Machine Learning for Healthcare

What makes healthcare unique?

Prof. David Sontag & Pete Szolovits

CSAIL

INSTITUTE FOR MEDICAL ENGINEERING & SCIENCE

HEALTH SCIENCES & TECHNOLOGY

The Problem

- Cost of health care expenditures in the US are over $3 trillion, and rising
- Despite having some of the best clinicians in the world, chronic conditions are
  - Often diagnosed late
  - Often inappropriately managed
- Medical errors are pervasive

Outline for today's class

1. **Brief history of AI and ML in healthcare**
2. Why *now*?
3. Examples of how ML will transform healthcare
4. What is *unique* about ML in healthcare?

1970's: MYCIN expert system

- 1970's (Stanford): MYCIN expert system for identifying bacteria causing severe infections
- Proposed a good therapy in ~69% of cases. Better than infectious disease experts

information flow.

to help build a knowledge base, to explain a line of reasoning, and sc
The knowledge base is the program's store of facts and associa
"knows" about a subject area such as medicine. A critical design d
is how such knowledge is to be represented within the program. Th
many choices, in general. For MYCIN, we chose to represent kno
mostly as conditional statements, or rules, of the following form:

IF: There is evidence that A and B are true,

THEN: Conclude there is evidence that C is true.

This form is often abbreviated to one of the following:

If A and B, then C

Dialogue interface

I am ready

** THIS IS A 26 YEAR OLD MALE PATIENT

My understanding is:
The age of the patient is 26
The sex of the patient is male

** FIVE DAYS AGO, HE HAD RESPIRATORY-TRACT SYMPTOMS

What is his name?

** JO

My understanding is:
The name of the patient is Jo
Respiratory-tract is one of the symptoms that the patient had

** A COUPLE OF DAYS BEFORE THE ADMISSION, HE HAD A MALAISE

Please give me the date of admission

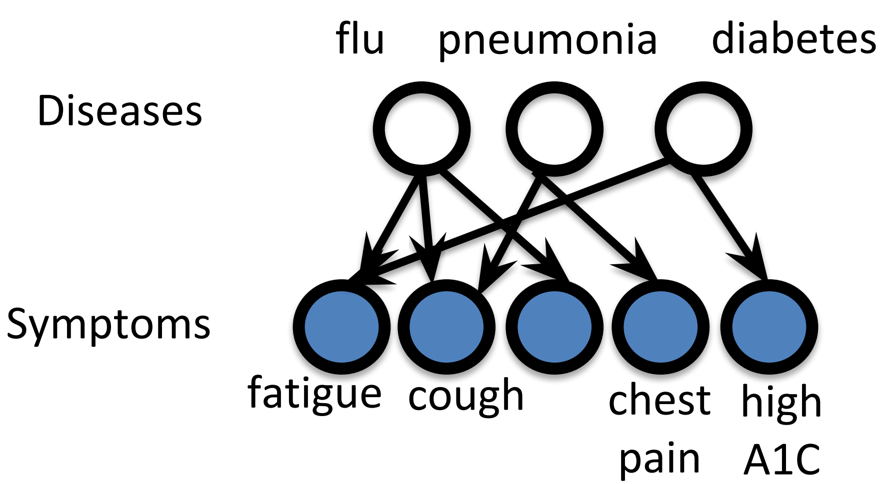
** MARCH 12, 1979

My understanding is:
The patient was admitted at the hospital 3 days ago
Malaise is one of the symptoms that the patient had 5 days ago

FIGURE 33-1 Short sample dialogue. The physician's inputs appear in capital letters after the double asterisks.

1980's: INTERNIST-1/QMR model

- 1980's (Univ. of Pittsburgh): INTERNIST-1/Quick Medical Reference
- Diagnosis for internal medicine



Probabilistic model relating:

570 binary disease variables
4,075 binary symptom variables
45,470 directed edges

Elicited from doctors:
15 person-years of work

Led to advances in ML & AI (Bayesian networks, approximate inference)

Problems:
1. Clinicians entered symptoms *manually*
2. Difficult to maintain, difficult to generalize

[Miller et al., '86, Shwe et al., '91]

