Health Informatics

Lecture 9

Samantha Kleinberg
samantha.kleinberg@stevens.edu
Next week: journal club

• For all papers: read, and prepare to comment on
• For your paper:
  – Read the article (+ other references if needed for context)
  – Prepare a brief summary
  – Prepare questions to lead discussion
Format

• Brief overview of topic
• Summary of paper
  – What’s the main argument/hypothesis?
  – How is this supported?
    • Overview experiments, figures
• Discussion
  – Prepare questions

Note: Everyone must participate
Total of 30 min/paper
Phenotyping

Data-> cohort with particular traits
Uses

- Retrospective research
- Clinical trials
- Epidemiology/population health
- Discovering subgroups
Using EHRs for research

• First challenge is identifying cohort
  – Who has disease? Who doesn’t?
  – Accuracy affects everything that follows

• Usually based on manually defined rules
  – Iterative process
  – Extremely time consuming

• Clinical trials
  – Usually manual identification of eligibility
Mining electronic health records: towards better research applications and clinical care

Peter B. Jensen¹, Lars J. Jensen¹ and Søren Brunak¹,²
Usual process:
- Define criteria
- Evaluate results
- Iterate

Need some labeled data or ability to determine whether results are correct
Challenges

• Representation
  – How to describe/define a phenotype

• Complexity
  – Scale of data, complexity of definition

• Data quality/completeness/bias
  – Are the criteria we need there? Can we trust the data?

• Narrative text
  – Extracting criteria

• Portability
  – Does phenotype in one hospital work at another?
EXAMPLE: CHF case definition

Geisinger:
- 2 medication orders for CHF or
- 2 outpatient visits with CHF diagnosis or
- 1 medication order and 1 outpatient CHF diagnosis or
- CHF on problem list

Columbia:
- 2 ICD-9 codes for CHF or
- 1 ICD-9 code for CHF and 1 mention of CHF in note on the same date or
- 1 ICD-9 code for CHF and one medication typically indicated for CHF
<table>
<thead>
<tr>
<th>Operational criteria</th>
<th>Number of Framingham criteria</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>none</td>
<td>1322</td>
</tr>
<tr>
<td></td>
<td>1 minor, 0 major</td>
<td>475</td>
</tr>
<tr>
<td></td>
<td>2+ minor, 0 major</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>0 minor, 1 major</td>
<td>1123</td>
</tr>
<tr>
<td></td>
<td>1 minor, 1 major</td>
<td>764</td>
</tr>
<tr>
<td></td>
<td>1 major, 2+ minor</td>
<td>289</td>
</tr>
<tr>
<td></td>
<td>2+ major</td>
<td>2429</td>
</tr>
<tr>
<td>1 outpatient visit or 1 medication order</td>
<td></td>
<td>1303</td>
</tr>
<tr>
<td>2+ medication orders only</td>
<td></td>
<td>909</td>
</tr>
<tr>
<td>2+ outpatient visits only</td>
<td></td>
<td>852</td>
</tr>
<tr>
<td>1 outpatient visit and 1 medication order</td>
<td></td>
<td>222</td>
</tr>
<tr>
<td>Problem list only</td>
<td></td>
<td>481</td>
</tr>
<tr>
<td>Meet 2 of 3 criteria</td>
<td></td>
<td>1869</td>
</tr>
<tr>
<td>Meet all 3 criteria</td>
<td></td>
<td>861</td>
</tr>
</tbody>
</table>

Meets operational criteria and Framingham criteria, N=2294 (35%)
Meets operational criteria but not Framingham criteria, N=2900 (45%)
Meets Framingham criteria but not operational criteria, N=424 (7%)
Does not meet operational criteria or Framingham criteria, N=879 (14%)

Phenotyping algorithms

A review of approaches to identifying patient

Table 2  Summary of number of studies using different data sources across the top 10 phenotypes

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Demographics</th>
<th>Medications</th>
<th>Lab reports</th>
<th>Vitals</th>
<th>Clinical</th>
<th>Diagnosis</th>
<th>Treatment</th>
<th>Notes</th>
<th>Genomic</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer</td>
<td>12</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>9</td>
<td>10</td>
<td>4</td>
<td>9</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Diabetes</td>
<td>11</td>
<td>10</td>
<td>14</td>
<td>4</td>
<td>7</td>
<td>14</td>
<td>7</td>
<td>8</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Heart failure</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Rheumatoid arthritis</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cataract</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Drug side effect</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Asthma</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Peripheral arterial disease</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Hypertension</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Journal of Biomedical Informatics, (3) Proceedings of the Annual American Medical Informatics Association Symposium, and (4) Proceedings of Clinical Research Informatics Conference within the past 3 years was assessed for inclusion in the review. Only articles using automated techniques were included.

