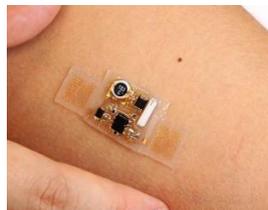# Harnessing Big Data to Understand Pediatric Asthma

The Utah PRISMS project

Katherine Sward PhD RN

Utah informatics center for the NIH *Pediatric Research using Integrated Sensor Monitoring Systems* (PRISMS) program: Federated Integration Architecture

ATS news 5/12/2016

WHO Says Air Pollution Is Worsening

The Washington Post (5/11, Mooney, Dennis) reports the WHO said Thursday that air pollution is "growing worse in urban areas across much of the globe, hitting the poorest city dwellers hardest and contributing to a wide range of potentially life-shortening health problems," including heart disease and asthma. New data released by the organization "detailed how 4 of every 5 residents of cities with reliable measurements face levels of particulate air pollution that exceed what the WHO recommends." The New York Times …reports the new WHO data, which compiled air quality readings from 3,000 cities in 103 countries…

Salt Lake City Air Quality

Courtesy: Dr. K. Kelly

- Winter inversions: colder surface temperatures trap fine particulate matter ($PM_{2.5}$) = serious health concerns.

- Summer months in the valley have increased ozone (O^3) levels.
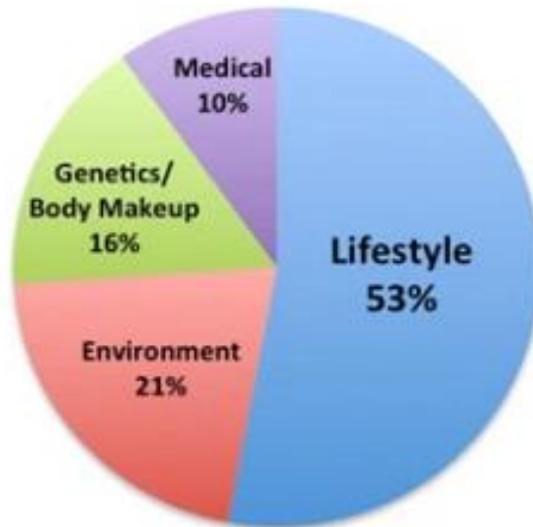
- Natural/Quasi-experimental conditions.

Utah Concludes Winter Inversion Season, Residents Proactively Engaged.
http://www.deq.utah.gov/News/docs/2014/03Mar/DAQ_NewRelease_AirQualityStats_draftv2.pdf

What Determines Your Health?

Medical 10%
Genetics/Body Makeup 16%
Lifestyle 53%
Environment 21%

Center for Disease Control, 2004
Lifestyle...........................53%
Medical............................10%
Environment.................21%
Genetics/Body Makeup.................16%

"Precision Medicine" takes into account uniqueness of individuals or groups in prevention and treatment strategies.

Genome, proteome, metabolome are widely talked about – but lifestyle and exposures are likely MORE influential

Another important "omics" area:
Clinically Relevant Exposome

Exposome: the total set of environmental factors to which a person is exposed, including
- the complex interplay between environmental factors;
- and between environmental, behavioral & psychosocial factors;

that in turn influence health and disease

Initial need for asthma: understand the risks associated with exposure to air pollutants.

Current research mainly associates single pollutant and clinical conditions, future research needs to include exposures to multiple environmental factors.

Types of Air Quality Sources

Personal Sensors
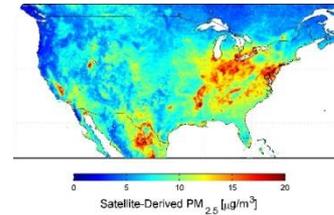

Laser Ceilometers

Novel Sensors

Mobile Sensors

Balloon Sensors

Satellite-derived aerosol optical depth measurements

State Environmental Department

Environmental Protection Agency

Sensors & Big Data V's

- No standards – heterogeneous [Variety]
- Gathered many times per second [velocity]
- Measured over long times [volume]
- Lots of uncertainty, variability, gaps – NOISY data [veracity]

More data is not the same as more information – we need the right data, presented in such a way that we can make sense of it