1980's: automating medical discovery

RX PROJECT: AUTOMATED KNOWLEDGE ACQUISITION

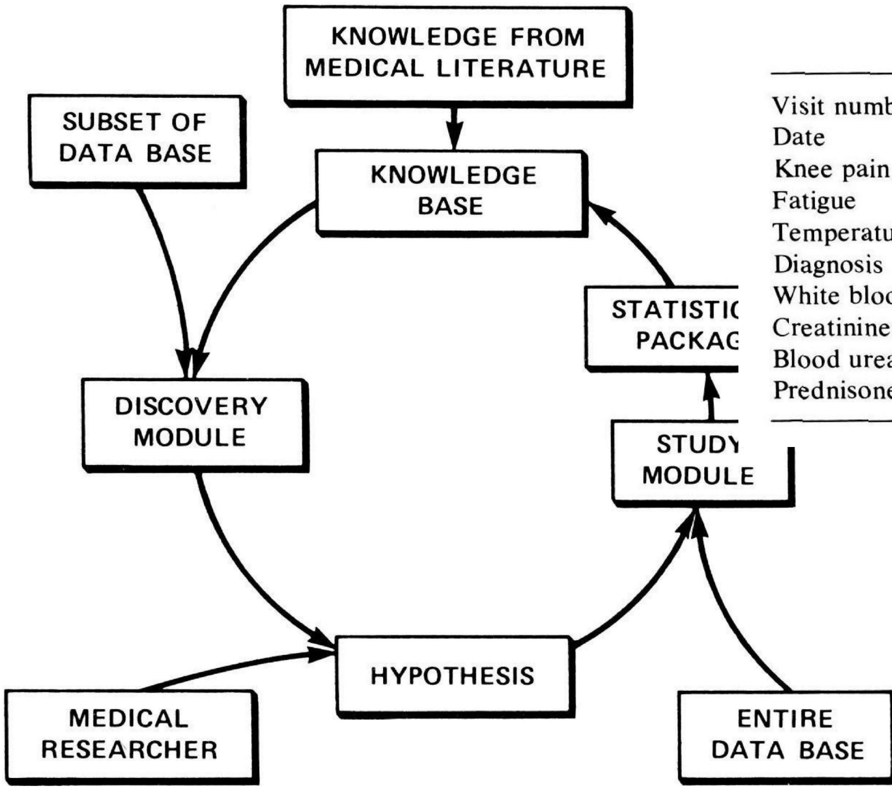


TABLE 1

HYPOTHETICAL TIME-ORIENTED RECORD FOR ONE PATIENT

Visit number	1	2	3
Date	January 17, 79	June 23, 79	July 1, 79
Knee pain	Severe	Mild	Mild
Fatigue	Moderate	—	Moderate
Temperature	38.5	37.5	36.9
Diagnosis	Systemic lupus		
White blood count	3500	4700	4300
Creatinine clearance	45	—	65
Blood urea nitrogen	36	33	—
Prednisone	30	25	20

Discovers that prednisone elevates cholesterol
(Annals of Internal Medicine, '86)

[Robert Blum, "Discovery, Confirmation and Incorporation of Causal Relationships from a Large Time-Oriented Clinical Data Base: The RX Project". Dept. of Computer Science, Stanford. 1981]

1990's: neural networks in medicine

- Neural networks with clinical data took off in 1990, with 88 new studies that year
- Small number of features (inputs)
- Data often collected by chart review

FIGURE 2. A multilayer perceptron. This is a two-layer perceptron with four inputs, four hidden units, and one output unit.

Problems:
1. Did not fit well into clinical workflow
2. Hard to get enough training data
3. Poor generalization to new places

[Penny & Frost, Neural Networks in Clinical Medicine. Med Decis Making, 1996]

Table 1 • 25 Neural Network Studies in Medical Decision Making*

Subject	No. of Examples		P†	Network	D‡	Accuracy§	
	Training	Test				Neural	Other
Breast cancer[4]	57	20	60	9–15–2	0.6	80	75
Vasculitis[2]	404	403	73	8–5–1	8.0	94	—
Myocardial infarction[6]	351	331	89	20–10–10–1	1.1	97	84
Myocardial infarction[8]	356	350	87	20–10–10–1	1.1	97	94
Low back pain[11]	100	100	25	50–48–2	0.2	90	90
Cancer outcome[13]	5,169	3,102	—	54–40–1	1.4	0.779	0.776
Psychiatric length of stay[17]	957	106	73	48–400–4	0.2	74	76
Intensive care outcome[23]	284	138	91	27–18–1	0.5	0.82	0.82
Skin tumor[21]	150	100	80	18	—	80	90
Evoked potentials[35]	100	67	52	14–4–3	3.8	77	77
Head injury[47]	500	500	50	6–3–3	20	66	77
Psychiatric outcome[54]	289	92	60	41–10–1	0.7	79	—
Tumor classification[55]	53	6	38	8–9–3	1.4	99	88
Dementia[57]	75	18	19	80–10–7–7	0.6	61	—
Pulmonary embolism[59]	607	606	69	50–4–1	2.9	0.82	0.83
Heart disease[62]	460	230	54	35–16–8–2	3	83	84
Thyroid function[62]	3,600	1,800	93	21–16–8–3	22	98	93
Breast cancer[62]	350	175	66	9–4–4–2	10	97	96
Diabetes[62]	384	192	65	8–4–4–2	12	77	75
Mycardial infarction[63]	2,856	1,429	56	291–1	9.8	85	—
Hepatitis[65]	39	42	38	4–4–3	3.3	74	79
Psychiatric admission[76]	319	339	85	53–1–1	6.0	91	—
Cardiac length of stay[83]	713	696	73	15–12–1	3.5	0.70	—
Anti-cancer agents[89]	127	141	25	60–7–6	1.5	91	86
Ovarian cancer[91]	75	98	—	6–6–2	2.6	84	81
MEDIAN VALUE	350	175	71	20	2.8		

*For reference citations, see the reference list

†P = prior probability of most prevalent category.

