Medicine in the age of EHRs

Nigam Shah, MBBS, PhD
nigam@stanford.edu
Going from Medicine to Healthcare

• Big Data has a role in the science of medicine (biomedical research) and the practice of medicine (healthcare).

• Priorities are different for the science and the practice.

• The are areas of activity at the junction of the science and the practice.
Learning from a million patients

- **Problem**: 96% of medical care is best guesses

- **Opportunity**: Make Decisions based on what happened to people like you.

- **Solution**:
  - Find similar patients using all the data
  - See trajectories of similar patients, at the bedside

- **Main risk**: getting data
  - Access
  - Amount

---

Great Medical Mines
Stanford Medicine Mag, Spring 2014

Features

- Drugs
- Devices
- Procedures

Persons

- 9 million

Time

- 3 years
A 55 year old female of Vietnamese heritage with known asthma presents to her physician with new onset moderate hypertension.

**Intervention**

antihypertensives

**Outcome**

Diastolic pressure < 90 mm Hg
Answering clinical questions

Cole TS et al, Pediatr Rheumatol Online J. 2013 Dec 3;11(1):45
## Evidence-Based Medicine in the EMR Era

Jennifer Frankovich, M.D., Christopher A. Longhurst, M.D., and Scott M. Sutherland, M.D.

<table>
<thead>
<tr>
<th>Outcome or Risk Factor</th>
<th>Keywords Used to Conduct Expedited Electronic Search</th>
<th>Prevalence of Thrombosis</th>
<th>Relative Risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome — thrombosis</td>
<td>“Thrombus,” “Thrombosis,” “Blood clot”</td>
<td>10/98 (10)</td>
<td>Not applicable</td>
</tr>
<tr>
<td>Thrombosis risk factor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heavy proteinuria (&gt;2.5 g per deciliter)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present at any time</td>
<td>“Nephrosis,” “Nephrotic,” “Proteinuria”</td>
<td>8/36 (22)</td>
<td>7.8 (1.7–50)</td>
</tr>
<tr>
<td>Present &gt;60 days</td>
<td>“Urine protein”</td>
<td>7/23 (30)</td>
<td>14.7 (3.3–96)</td>
</tr>
<tr>
<td>Pancreatitis</td>
<td>“Pancreatitis,” “Lipase”</td>
<td>5/8 (63)</td>
<td>11.8 (3.8–27)</td>
</tr>
<tr>
<td>Antiphospholipid antibodies</td>
<td>“Aspirin”</td>
<td>6/51 (12)</td>
<td>1.0 (0.3–3.7)</td>
</tr>
</tbody>
</table>
Results of Electronic Search of Patient Medical Records (for a Cohort of 98 Pediatric Patients with Lupus) Focused on Risk Factors for Thrombosis Relevant to Our 13-Year-Old Patient with Systemic Lupus Erythematosus.*

<table>
<thead>
<tr>
<th>Outcome or Risk Factor</th>
<th>Keywords Used to Conduct Expedited Electronic Search</th>
<th>Prevalence of Thrombosis no./total no (%)</th>
<th>Relative Risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome — thrombosis</strong></td>
<td>“Thrombus,” “Thrombosis,” “Blood clot”</td>
<td>10/98 (10)</td>
<td>Not applicable</td>
</tr>
<tr>
<td><strong>Thrombosis risk factor</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heavy proteinuria (&gt;2.5 g per deciliter) Present at any time</td>
<td>“Nephrosis,” “Nephrotic,” “Proteinuria”</td>
<td>8/36 (22)</td>
<td>7.8 (1.7–50)</td>
</tr>
<tr>
<td>Present &gt;60 days</td>
<td>“Urine protein”</td>
<td>7/23 (30)</td>
<td>14.7 (3.3–96)</td>
</tr>
<tr>
<td>Pancreatitis</td>
<td>“Pancreatitis,” “Lipase”</td>
<td>5/8 (63)</td>
<td>11.8 (3.8–27)</td>
</tr>
<tr>
<td>Antiphospholipid antibodies</td>
<td>“Aspirin”</td>
<td>6/51 (12)</td>
<td>1.0 (0.3–3.7)</td>
</tr>
</tbody>
</table>

