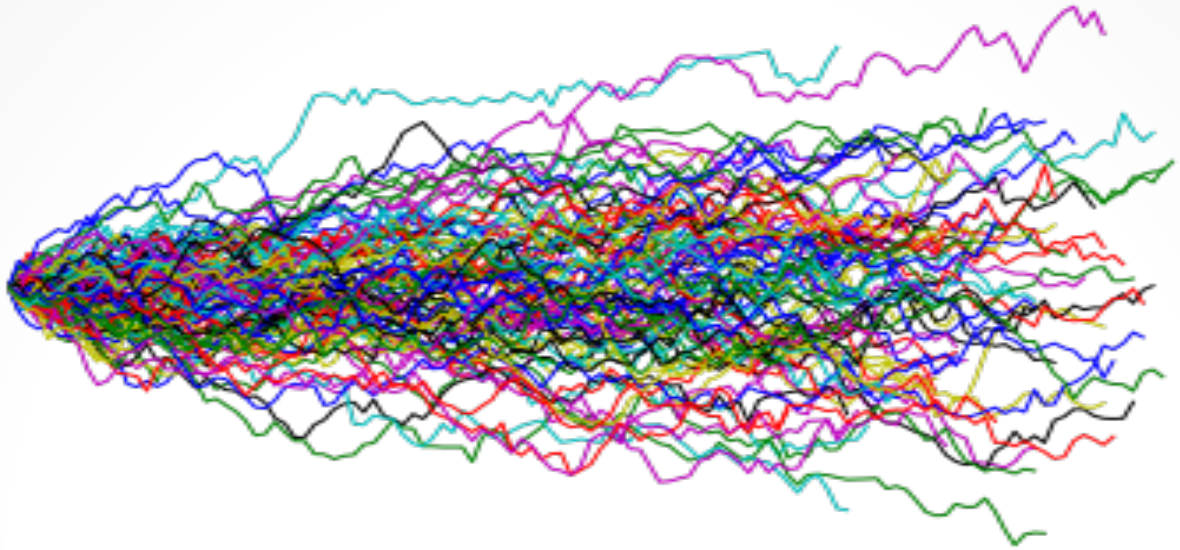


Penn Signals

Accelerating Machine Learning Solutions and predictive applications in health care