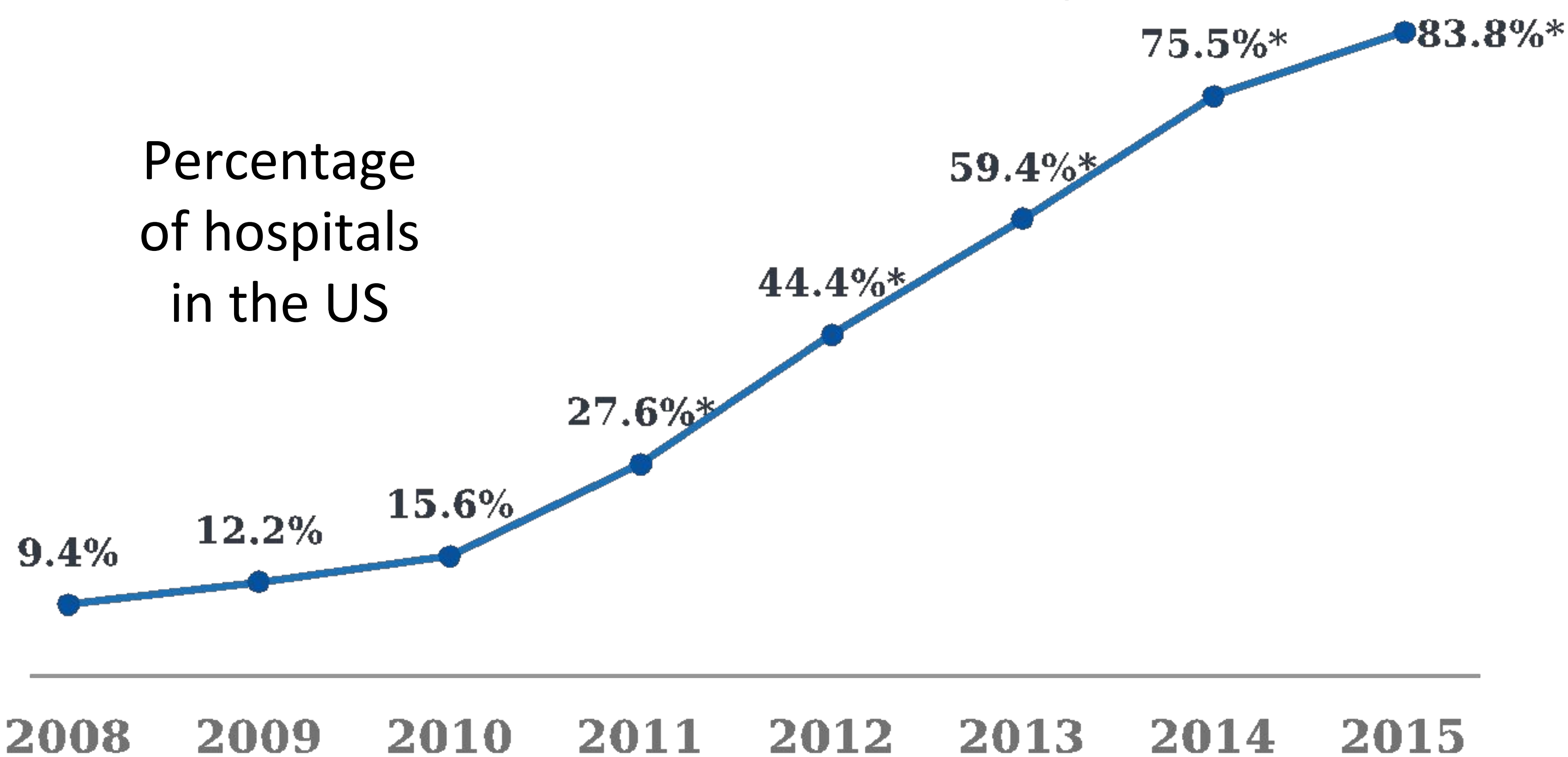
‡D = ratio of training examples to weights per output.

§A single integer in the accuracy column denotes percentage overall classification rate and a single real number between 0 and 1 indicates the AUROCC value. Neural = accuracy of neural net, Other = accuracy of best other method.

Outline for today's class

1. Brief history of AI and ML in healthcare
2. **Why *now*?**
3. Examples of how ML will transform healthcare
4. What is *unique* about ML in healthcare?
5. