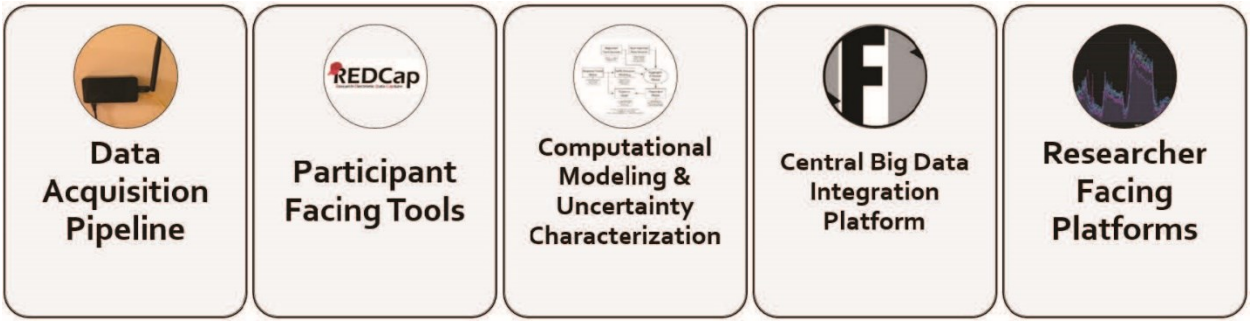
Title: The Utah PRISMS Informatics Ecosystem: An Infrastructure for Generating and Utilizing Exposomes for Translational Research

Co-Authors: Ram Gouripeddi, Randy Madsen, Scott Collingwood, Julio Facelli, Katherine Sward

Abstract:

Quantifying effects of the modern environment on health requires taking into account data from all contributing environmental exposures (*exposome*) which can span endogenous processes within the body, biological responses of adaptation to environment, and socio-behavioral factors. Exposomic research is translational in nature as the exposome includes direct biological pathway alterations as well as mutagenic and epigenetic mechanisms of environmental influences on the phenome. Generating exposomes requires integration of data from wearable and stationary sensors, environmental monitors, physiology, medication use and other clinical data, genomic and other biospecimen-derived, person-reported and computational models. This aggregation and integration requires to support variable spatio-temporal resolutions due to differences in study, experimental and analytical designs. Gaps in measured data may need to be filled with modeled data along with characterization of uncertainties.

We are developing a scalable computation infrastructure, the Utah PRISMS (Pediatric Research Using Integrated Sensor Monitoring Systems) Informatics Ecosystem (UPIE) to address these needs. UPIE is a comprehensive, standards-based, open-source informatics platform that provides semantically consistent, metadata-driven, event-based management of exposomic data. Using an event-driven architecture always the modeling and storage of all activities related to the study itself and its operations in their primitive form on a timeline as events that can be transformed to higher/analytical models based on use-cases. It is aligned with the goals of modern environmental health research supporting meaningful integration of sensor and biomedical data. It consists of the following components:



1. Data acquisition pipeline: Hardware and software tools, wireless networking, and protocols to support easy system deployment, robust sensor data collection, and feedback to study participants.
2. Participant facing tools: Collect and annotate various patient reported and activity data, as well as inform participants on their current clinical and environmental status.
3. Computational modeling: Generate comprehensive spatio-temporal data in the absence of measurements and for recognition of activity signatures from sensor measurements.
4. Central *big data* federation/integration platform: Standards-based, open-access infrastructure that integrates measured and computationally modeled data with biomedical information along with characterizing uncertainties associated with using these data.

5. Researcher facing platforms: Tools and processes for researchers undertaking exposomic studies for a variety of experimental designs and for clinical care.

At this point we have demonstrated the feasible of using UPIE to perform longitudinal cohort exposure health studies. A PRISMS pilot study has now an integrated store of over 25 million environmental (indoor and outdoor) as well as clinical events with ongoing study analysis and further data acquisition and integration. The ECHO study has so far 4,358,194 sensor readings over 2 months from 28 homes. Most participants were supportive of using their personal Wi-Fi for sensor data transmission, but we found that there is a need for flexible data transmission and storage methods and sensor deployment strategies. ECHO and PRISMS pilot studies demonstrate participant acceptance of personal and Internet-of-Things sensor monitoring; and integration of sensor and clinical/patient data to generate and utilize exposure profiles.

In this presentation, we discuss the architecture of UPIE, and the generalizability of this multi-scale and multi-omics platform for providing robust pipelines for reproducible exposomic research using results from pilot projects using real-time, low-cost air quality sensors to provide spatio-temporal records of particulate matter exposures. These pediatric studies are aligned with the National Institutes of Health (NIH) ECHO program and NIH PRISMS program and evaluated participant acceptance of personal and home-based *Internet-of-Things* sensor monitoring; demonstrated the feasibility of gathering environmental exposure data from low-cost sensors at a very fine level of resolution; demonstrated the efficacy in integrating heterogeneous data for building and analyzing exposure profiles. We are building on our initial efforts to establish a Center for Excellence in Exposure Health Informatics consisting of clinical domain experts, computer scientists, sensor experts, chemical engineers, atmospheric scientists, industrial engineers, electrical engineers and informaticians to support the next generation of novel exposomic studies.

This research is supported by the Utah PRISMS Informatics Center (NIH/NIBIB U54EB021973), the Utah Center for Clinical and Translational Science (NCATS UL1TR001067), and the Utah ECHO Program (NIH UG3OD023249). Content is sole responsibility of authors and does not necessarily represent official views of the NIH.