Corey Chivers, PhD

Penn Medicine

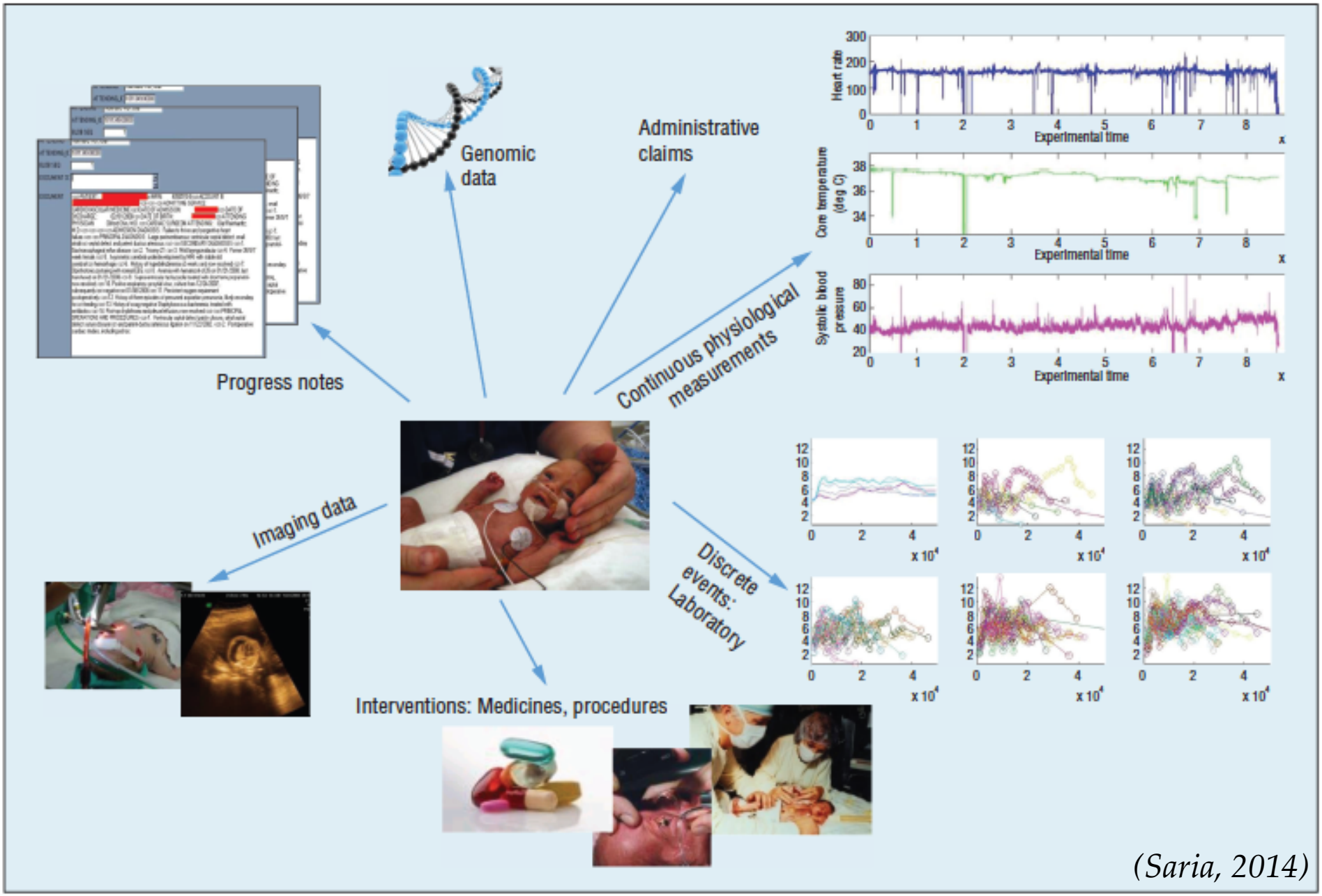


Figure 1. An illustration of the diverse electronic health data (EHD) that are routinely collected, including physiological measurements, laboratory test results, medications administered, imaging test results, progress and discharge reports, genomic profiles, and administrative claims.

Realizing the Opportunity: Detecting Severe Sepsis

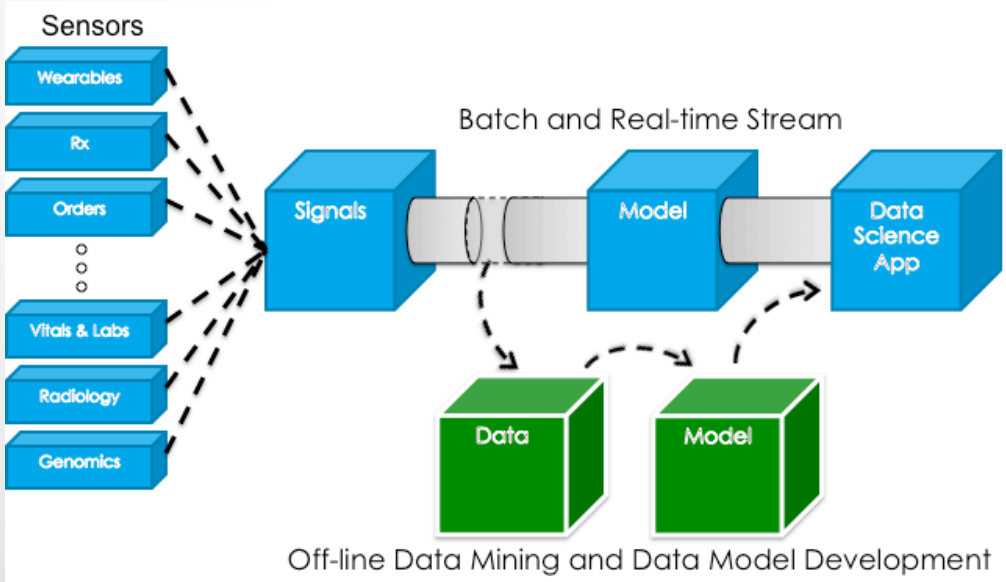
- Severe Sepsis
  - Every hour undiagnosed is a 7.6% increase in mortality
  - Only 50% of Septic Shock patient receive effective therapy in time
  - $30B annual Impact the US Health System

