Predicting Clinical Events from Electronic Health Records

David Page, University of Wisconsin-Madison

Microsoft Research © David Page 2013
Thanks to:

Michael Caldwell
Peggy Peissig
Jesse Davis
Vitor Santos Costa
Sriraam Natarajan
Jeremy Weiss
Eric Lantz
Elizabeth Burnside
Kendrick Boyd
Jie Liu
Aubrey Barnard
Humberto Vidaillet
Jude Shavlik
Houssam Nassif
International Warfarin Pharmacogenomics Consortium

Funding:

NIGMS grant R01GM097618
NLM grant R01LM011028
UW Institute for Clinical and Translational Research
UW Carbone Cancer Center
Wisconsin Genomics Initiative
Microsoft Research
Physicians
700,000 US Physicians
Accidental Annual Deaths > 120,000
Accidental Deaths/Physician = 0.171

Guns
80,000,000 Owners
Accidental Gun Deaths/Year~1500
Accidents/Gun Owner=0.000188

"Of course my blood pressure is high... from worrying about having a heart attack if I don't lower my cholesterol..."
...The top prescription is for your arthritis, but it may cause a heart attack. The second prescription should prevent a heart attack, but it could damage your liver. The third should prevent liver trouble, but it may destroy your spleen. The fourth protects the spleen but has been known to eat away the prostate. The fifth....
Predictive Models for Medicine

State-of-the-Art Machine Learning

Genetic, Clinical, & Environmental Data

Predictive Model for Disease Susceptibility & Treatment Response

Personalized Treatment

Courtesy of Wisconsin Genomics Initiative (WGI)

Microsoft Research
Why Now?

- Widespread use of electronic health records (EHRs) and data transfer to data warehouses
- US$1000 genome and SNP-based GWAS
- Recognition of need
  - Adverse drug events and other medical dangers
  - Rapidly-rising health care costs, sustainability

Microsoft Research
Some Prediction Tasks

- Stable dose of Warfarin (NEJM ‘09)
- Mammogram malignancies (Radiology ‘09)
- Myocardial infarction risk (AI Mag ‘12)
- Atrial fibrillation risk (under review)
- Age-related macular degeneration risk (u.r.)
- Risk of a known ADE if given drug (ICML ‘12)
- Discovery of unknown ADEs (AAAI ‘12)
## Patient ID | Gender | Birthdate
---|---|---
P1 | M | 3/22/1963

## Patient ID | Date | Physician | Symptoms | Diagnosis
---|---|---|---|---
P1 | 1/1/2001 | Smith | palpitations | hypoglycemic
P1 | 2/1/2001 | Jones | fever, aches | influenza

## Patient ID | Date | Lab Test | Result
---|---|---|---
P1 | 1/1/2001 | blood glucose | 42
P1 | 1/9/2001 | blood glucose | 45

## Patient ID | Date | Observation | Result
---|---|---|---
P1 | 1/1/2001 | Height | 5'11
P2 | 1/9/2001 | BMI | 34.5

## Patient ID | Date Prescribed | Date Filled | Physician | Medication | Dose | Duration
---|---|---|---|---|---|---
P1 | 5/17/1998 | 5/18/1998 | Jones | Prilosec | 10mg | 3 months
Two Very Different ADE Tasks

• **Given:** an EHR and a known ADE (a `<drug,condition>` pair)
  **Do:** learn model to predict (at prescription time) whether a patient will have the ADE if they take the drug

• **Given:** an EHR and a specified drug
  **Do:** find conditions caused by the drug (ADE)
Predicting Myocardial infarction given Cox2 Inhibitor

- Non-steroidal anti-inflammatory drug
- Cox-2 goal: reduce stomach trouble

Cox-1

Aspirin, Aleve, Ibuprofen, etc block both pathways

Cox-2

Vioxx, Bextra, Celebrex block this pathway
Dec. 1998-May 1999, Celebrex, Vioxx approved

2001, Cox-2 sales top $6 billion/year in US

2002, Beginning of APPROVe Study

Sept 2004, Vioxx voluntarily pulled from market

Dec. 2004, FDA issues warning

April 2005, FDA removes Bextra from market
SAYU-NB

Score = 0.00

Class Value

Rule 14
Rule N

seed 2
SAYU-View

[Davis et al. Intro to SRL 06]
Predicting MI Given Cox2 Inhibitor
# Mammography Structured Reports

<table>
<thead>
<tr>
<th>Patient</th>
<th>Abnormality</th>
<th>Date</th>
<th>Calcification Fine/Linear</th>
<th>...</th>
<th>Mass Size</th>
<th>Loc</th>
<th>Benign/ Malignant</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>1</td>
<td>5/02</td>
<td>No</td>
<td></td>
<td>0.03</td>
<td>RU4</td>
<td>B</td>
</tr>
<tr>
<td>P1</td>
<td>2</td>
<td>5/04</td>
<td>Yes</td>
<td></td>
<td>0.05</td>
<td>RU4</td>
<td>M</td>
</tr>
<tr>
<td>P1</td>
<td>3</td>
<td>5/04</td>
<td>No</td>
<td></td>
<td>0.04</td>
<td>LL3</td>
<td>B</td>
</tr>
<tr>
<td>P2</td>
<td>4</td>
<td>6/00</td>
<td>No</td>
<td></td>
<td>0.02</td>
<td>RL2</td>
<td>B</td>
</tr>
<tr>
<td>Patient</td>
<td>Abnormality</td>
<td>Date</td>
<td>Calcification</td>
<td>Mass Size</td>
<td>Incr. In Size</td>
<td>Loc</td>
<td>Benign/Malignant</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>-------</td>
<td>---------------</td>
<td>-----------</td>
<td>---------------</td>
<td>------</td>
<td>-----------------</td>
</tr>
<tr>
<td>P1</td>
<td>1</td>
<td>5/02</td>
<td>No</td>
<td>0.03</td>
<td>No</td>
<td>RU4</td>
<td>B</td>
</tr>
<tr>
<td>P1</td>
<td>2</td>
<td>5/04</td>
<td>Yes</td>
<td>0.05</td>
<td>Yes</td>
<td>RU4</td>
<td>M</td>
</tr>
<tr>
<td>P1</td>
<td>3</td>
<td>5/04</td>
<td>No</td>
<td>0.04</td>
<td>No</td>
<td>LL3</td>
<td>B</td>
</tr>
<tr>
<td>P2</td>
<td>4</td>
<td>6/00</td>
<td>No</td>
<td>0.02</td>
<td>No</td>
<td>RL2</td>
<td>B</td>
</tr>
</tbody>
</table>
Identifying Malignant Abnormalities from Mammography Structured Reports
## Electronic Medical Record (EMR)

### Demographics

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Gender</th>
<th>Birthdate</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>M</td>
<td>3/22/1963</td>
</tr>
</tbody>
</table>

### Diagnoses

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Date</th>
<th>Physician</th>
<th>Symptoms</th>
<th>Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>1/1/2001</td>
<td>Smith</td>
<td>palpitations</td>
<td>hypoglycemic</td>
</tr>
<tr>
<td>P1</td>
<td>2/1/2001</td>
<td>Jones</td>
<td>fever, aches</td>
<td>influenza</td>
</tr>
</tbody>
</table>

### Lab Results

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Date</th>
<th>Lab Test</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>1/1/2001</td>
<td>blood glucose</td>
<td>42</td>
</tr>
<tr>
<td>P1</td>
<td>1/9/2001</td>
<td>blood glucose</td>
<td>45</td>
</tr>
</tbody>
</table>

### Vitals

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Date</th>
<th>Observation</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>1/1/2001</td>
<td>Height</td>
<td>5'11</td>
</tr>
<tr>
<td>P2</td>
<td>1/9/2001</td>
<td>BMI</td>
<td>34.5</td>
</tr>
</tbody>
</table>