Data Science Challenges

- Selection of Relevant Sensor Data Sources
- Modeling for a High Spatio-temporal Grid
- Characterizing Uncertainty
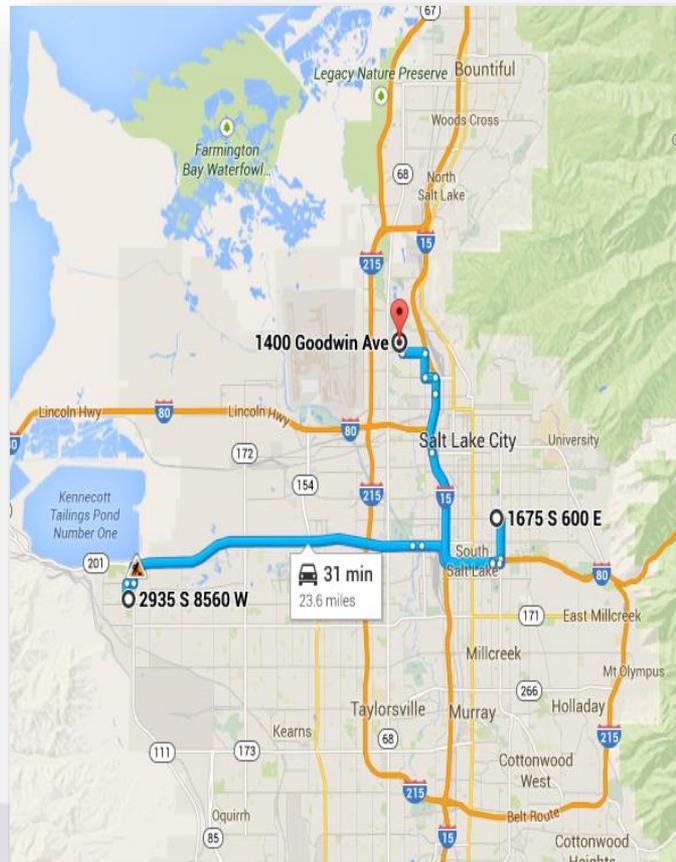- Data Integration to Support Ease of Use

⬌

Exposome: Totality of human environmental exposures from conception onwards, complementing the genome.

$$Exposure = f \begin{pmatrix} Quantity\ of\ Air\ Pollutant, \\ Duration, \\ Frequency, \\ Person\ and\ Biological\ Charateristics \end{pmatrix}$$

Pollutant & Quantity — Travel — Home — Ventilation

Outdoors — Clinical Conditions — Biological Membranes — Others

Air Quality Monitoring in Salt Lake County

- Three monitoring stations in Salt Lake County
- What is the air quality at any other location?
- Need for cross-linking patient locations and condition occurrences: <u>High Resolution Spatio-temporal Air Quality Grid</u>

Manifestations following exposure can occur immediately, after a lag phase, progress or persist over long durations

Bridging clinical process and Health IT to advance patient-centered care & improve performance