- Methods
  - **OLD: Clinical guidelines use 6 vitals and labs values with threshold rule**
  - **NEW:** Uses over 200 clinical variables and 100,000's of training examples

- Comparison
  - 8 alerts/day (EWS2.0) vs 20 alerts/day (EWS1.0)
  - 92% True Positive Rate (EWS2.0) vs 59% TP (EWS1.0)
  - Alerts are day's before onset of patient shock

The What
Real-time Detection & Predictive Apps

Alerting when patient at risk of of an acute event



Fast Deployment in the hands of care team

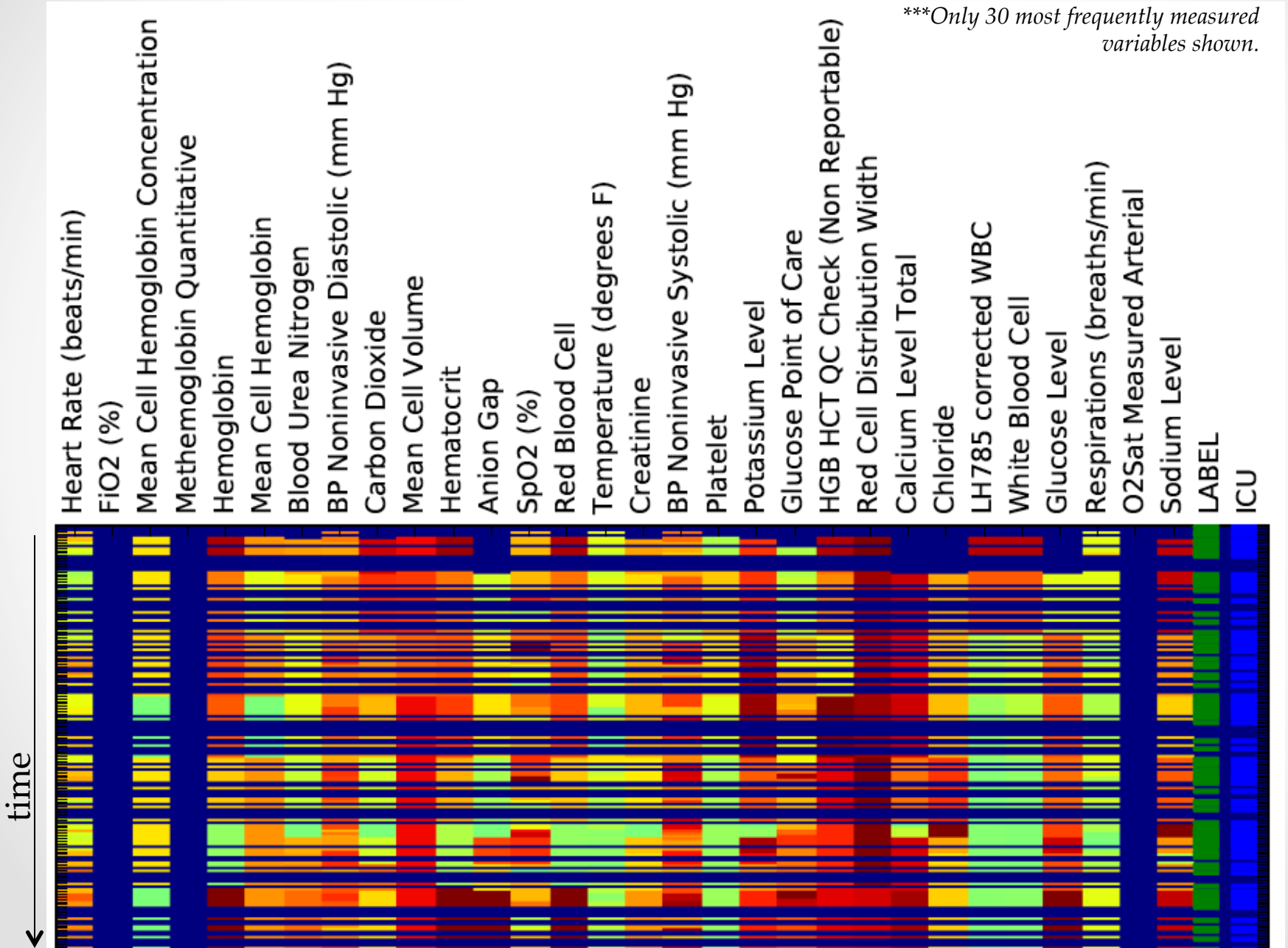
Fast Research, Data Mining & App Development

