Penn Signals
Accelerating Machine Learning Solutions and predictive applications in health care
Corey Chivers, PhD
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.

(Saria, 2014)
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 an acute event

The How
Combined Data Pipeline for Research and App workflow

Fast Deployment in the hands of care team

Fast Research, Data Mining & App Development
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’

***Only 30 most frequently measured variables shown.***

<table>
<thead>
<tr>
<th>Heart Rate (beats/min)</th>
<th>FIO2 (%)</th>
<th>Mean Cell Hemoglobin Concentration</th>
<th>Methemoglobin Quantitative</th>
<th>Hemoglobin</th>
<th>Mean Cell Hemoglobin</th>
<th>Blood Urea Nitrogen</th>
<th>BP Noninvasive Diastolic (mm Hg)</th>
<th>Mean Cell Volume</th>
<th>Hematocrit</th>
<th>Anion Gap</th>
<th>Red Blood Cell</th>
<th>Temperature (degrees F)</th>
<th>Creatinine</th>
<th>BP Noninvasive Systolic (mm Hg)</th>
<th>Platelet</th>
<th>Potassium Level</th>
<th>Glucose Point of Care</th>
<th>HGB HCT QC Check (Non Reportable)</th>
<th>Red Cell Distribution Width</th>
<th>Calcium Level Total</th>
<th>Chloride</th>
<th>LH785 corrected WBC</th>
<th>White Blood Cell</th>
<th>Glucose Level</th>
<th>O2 Sat Measured Arterial</th>
<th>Sodium Level</th>
<th>LABEL</th>
<th>ICU</th>
</tr>
</thead>
</table>

**Time**
'Pre-septic'

'12 hours pre-onset window'

(Positive Label)

ICU
Data Schema: Simple, complete and self contained

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

Open source processing tools

mongoDB

{ name: mongo, type: DB }

pandas

\[ u_{it} = \beta x_{it} + \mu_i + \epsilon_{it} \]
From start to finish

• SignalsML provides a framework on which to build predictive applications in health care
  o 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

Penn Signals

-- HUP ( ).

year old Male admitted 2015-04- in CARDIOVASCULAR MEDICINE

Heart Rate (beats/min): 73.0(↓)
Respirations (breaths/min): 18.0(↓)

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