Integrating Sensors into eAT
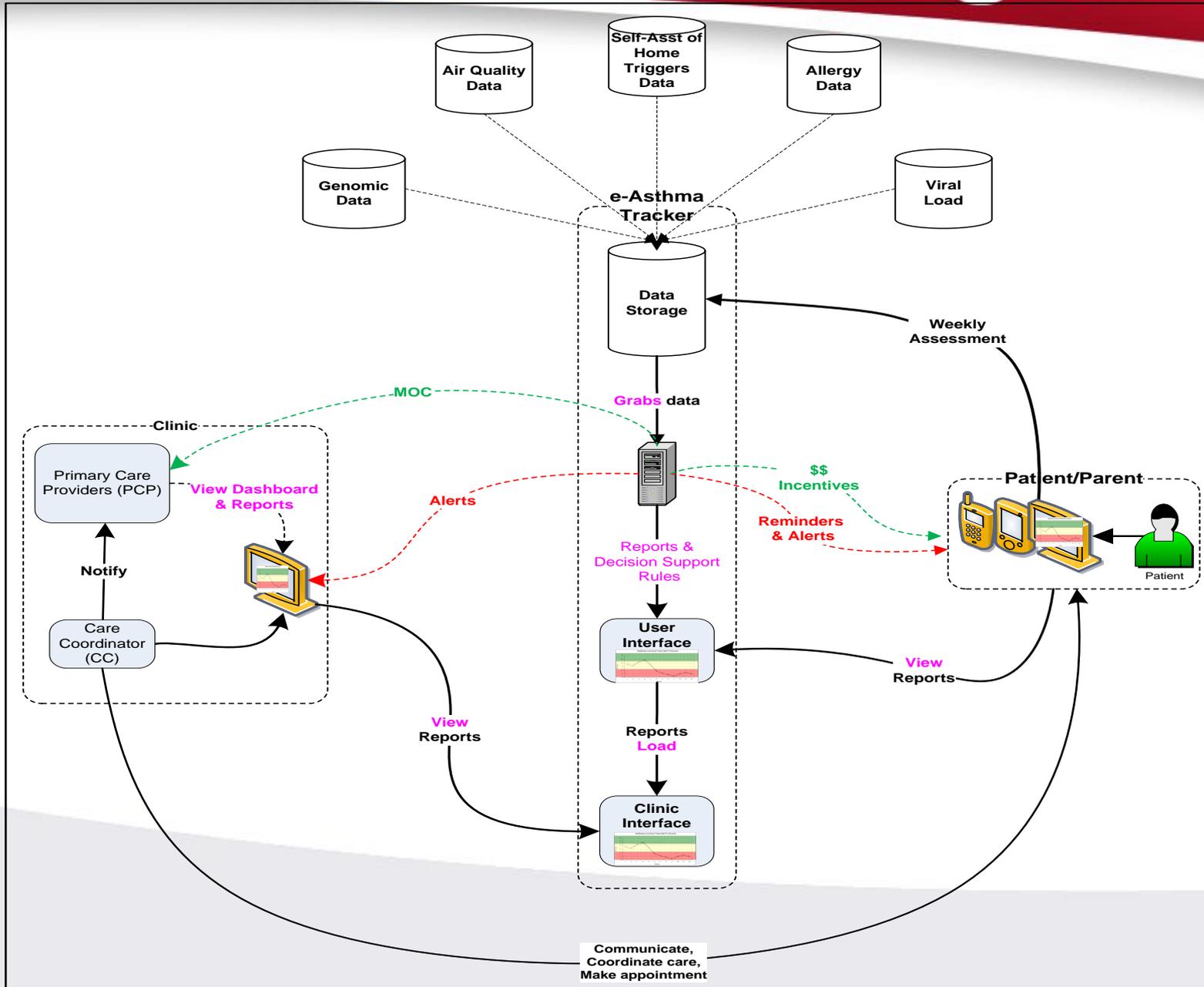
PRISMS Informatics

- A modern data infrastructure is needed to broadly and quantitatively study exposome factors (lifestyle, environment, genetics/biology, medical care)

- PRISMS center will develop efficient ways to link information from different sensors together, and how to link sensor data with other kinds of information, such as clinical records or research outcomes.