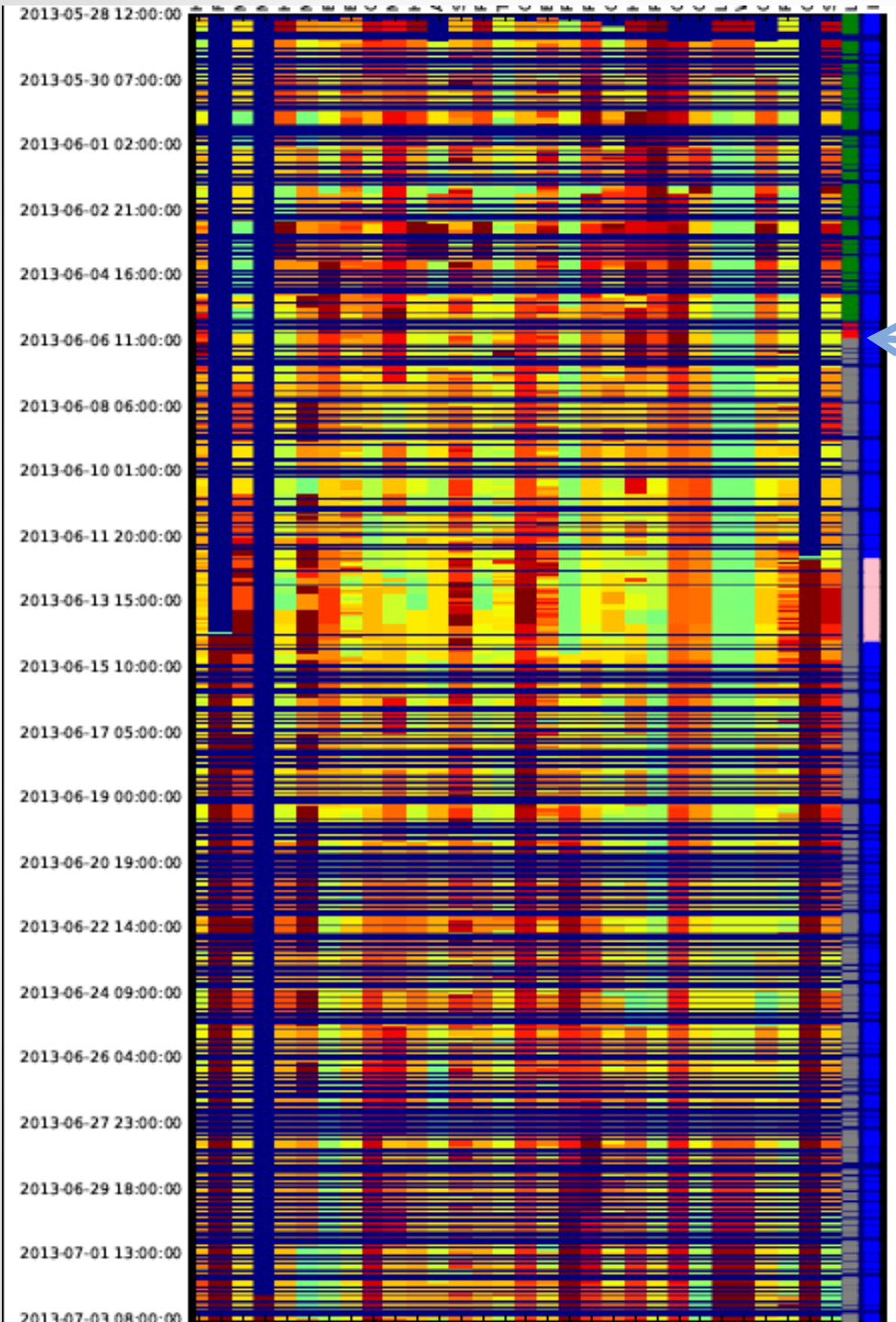
The How
Combined Data Pipeline for Research and App workflow