Results Ninety-seven articles met our inclusion criteria. Forty-six used natural language processing (NLP)-based techniques, 24 described rule-based systems, 41 used

...
Main approaches

• Logical rules
  – Where do rules come from? Guidelines, expert knowledge

• Text-based
  – Keywords, deeper semantic processing

• Statistical/ML(classification)
High throughput phenotyping

• Previously: laboriously define criteria for one phenotype at a time

• Limitations
  – Large-scale GWAS
  – Discovering new phenotypes
  – Stratification
  – Only as good as prior knowledge
Challenges?

• Missing data
• Bias
• Error
• Standardization
Open problems

• Phenotype detection
• Time
  – Disease state changing over time, patient population changing over time, medical process changing over time
• Interpretation
  – Making automated phenotypes understandable
• Evaluation
  – What’s the ground truth?
• Fragmented data
  – Nonclinical data
  – HIE
Pharmacovigilance

• Drug trials involve small populations

• What about
  – Rare side effects?
  – Interactions with other medications?
  – Interactions with foods?
  – Repurposing drugs?

• Postmarket analysis
How are ADEs reported?

• Clinical trials
  – Every event reported

• FDA AERS
  – Selected events submitted by patient, doctor

• But also indirectly
  – Searching for information
  – Evidence of side effect in EHR
Biomedical informatics goals

- Drug side effects
- Drug-drug interactions
- Dosing
- Ideal treatment
- Subgroup
- Finding new indications
Detecting Drug Interactions From Adverse-Event Reports: Interaction Between Paroxetine and Pravastatin Increases Blood Glucose Levels

NP Tatonetti¹,²,³, JC Denny⁴,⁵, SN Murphy⁶,⁷, GH Fernald¹,²,³, G Krishnan⁸, V Castro⁶, P Yue⁹, PS Tsau⁹, I Kohane⁷,¹⁰,¹¹, DM Roden⁵, and RB Altman²,³

¹Biomedical Informatics Training Program, Stanford University, Stanford, California, USA
²Department of Bioengineering, Stanford University, Stanford, California, USA
³Department of Genetics, Stanford University, Stanford, California, USA
⁴Department of Biomedical Informatics, Vanderbilt University, Nashville, Tennessee, USA
⁵Department of Medicine, Vanderbilt University, Nashville, Tennessee, USA
⁶Laboratory of Computer Science, Massachusetts General Hospital, Boston, Massachusetts, USA
⁷Harvard Medical School, Boston, Massachusetts, USA
⁸Stanford Center for Clinical Informatics, Stanford University, Stanford, California, USA
⁹Department of Medicine/Cardiovascular Medicine, Stanford University, Stanford, California, USA
¹⁰Children’s Hospital, Boston, Massachusetts, USA
¹¹Brigham and Women’s Hospital, Boston, Massachusetts, USA
• Many patients have comorbidities
  – Two common illnesses will lead to many drug-pairs in common
• One effect: raised glucose
  – T2DM increasing, major public health problem
• DDI may not be reported as such, need to look at co-occurrences
Methods

- 12,627 adverse event reports
  - 37 drugs
- 4 pairs
  - 3 were infrequent combo
- 1 combination left: Paroxetine, Pravastatin
  - Took glucose measurements before/during treatment for patients at Stanford
    - 374 on paroxetine, 449 on pravastatin, 8 on both
  - Replicated at Vanderbilt, and Partners
Web-scale pharmacovigilance: listening to signals from the crowd