OpenFurther

- Query Tool
- Federated Query Engine
- Data Source Adapters
- Admin & Security Components
- Virtual Identity Resolution on the GO (VIRGO)
- Quality & Analytics Framework
- Metadata Repository
- Terminology/Ontology Server
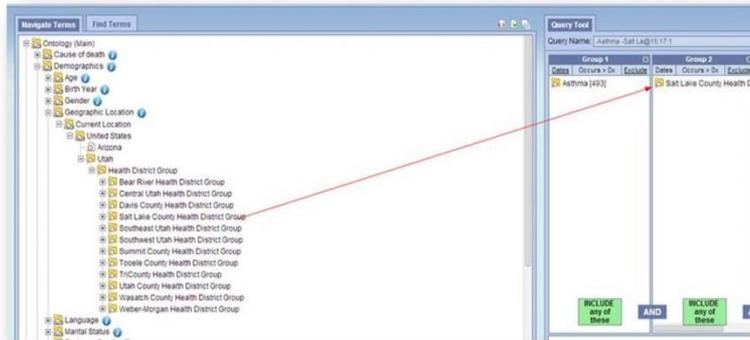- Air Quality Modelling Unit
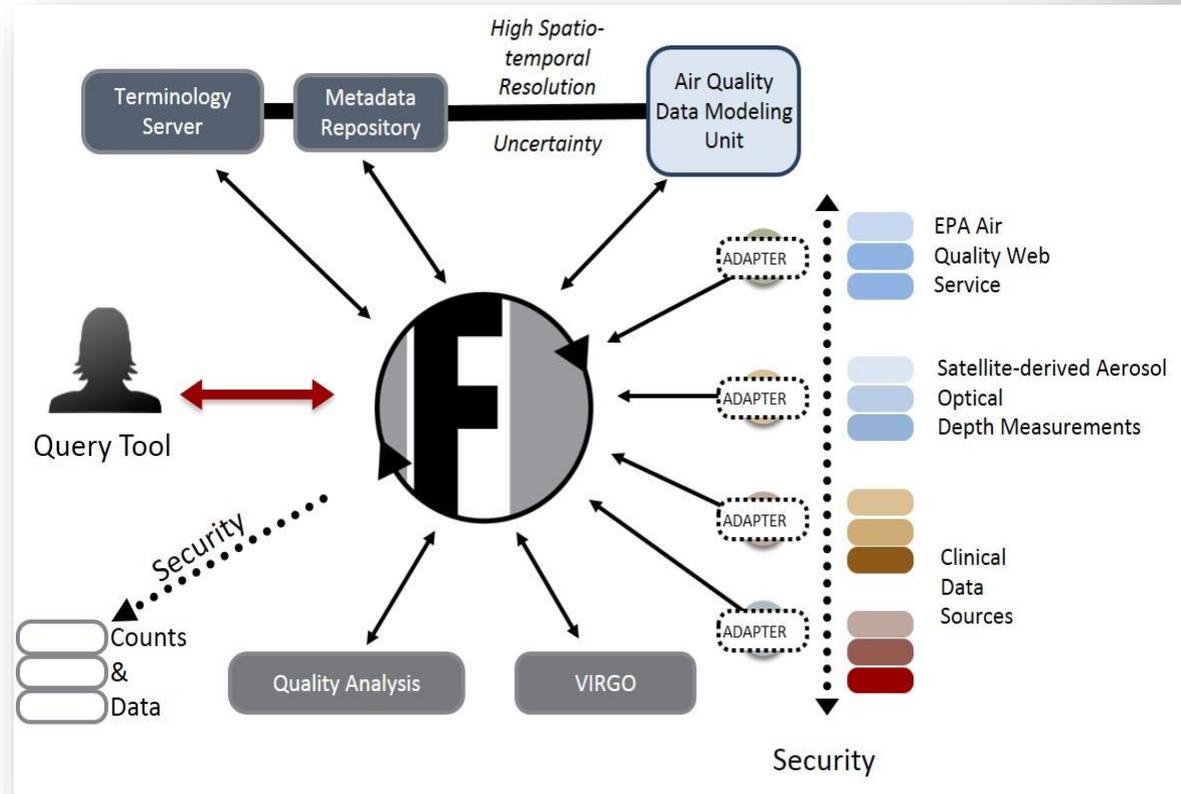
Openfurther.org
Gouripeddi R, Schultz ND, Bradshaw RL, et al. FURTHeR: An Infrastructure for Clinical, Translational and Comparative Effectiveness Research. American Medical Informatics Association, 2013 Annual Symposium; 2013 Nov 16; Washington, D.C.
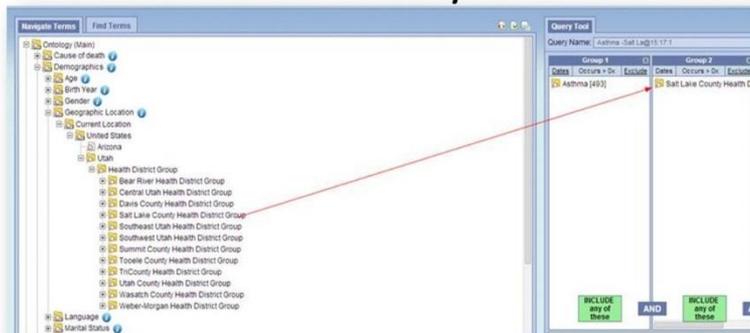http://knowledge.amia.org/amia-55142-a2013e-1.580047/t-10-1.581994/f-010-1.581995/a-184-1.582011/ap-247-1.582014