Pain points

- Integrating and fusing new data sources
- Truth labels can be elusive
- Hard to iterate
  explore -> solve -> evaluate -> repeat
- How well will it work *in the wild?*
- Going from a nifty model to production deployment of a predictive application

Our Class of Problems: Time Series
What the machine 'sees'



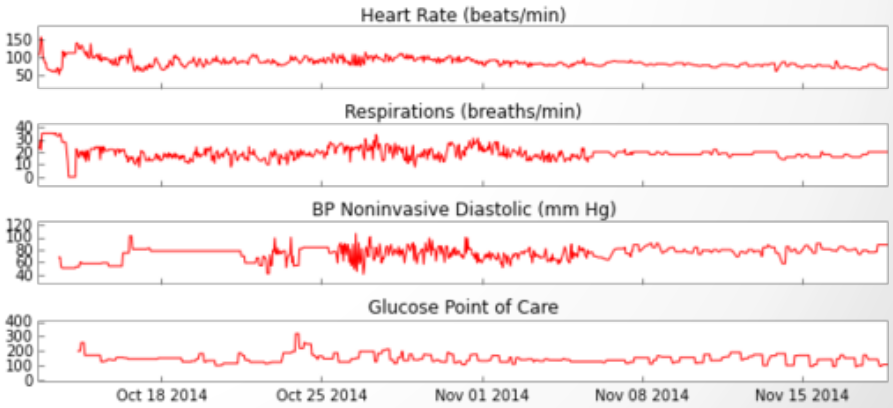


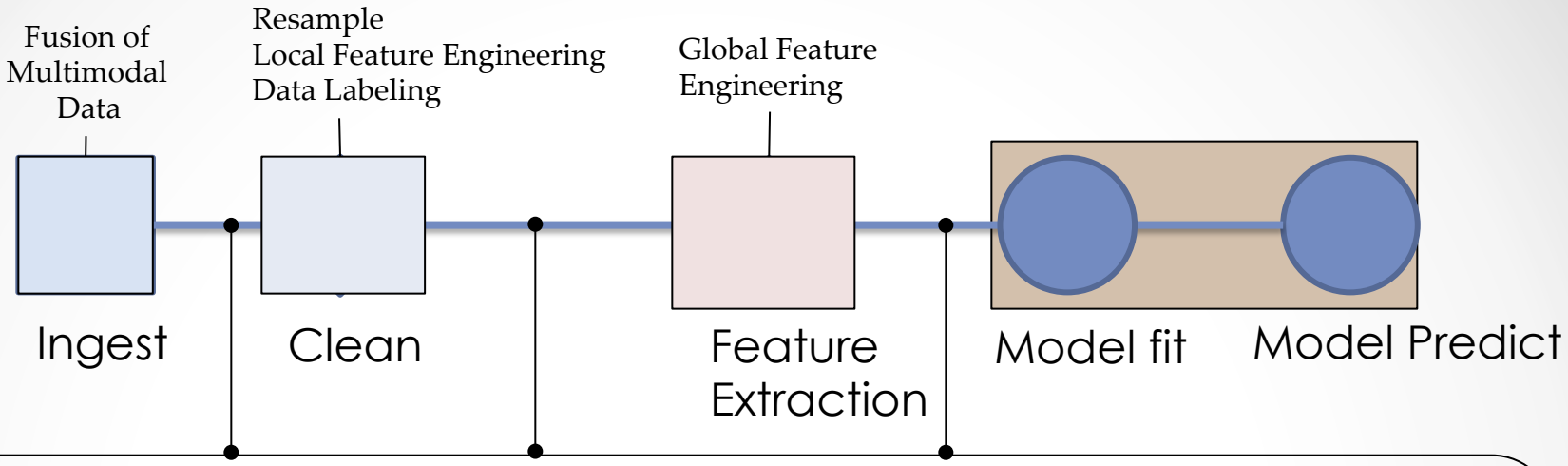
'Pre-septic'

'12 hours pre-onset window'

(Positive Label)

ICU

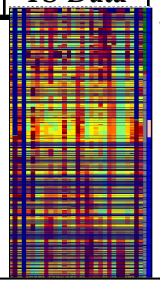




Data Schema: Simple, complete and self contained

Pt ID	Time	Adm	Age	Lab Stats	…	Zip	TS Data

Fast query and data mining:
Patient static data with statistical summary of temporal data.