*Note: This table provides a summary of the results of an electronic search of patient medical records to identify risk factors for thrombosis in a cohort of 98 pediatric patients with lupus. The table includes keywords used for the search, the prevalence of thrombosis, and the relative risk with 95% confidence intervals.
Count positive, present mentions

- Keeping track of negation
- Keeping track of history, family history
Juvenile Idiopathic Arthritis and Uveitis

**Primary Cohort** (Juvenile Idiopathic Arthritis)  
ICD 9 codes  
696.0, 714.0, 714.2, 714.3, 714.9, 720.2, 720.9

**Terms:**  
Juvenile idiopathic arthritis, JIA  
Juvenile rheumatoid arthritis, JRA  
Psoriatic arthritis  
Juvenile spondyloarthropathy, spondyloarthritis, enthesitis related arthritis, sacroiliitis, reactive arthritis

**Outcome of interest** (Chronic Uveitis)  
ICD 9 codes  
364.00 (acute)  
364.10 (chronic)

**Terms:**  
Uveitis  
Iridocyclitis  
Iritis

**Patient factors associated with uveitis**  
ANA positive, positive ANA  
psoriasis  
allergic, allergy  
oligoarticular, oligo-onset, pauciarticular, pauci-onset, monoarthritis, monoarticular  
rheumatoid factor positive, rf positive, positive rheumatoid factor, positive rf

**Examples of allergy medications in clinical records:**  
Nasal steroids: Flonase, Nasacort  
Oral Antihistamines: Allegra, Zyrtec, Claritin, Clarinex, Benadryl, Xyzal  
Nasal antihistamines: Astelin  
Leukotriene inhibitors: Singulair  
Decongestant: Sudafed
Peripheral artery disease (PAD): obstruction of infra-renal abdominal aorta and lower extremity arteries

➔ Cilostazol
Peripheral Artery Disease

- No difference in MACE among patients taking Cilostazol
- Uncovered a “practice-pattern” that clinicians speculated about
Finding “similar patients”

14 methods

- expertPS
- hdPS
- lassoPS
- rfPS
- Euclidean
- Jaccard
- Dice
- Cosine
- Pearson
- Spearman
- lassoMV
- rdmSamp1x
- rdmBagging
- rdmJackknife

20 outcomes →

Beta coefficients

- Risky
- Protective

Circle size is proportional to confidence intervals
Insights from the data

Turning text into safety signals

Patient-feature matrix
• rows = patients
• columns = medical concepts

Can detect 6 out of 9 recalls in the past decade
Gold Standard based on the EU-ADR validation set:
- 28 positive test cases
- 165 negative test cases
- 12 events
- 78 drugs

AERS: AUC 0.72 – 0.83*
Notes: AUC 0.75 – 0.80

New discovery or false positive?

T/SICU Nursing Admission Note:
This is a 31 year old male s/p seizure disorder who fell 15-20 feet on 1/10-17, sustaining a left femur fracture, right humeral fracture, and left distal radius fracture. He requires intubation for 48 hours post-op. His past medical history is significant only for seizure disorder, and his only medication is diazepam. He has no known allergies.