### Medications

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Date Prescribed</th>
<th>Date Filled</th>
<th>Physician</th>
<th>Medication</th>
<th>Dose</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>5/17/1998</td>
<td>5/18/1998</td>
<td>Jones</td>
<td>Prilosec</td>
<td>10mg</td>
<td>3 months</td>
</tr>
</tbody>
</table>
Random Forest Prediction of Atrial Fibrillation/Flutter and Subsequent Mortality or Stroke

Figure 1. ROC Curves for our four target prediction tasks: predicting AF/F onset and, given AF/F onset, predicting each...features have non-negligible importance scores. The top features are presented to gain some insight from the model.
Some Lessons So Far

- EHR data multi-relational: use SRL and ILP
- Often can do ~ as well with random forests
  - Features: counts of diagnoses, abnormal labs, prescriptions in 1, 3, 5, n years
  - Linear SVMs with 1000 top features second best
- Genetic data rarely adds much to accuracy past EHR data (exceptions: Warfarin, AMD)
Lessons (Continued)

- Data encoding not designed for learning
- Phenotyping is hard but is key to accuracy
  - Can’t assume ICD9 of 410 means heart attack
  - Machine learning now being used to phenotype
- Other Lessons of Pop Health/Epidemiology
  - Where to censor data
  - Reasoning about treatment effects, confounding
Lessons (Continued)

- Patient privacy is important but can hinder data sharing, accurate model learning, and even model sharing across institutions
  - Even if de-identified, risk of re-identification
  - Just publishing a resulting model and/or ROC curve could violate differential privacy
  - Text data and genomic data especially hard to de-identify, and some information is only in text
Lessons (Continued)

- For some tasks, causal inference is needed; for others, accurate prediction is enough
- Patient data is a timeline
  - CTBNs and related methods are relevant
  - Timeline methods good for feature construction, not as accurate as supervised learning so far
  - Not a regular time series – DBNs less relevant
The Other ADE Task

• **Given**: an EHR and a known ADE (a `<drug,condition>` pair)
  **Do**: learn model to predict (at prescription time) whether a patient will have the ADE if they take the drug

• **Given**: an EHR and a specified drug
  **Do**: find conditions caused by the drug (ADE)
# Observational Medical Outcomes Partnership 2011

<table>
<thead>
<tr>
<th>Drug</th>
<th>ACE Inhibitors</th>
<th>Amphotericin B</th>
<th>Antibiotics: erythromycins, sulfonamides, tetracyclines</th>
<th>Antiepileptics: carbamazepine, phenytoin</th>
<th>Benzodiazepines</th>
<th>Beta blockers</th>
<th>Bisphosphonates: alendronate</th>
<th>Tricyclic antidepressants</th>
<th>Typical antipsychotics</th>
<th>Wartarin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angioedema</td>
<td>Red</td>
<td>Blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aplastic Anemia</td>
<td>Blue</td>
<td>Red</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute Liver Injury</td>
<td>Blue</td>
<td>Blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bleeding</td>
<td>Blue</td>
<td>Blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hip Fracture</td>
<td>Blue</td>
<td>Blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospitalization</td>
<td>Blue</td>
<td>Green</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myocardial Infarction</td>
<td>Blue</td>
<td>Blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality after MI</td>
<td>Blue</td>
<td>Blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Renal Failure</td>
<td>Blue</td>
<td>Blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GI Ulcer Hospitalization</td>
<td>Blue</td>
<td>Blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Approach

- Search for events that occur more frequently after drug initiation than before

- Basic scoring function:
  \[ P(t_c > t_d | c, d) \]

- Normalize by dividing by:
  \[ \frac{P(t_c > t_d | C, d)}{P(t_c > t_D | c, D)} \]
OMOP Task ROC Curve: MSLR

AUC: 0.764
Relational Approach

Cox2 inhibitor(P,D) → hypertension(P)
older(P,55), vioxx(D)

ACE inhibitor

Warfarin
Cox2 Rules

- Found myocardial infarction (MI, or heart attack) association, and could have found it just two years into use.
- Found the Vioxx-specific rule for increased blood pressure in older people.
- Other rules just associated with reason for taking drug (indications).