Ryen W White, Nicholas P Tatonetti, Nigam H Shah, Russ B Altman, Eric Horvitz

ABSTRACT
Adverse drug events cause substantial morbidity and mortality and are often discovered after a drug comes to market. We hypothesized that Internet users may provide early clues about adverse drug events via their online information-seeking. We conducted a large-scale study of Web search log data gathered during 2010. We pay particular attention to the specific drug pairing of paroxetine and pravastatin, whose interaction was reported to cause hyperglycemia after the time period of the online logs used in the analysis. We also examine sets of drug pairs known to be associated with hyperglycemia and those not associated with hyperglycemia. We find that anonymized signals on drug case an interaction between paroxetine (an antidepressant) and pravastatin (a cholesterol-lowering drug), which was recently reported to create hyperglycemia. This association was extracted from the US Food and Drug Administration adverse event reporting system (AERS) using a data-mining algorithm that aggregates reports to identify drug–drug interactions. The finding was confirmed in a retrospective analysis of the electronic health records of three regionally distinct medical institutions and confirmed in a mouse model. We hypothesized that patients taking these two drugs might experience symptoms of hyperglycemia and may have conducted internet searches on these symptoms and concerns.
Google flu

Figure 2: A comparison of model estimates for the Mid-Atlantic Region (black) against CDC-reported ILI percentages (red), including points over which the model was fit and validated. A correlation of 0.85 was obtained over 128 points from this region to which the model was fit, while a correlation of 0.96 was obtained over 42 validation points. 95% prediction intervals are indicated.

FEVER PEAKS
A comparison of three different methods of measuring the proportion of the US population with an influenza-like illness.

Google's algorithms overestimated peak flu levels this year.
You Are What You Tweet: Analyzing Twitter for Public Health

Michael J. Paul and Mark Dredze
Human Language Technology Center of Excellence
Center for Language and Speech Processing
Johns Hopkins University
Baltimore, MD 21218
{mcpaul, mdredze}@cs.jhu.edu
<table>
<thead>
<tr>
<th>Ailment</th>
<th>Allergies</th>
<th>Depression</th>
<th>Aches/Pains</th>
<th>Cancer</th>
<th>Obesity</th>
<th>Flu</th>
<th>Dental</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prior</strong></td>
<td>Allergies</td>
<td>Anxiety</td>
<td>Back Pain</td>
<td>Breast Cancer</td>
<td>Diabetes</td>
<td>Flu</td>
<td>Oral Health</td>
</tr>
<tr>
<td><strong>Frequency</strong></td>
<td>6.4%</td>
<td>5.8%</td>
<td>10.8%</td>
<td>8.0%</td>
<td>2.3%</td>
<td>8.1%</td>
<td>4.6%</td>
</tr>
<tr>
<td><strong>General Words</strong></td>
<td>allergies, stop, eyes, allergic</td>
<td>help, dont, body, depression</td>
<td>body, head, need, hurts</td>
<td>cancer, pray, mom, shes</td>
<td>blood, doctor, high, meds</td>
<td>flu, “swine flu”, “flu shot”, dont</td>
<td>meds, killers, dentist, teeth</td>
</tr>
<tr>
<td><strong>Symptoms</strong></td>
<td>sneezing, cold, coughing</td>
<td>pain, anxiety, stomach</td>
<td>pain, aches, stomach</td>
<td>pain, sad, “breast cancer”</td>
<td>pressure, “high blood pressure”</td>
<td>fever, cold, “sore throat”</td>
<td>pain, toothache, sore</td>
</tr>
<tr>
<td><strong>Treatments</strong></td>
<td>medicine, benadryl, claritin</td>
<td>surgery, treatment, plastic</td>
<td>massage, “hot bath”, ibuprofen</td>
<td>surgery, hospital, “heart surgery”</td>
<td>hospital, diet, exercise</td>
<td>hospital, vaccine, medicine</td>
<td>braces, surgery, antibiotics</td>
</tr>
</tbody>
</table>
Separating Fact from Fear: Tracking Flu Infections on Twitter