Air Quality - Clinical Data

Asthma in January 2014

615 patients with a diagnosis of asthma in Salt Lake County and average $PM_{2.5}$ 28 micrograms
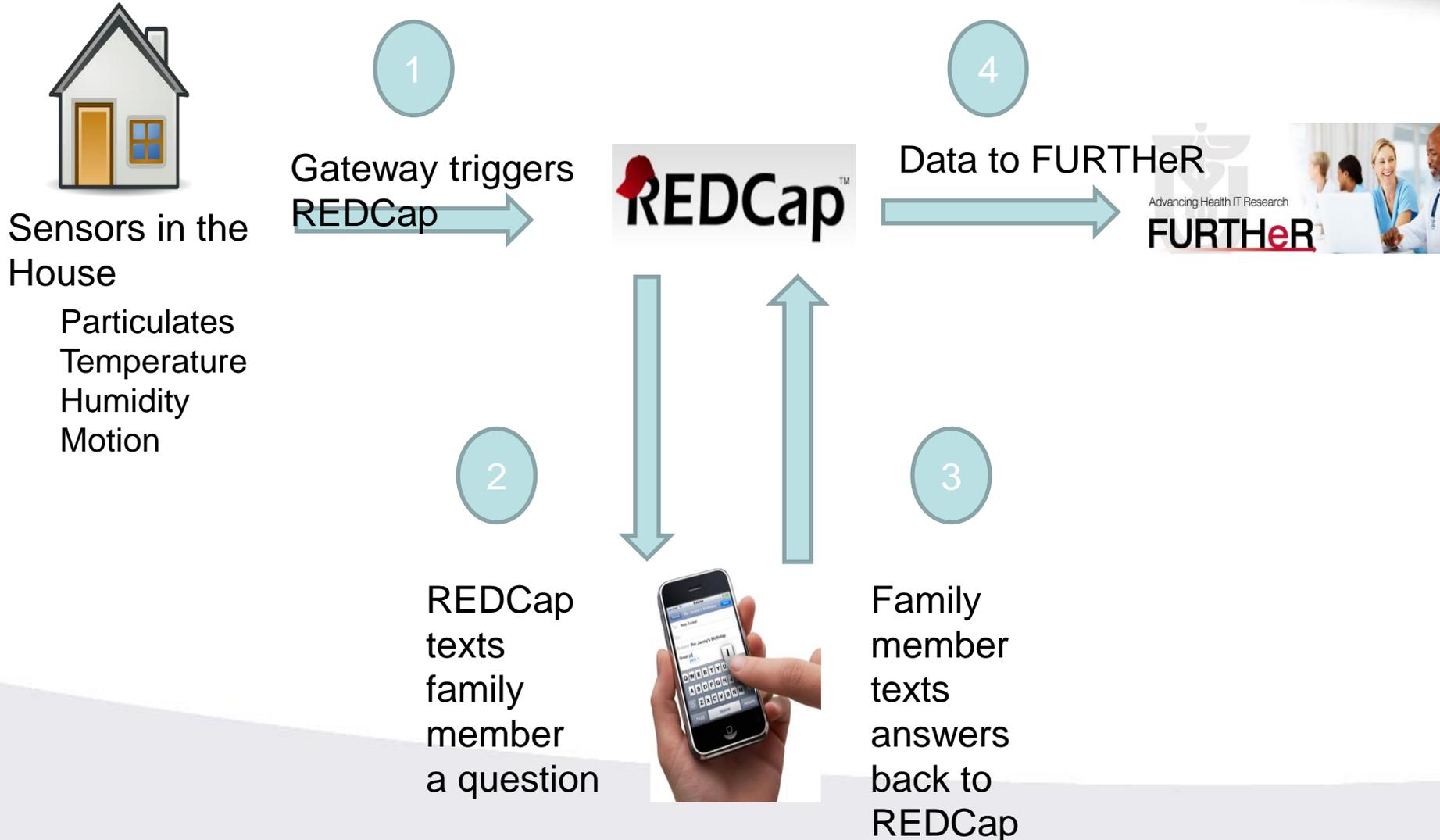
Asthma in January 20th 2014

25 patients with a diagnosis of asthma who reside in Salt Lake County and average $PM_{2.5}$ 50 micrograms

Worst Inversion Day

- Demonstrated the feasibility of federating air quality data from the Environmental Protection Agency (EPA) with clinical data from the University of Utah using OpenFurther.

- We were able to select different cohorts of patients living in SLC county and having clinical conditions (e.g. asthma) occurrences that were related to the temporal variations of air pollutant concentration.

Information from the Home

Sensors in the House
- Particulates
- Temperature
- Humidity
- Motion

1 — Gateway triggers REDCap

REDCap

4 — Data to FURTHeR

2 — REDCap texts family member a question

3 — Family member texts answers back to REDCap

Our vision

A transdisciplinary (team science) center where researchers integrage genetic, environmental, behavioral, patient-generated and clinical information to advance precision medicine (patient-centered care) and health care transformation.