Predicting asthma exacerbations using personal sensor monitoring systems

Sandrah P. Eckel, PhD
Assistant Professor
Division of Biostatistics
University of Southern California
eckel@usc.edu

Joint work with: R Habre, K Li, H Deng, R Urman, J Morrison, WJ Gauderman, JL Ambite, YY Chiang, D Stripelis, F Gilliland
Support: NIBIB U24EB021996 (data center), U54EB022002 (informatics platform & epi study)
Motivation: improve asthma management

• 1 in 12 people have asthma in the US (25 million people)
• ~50% have an asthma attack each year
• Cost of $56 billion per year (medical costs, lost school & work days, early deaths)
• Many asthma attacks can be prevented by using long-term controller medications correctly and avoiding triggers
→ Mixed success of asthma management plans

Data: https://www.cdc.gov/vitalsigns/asthma/index.html
Launched in 2015 by the National Institute of Biomedical Imaging and Bioengineering (NIH/NIBIB)

**Goal:** Develop sensor-based, integrated health monitoring systems for measuring environmental, physiological, and behavioral factors in pediatric epidemiological studies of asthma, and eventually other chronic diseases

Three arms of PRISMS:
- 6 Sensor Development Projects
- 2 Informatics Platforms
- 1 Data and Software Coordination and Integration Center
PRISMS ecosystem

U01s and off-the-shelf sensors for physiology (spirometry, ventilation, heart rate), medication usage, etc.

Sensor discovery, configuration

Data collection protocol design workbench, deployment dashboards

Integrated, synchronized, data views

Data quality assessment

Calculated exposure summary measures, health outcomes (ACT, PEF)

Risk assessment and warnings

Tailored predictive model

Targeted material

About your asthma...

Healthcare provider feedback

Real-time environmental data

U24 data center

Child and/or caregiver

Automatic, secure data upload to cloud

Self-reported symptoms, stress, behaviors

PRISMS ecosystem (from UCLA/USC LA BREATHE U54 platform, PI: Bui)
Types of data to be collected

Typical data structure:
- Time stamp (GPS stamp)
- Multiple or single features (possibly pre-processed)
- Recorded continuously or on-demand, upload frequency to optimize power

Sensors
- GPS
- Accelerometer/gyroscope to classify physical activity
- Spirometry
- Inhaler use
- Environmental measures (PM, NO₂, near roadway pollution, etc.)

Self-reported measures
- Ecological momentary assessment (EMA) for asthma symptoms, inhaler usage, stress
- Validated questionnaires (health status, physical activity, etc.)

Real-time environmental data
- Weather
- Pollen
- Air quality indices
- Nearby traffic volumes
- Indoor/outdoor metrics

Electronic health record
- Demographics, vitals
- Medications
- Allergies and documented triggers
- Health status and comorbidities
- Pulmonary function tests, labs
- Past exacerbations (e.g., ER visits)
Example data: Contextual, real-time info

Evening Commute Example

Phone Sensors, including Geolocation Accuracy

Physiological Sensors

Meteorology Sensors

(from R Habre UCLA/USC LA BREATHE U54 platform, PI: Bui)
Example data: Spatial patterns, evening commute

Air Pollution

Heart Rate

(from R Habre UCLA/USC LA BREATHE U54 platform, PI: Bui)
Conceptual overview of PRISMS data analysis

1. Evaluate sensors
   [reliability, validity, etc.]

2. Data collection
   • Baseline info
   • Ongoing collection [user adherence]

3. Key themes of planned PRISMS data analysis

   • Unsupervised cluster analysis/pattern detection
     [e.g., identify asthma phenotype groupings]

   • Identify individual baselines
     (e.g., Li et al. *PLoS Biology* 15.1 (2017): e2001402)

   • Supervised prediction of deviations from typical patterns
     A. Population-based models
     B. Cluster-specific population-based models
     C. Individualized prediction models

   • Real-time prediction
     • Train models offline (nightly) and apply to real-time streaming data
Statistical analysis pipeline: raw data $\rightarrow$ health model

1. Data processing

- Raw data files
- Noise filters
- Temporal alignment
- Transformation

2. Feature engineering

- $x \rightarrow f_1(x)$
- $f_k(x)$

3. Modeling

- Model performance, prediction, interpretation
- Machine learning: model $Y \sim X$
- Summarize within windows
Broad challenge in health modeling: $X \rightarrow Y$

Summarizing/integrating exposures ($X$):
- Most assessed ~continuously
- At high spatial & temporal resolution

Matching to health outcomes ($Y$):
1. Assessed continuously (e.g., heart rate)
2. Assessed at regular intervals (e.g., twice daily peak flow, daily asthma control test score/symptoms diary, EMA symptoms)
3. Intermittent report of (rare) events (e.g., rescue use of smart inhaler, ER hospitalization)
Exposure assignment: GPS trajectories

Time-weighted kernel density smooth


PRISMS can impact two major areas

1. Environmental epidemiology research
   - Understand environmental contributions to asthma exacerbations
   - New paradigm for exposure sciences
     (fine-scale personal exposures, with spatial and contextual info)
   → New public health policies

2. Personalized medicine
   - Trigger identification and avoidance
   - Personalized decision-making
   → Improved personal asthma management
What questions can these data answer? Policy implications?

• Time course of exposure-response
  ▪ Relevant averaging time for air quality standards

• Context: Health effects of personal vs ambient exposures
  ▪ PM$_{2.5}$ from cooking or commuting: which is more toxic?

• Identify new sources/triggers
  ▪ Are we missing a “smoking gun”?

• Heterogeneity of response to exposures (personal models)
  ▪ Standards to protect health of vulnerable groups

• Can personalized data improve asthma management?
  ▪ EMA questionnaires: symptoms, stress, etc. in context
  ▪ Patient engagement and empowerment
Patient engagement: a personal example

First trimester weight gain: -0.4 lbs/week, 95% CI: (-0.5, -0.3)
Second trimester weight gain: 1 lbs/week, 95% CI: (1, 1.1)
Third trimester weight gain: 1.3 lbs/week, 95% CI: (1.2, 1.3)

Environmental epidemiology: PRISMS-like study

- 30 children, 8 days
- Personal exposure to fine particulate matter divided by microenvironment
- Exposure “spikes” during transit (vs. home or school) were most strongly related to biomarker of exposure

Challenges and open questions

• Wearable sensors have to be worn
  ▪ Compliance, missing data
  ▪ How do specific sensors need to be worn? Is GPS enough?

• Sensors sense
  ▪ Will we measure the right things?

• Cheap sensors are cheap
  ▪ Requires calibration, more expensive QA/QC, processing
  ▪ Incorporate data quality metrics in models?

• Personal monitoring is personal
  ▪ Privacy/Ethics issues: GPS trajectories, heart rate, etc..

• Real-time sensors are real-time
  ▪ Large volumes of data, potentially not relevant timescale
  ▪ Feedback to user can influence behavior
Thank you!

eckel@usc.edu