Patient-feature matrix
- rows = patients
- columns = medical concepts

- Table showing patient features

Graph showing exposure and outcome over time.
Identifying Off-label drug use

Classifier for "used-to-treat" relationship

EMR

Prior knowledge

8,861 positive examples

34,973 negative examples

Clinical Notes

Medi-span, Drugbank

Prior knowledge

Example features:
- Co-mentions
- Drug-first fraction

Example features:
- Similarity with on-label uses
- Fraction of uses that are approved

Prioritize 407 well-supported usages

2,362,950 drug-indication pairs

Cost

Risk

32

Low

High

17

35

Medi-span, Drugbank

Classifier for "used-to-treat" relationship

Example features:
- Similarity with on-label uses
- Fraction of uses that are approved

8,861 positive

27,938 negative training examples

7,112 positive and

1,749 positive and

27,938 negative test examples

1,749 positive and

7,035 negative test examples

Evaluate in test set

PPV

0.952

Spec

0.990

Recall

0.758

F1

0.844
Building the graph of medicine

**Text annotation**

**positive-present mentions**

**Concept Occurrence Matrix**

<table>
<thead>
<tr>
<th>Drugs</th>
<th>Diseases</th>
<th>Devices</th>
<th>Procedures</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="chart.png" alt="Concept Occurrence Matrix Chart" /></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**2-by-2 tables**

- $f(X) = B + A$
- $f(X, Y) = A$
- $f(Y) = A + C$

**conditional probabilities**

$P(X, Y) = P(X) \cdot P(Y | X)$

**Patient Timeline Bins**

- ADMITTING DIAGNOSIS: Syncope.
- CHIEF COMPLAINT: Vertigo or dizziness.
- HISTORY OF PRESENT ILLNESS: This is an (XX)-year-old male with a past medical history of coronary artery disease, CABG done a few years ago, atrial fibrillation, peripheral arterial disease, peripheral neuropathy, recently retired one year ago secondary to leg pain. The patient came to the ER for an episode of vertigo while reaching for some books. The patient was able to reach the books, to support self, but did not have any syncope. No nausea or vomiting. No chest pain. No shortness of breath. Came to ER and had a CT head, which was within normal limits. The impression was atrophy with old ischemic changes but no acute intracranial findings. No focal weakness, headache, vision changes or speech changes. The patient has had similar episodes since one year. Peripheral neuropathy since one year and not relieved with multiple medications. The patient also complains of weight loss of 25 pounds in the last 6 months.
Making predictions
Predicting ‘problem wounds’

Total: 1,079 features  ~1 - 14% outcomes
Advanced care decisions for wound care specialists

Screening for referral to wound care center
Early results: Patient trajectories

- Consider a patient record as a sequence of events comprised of mentions of drugs, diseases, procedures, and devices.

- Given such event sequences
  1. categorize sequences into groups based on how they evolve
  2. divide each sequence into stages.

Results on Chronic Kidney Disease

- 30 % of CKD Patients do not manifest albuminuria.
- We can learn the top symptom at each stage.
Interested?

nigam@stanford.edu
Acknowledgements

Group Members (13):
- **Fellows:** Suzanne Tamang, Yen Low, Alison Callahan, Elsie Gyang, Juan Banda
- **BMI Students:** Ken Jung, David Odgers, Sarah Poole, Vibhu Agarwal, Will Chen, Tanya Podchiyska
- **Med Students:** Greg Gaskin

**Alums:** Anna Bauer-Mehren (Roche), Srini Iyer (Facebook), Amogh Vasekar (Citrix), Sandy Huang (Berkeley). Paea Lependu (PCCI), Rave Harpaz (Oracle), Sam Finlayson (Harvard), Rick Barber (NunaHealth)

**Funding:**
- NLM – R01 LM011369 (PI, Nigam Shah)
- NIGMS – R01 GM101430 (PI, Nigam Shah)
- NHGRI – U54 HG004028 (PI, Mark Musen)
- NINDS – R01 NS080145 (PI, Vance Lemmon)
- NIMH – P50 MH094267 (PI, Andrey Rzhetsky)
- Seed funds – Dept. of Medicine, TRAM, Population Health Sciences, Clinical Excellence Research Center
- Fellowships – Med Scholars, Siebel Scholars Foundation, Stanford Graduate Fellowship
- Industry – Apixio, CollabRx, Healogics, J&J[OHDSI]

STRIDE Team members: Todd Ferris, Tanya Podchiyska
BMIR-IT: Alex Skrenchuk