Alex Lamb, Michael J. Paul, Mark Dredze
Human Language Technology Center of Excellence
Department of Computer Science
Johns Hopkins University
Baltimore, MD 21218
{alamb3, mpaull19, mdredze}@jhu.edu
<table>
<thead>
<tr>
<th>Class Name</th>
<th>Words in Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infection</td>
<td>getting, got, recovered, have, having, had, has, catching, catch, cured, infected</td>
</tr>
<tr>
<td>Possession</td>
<td>bird, the flu, flu, sick, epidemic</td>
</tr>
<tr>
<td>Concern</td>
<td>afraid, worried, scared, fear, worry, nervous, dread, dreaded, terrified</td>
</tr>
<tr>
<td>Vaccination</td>
<td>vaccine, vaccines, shot, shots, mist, tamiflu, jab, nasal spray</td>
</tr>
<tr>
<td>Past Tense</td>
<td>was, did, had, got, were, or verb with the suffix “ed”</td>
</tr>
<tr>
<td>Present Tense</td>
<td>is, am, are, have, has, or verb with the suffix “ing”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Self</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Awareness</td>
<td>23.15%</td>
<td>24.07%</td>
<td>47.22%</td>
</tr>
<tr>
<td>Infection</td>
<td>37.21%</td>
<td>15.57%</td>
<td>52.78%</td>
</tr>
<tr>
<td>Total</td>
<td>60.36%</td>
<td>39.64%</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: The distribution over labels of the data set. Infection tweets are more likely to be about the author (self) than those expressing awareness.
<table>
<thead>
<tr>
<th>Data</th>
<th>System</th>
<th>2009</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google</td>
<td>Flu Trends</td>
<td>0.9929</td>
<td>0.8829</td>
</tr>
<tr>
<td></td>
<td>ATAM</td>
<td>0.9698</td>
<td>0.5131</td>
</tr>
<tr>
<td></td>
<td>Keywords</td>
<td>0.9771</td>
<td>0.6597</td>
</tr>
<tr>
<td>Twitter</td>
<td>All Flu</td>
<td>0.9833</td>
<td>0.7247</td>
</tr>
<tr>
<td></td>
<td>Infection</td>
<td>0.9897</td>
<td>0.7987</td>
</tr>
<tr>
<td></td>
<td>Infection+Self</td>
<td>0.9752</td>
<td>0.6662</td>
</tr>
</tbody>
</table>

Figure 2: The Twitter flu rate for two years alongside the ILI rates provided by the CDC. The y-axes are not comparable between the two years due to differences in data collection, but we note that the 2011-12 season was much milder.
Main points

• Data is changing over time
• Publicizing results can change searcher behavior
• Correlation vs causation
  – Robustness
• Big data not necessarily good data
• How to discover new uses for drugs? (why important?)
• How to find out when drugs being used off-label?
Reprint PDF

The application of established drug compounds to new therapeutic indications, known as drug repositioning, offers several advantages over traditional drug development, including reduced development costs and shorter paths to approval. Recent approaches to drug repositioning use high-throughput experimental approaches to assess a compound’s potential therapeutic qualities. Here, we present a systematic computational approach to predict novel therapeutic indications on the basis of comprehensive testing of molecular signatures in drug-disease pairs. We integrated gene expression measurements from 100 diseases and gene expression measurements on 164 drug compounds, yielding predicted therapeutic potentials for these drugs. We recovered many known drug and disease relationships using computationally derived therapeutic potentials and also predict many new indications for these 164 drugs. We experimentally validated a prediction for the antiulcer drug cimetidine as a candidate therapeutic in the treatment of lung adenocarcinoma, and demonstrate its efficacy both in vitro and in vivo using mouse xenograft models. This computational method provides a systematic approach for repositioning established drugs to treat a wide range of human diseases.
• Lots of $$ and time to develop new drugs
• Can we repurpose old ones for new indications?

• 100 diseases, 164 drugs
  – 16000 pairings -> 2664 stat. sig

• Assumption: if drug leads to opposite expression from disease, could be useful treatment
with 2000 μM cimetidine exhibit a significant increase in TUNEL-positive (green) nuclei compared to vehicle (PBS)–treated control. The scale bar is 75 μM. (C) Results from a tumor xenograft experiment testing the efficacy of H2 agonist cimetidine in inhibiting the growth of A549 LA cell line tumors in SCID mice. Three treatment groups (25, 50, or 100 mg/kg per injection) and one control group (PBS) were used. Another group was treated with doxorubicin as a positive control. (D) Representative images of tumors treated with high dose of cimetidine (left) and control (right).
Next week

Journal club!