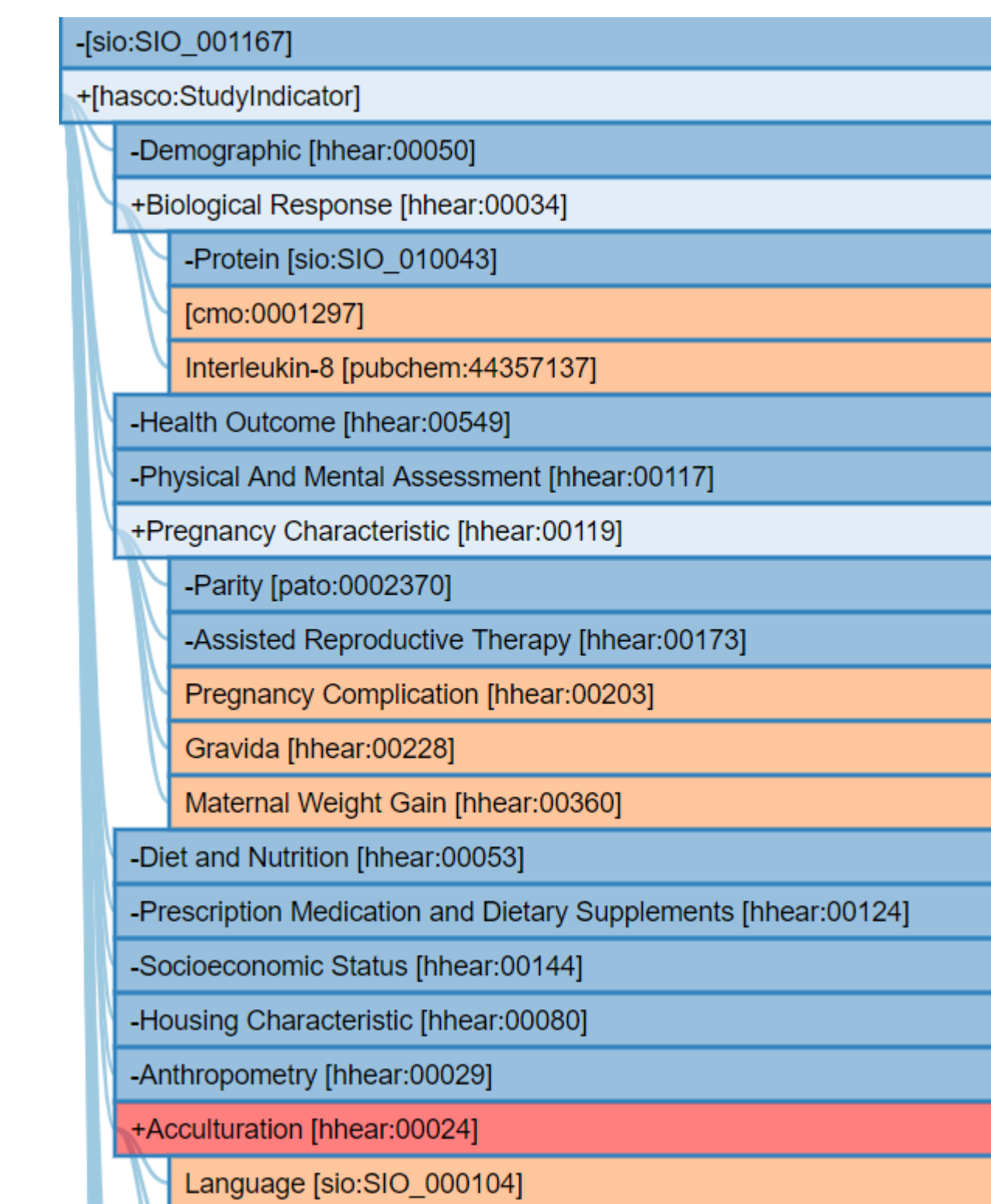### Human Health Exposure Analysis Resource

GOAL: infrastructure to provide the scientific community access to laboratory and statistical analyses aimed at adding or expanding the inclusion of environmental exposures in their research

- Coordinating Center
  - National Exposure Assessment Lab Network
  - Data Repository Analysis and Science Center

Number of Studies: **35**	Number of Specimens Analyzed: **84,065**
Number of Subjects: **3,697**	Number of Chemicals Targeted: **139**

Health Outcomes in HHEAR

HHEAR Studies by Health Outcome
(Asthma, Obesity/Growth, Neurologic/Cognitive Development, Pregnancy Outcomes, Cardiovascular Risk/Disease, Diabetes, Autism, Respiratory Health, Biomarker Validation, Infectious Disease, Liver Disease, Cancer)

Supported by the HHEAR Ontology

- [sio:SIO_001167]
  - +[hasco:StudyIndicator]
    - -Demographic [hhear:00050]
    - +Biological Response [hhear:00034]
      - -Protein [sio:SIO_010043]
        - [cmo:0001297]
        - Interleukin-8 [pubchem:44357137]
    - -Health Outcome [hhear:00549]
    - -Physical And Mental Assessment [hhear:00117]
    - +Pregnancy Characteristic [hhear:00119]
      - -Parity [pato:0002370]
      - -Assisted Reproductive Therapy [hhear:00173]
      - Pregnancy Complication [hhear:00203]
      - Gravida [hhear:00228]
      - Maternal Weight Gain [hhear:00360]
    - -Diet and Nutrition [hhear:00053]
    - -Prescription Medication and Dietary Supplements [hhear:00124]
    - -Socioeconomic Status [hhear:00144]
    - -Housing Characteristic [hhear:00080]
    - -Anthropometry [hhear:00029]
    - +Acculturation [hhear:00024]
      - Language [sio:SIO_000104]

Full Ontology Accessible via Bioportal

Designed for interoperability with existing and commonly used ontologies

The HHEAR Data Repository is the public interface for searching and downloading HHEAR Study Data

Visualize content of the repository

Search and select by Study

Search and select by Study Features

Search and select by Developmental Window

POWERED BY SEMANTIC TECHNOLOGY

> Interface enabled by open-source HADatAC software infrastructure composed of data and metadata repositories, including Solr and Blazegraph www.hadatac.org

> Individual studies mapped to data standards and ontology classes using semantic data dictionaries (Rashid et al Data Intelligence 2020)

Findable: DOI indexed; Searchable by study and multiple features

Accessible: Use of open file formats; Web-based access

Interoperable: Terms mapped to standard ontologies; Harmonized variable names and codes; Automatically generated data dictionaries

Reusable: Customizable downloads with corresponding data dictionaries; Version control; Codes stable across versions.

Sign up to Access the Repository