Feeding the Data Model
Payload
Raw Time Series
Column Form

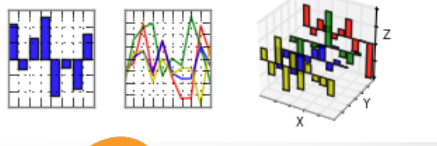
Open source processing tools

mongoDB

{ name: mongo, type: DB }

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



ØMQ

scikit learn

Penn Medicine

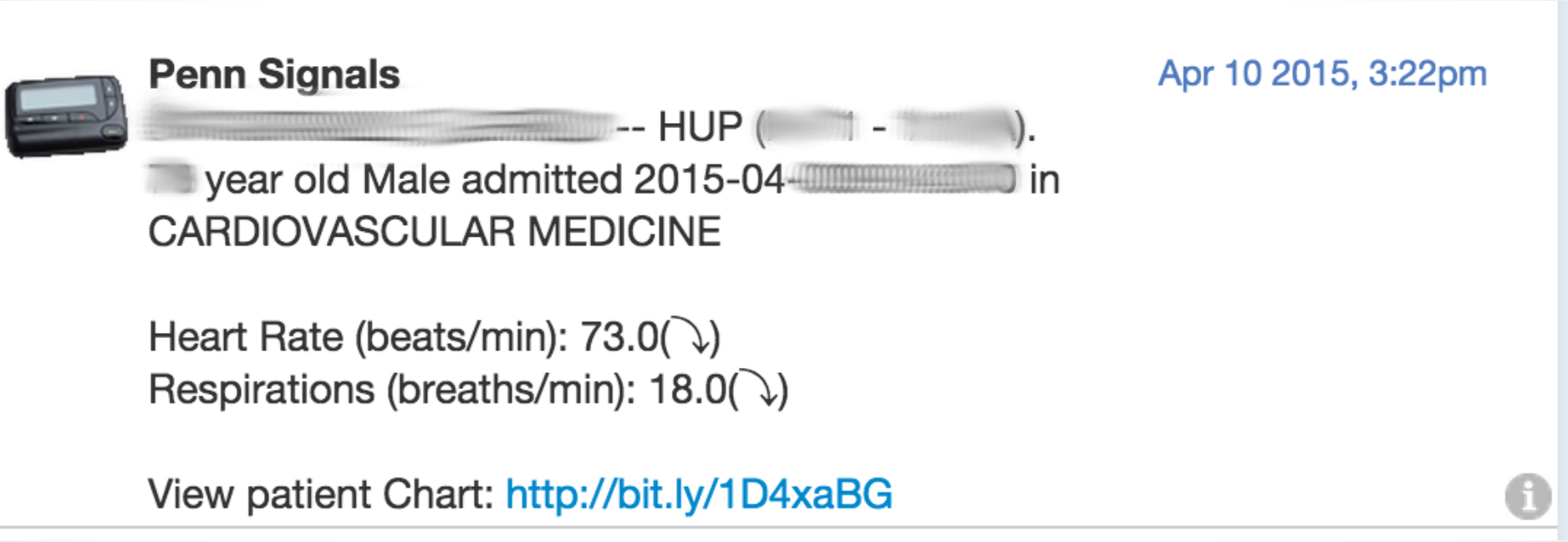
From start to finish

- SignalsML provides a framework on which to build predictive applications in health care
  - Data -> model -> test -> deploy

1. Importing historical data
2. Adding truth labels
3. Feature Engineering
4. Careful k-folds generation
5. Model building and CV
6. Operational Performance
7. Deployment

Demo!

Acute Event Alert



On the horizon

- Deep learning (Terabytes of radiology imaging data)
- Bringing data science to health beyond the hospital walls
- Empowering patients and clinicians with data

Join us!

- Develop data products and predictive applications
- Apply cutting edge machine learning and computational statistics.
- *Bathe* in *all* the data (eat, sleep and breath it!)
- Collaborate with top medical professionals
- Revolutionize Health care delivery

Contact:
corey.chivers@uphs.upenn.edu (@cjbayesian)
or
michael.draugelis@uphs.upenn.edu

Penn Medicine