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



 \leftarrow Features \rightarrow

Green button-using aggregate patient data



Answering clinical questions

Cole TS et al, Pediatr Rheumatol Online J. 2013 Dec 3;11(1):45 Bauer-Mehren et al, PLoS One. 2013 May 23;8(5):e63499

Evidence-Based Medicine in the EMR Era

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

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



http://clever.stanford.edu



3

Count positive, present mentions

- Keeping track of negation
- Keeping track of history, family history

Sens	73%
Spec	96%
PPV	83%
F1	78%

Juvenile Idiopathic Arthritis and Uveitis

	Event / Outcome concepts		Root]	
Juvenile Idiopathic Art Juvenile spondyloarth	Juvenile Idiopathic Arthritis	Root Term A Term B Term D Ontobgy	Root Term A Term B		
	Juvenile spondyloarthropathy				
4	:	Τ.	Ontology Roo	_	
	Uveitits		Term B		
	Iridocylitis		CTerm D Ontology		Oral Antinistamines

Primary Cohort (Juvenile Idiopathic Arthritis) ICD 9 codes 696.0, 714.0, 714.2, 714.3, 714.9, 720.2, 720.9	Outcome of interest (Chronic Uveitis) ICD 9 codes 364.00 (acute) 364.10 (chronic)	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
Terms: Juvenile idiopathic arthritis, JIA Juvenile rheumatoid arthritis, JRA Psoriatic arthritis Juvenile spondyloarthropathy, spondyloarthritis, enthesitis related arthritis, sacroiliitis, reactive arthritis	Terms: Uveitis Iridocyclitis Iritis	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

Peripheral artery disease (PAD): obstruction of infra-renal abdominal aorta and lower extremity arteries





Finding "similar patients"



Insights from the data

LePendu et al, Clin Pharmacol Ther. 2013 Jun;93(6):547-55. Jung et al, PLoS One, 2014 February, 9(2): e89324

Turning text into safety signals

T/SICU Nursing Admission Note:

Text clinical note

This is a 31 year old male s/p seizure on issues size cauting sei 15-20 feet on [**09-17**] now presenting to the 7/SICU post surgical repair of multiple facial fractures, right mandibular fracture, and left distal radius fracture. He needs to remain intubated for 48 hours post-op. His past medical history is significant only for seizure disorder, and his only medication is depakote. He has no known allergies.



Patient-feature matrix

- rows = patients
- columns = medical concepts

	MI	No MI
Vioxx	а	b
No Vioxx	С	d



Can detect 6 out of 9 recalls in the past decade

Overall Performance

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

* Harpaz et al. Performance of Pharmacovigilance Signal Detection Algorithms for the FDA Adverse Event Reporting System Nature - Clinical Pharmacology & Therapeutics 2013

New discovery or false positive?



Identifying Off-label drug use



Building the graph of medicine



Concept Occurrence Matrix

Making predictions

Predicting 'problem wounds'



Total: 1,079 features

~1 - 14% outcomes



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.

Symptom	Ground-truth	Ours	Baseline
secondary pulmonary hypertension	2	2.65	3.45
proteinuria	3	3.19	2.94
hyperphosphatemia	4	3.99	3.71
acidosis	4	3.97	3.21

Table 5: Examples of symptoms and their ground-truth stages.

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

THE AGE OF DATA-DRIVEN MEDICINE

Mining structured and unstructured health records as a form of medical research

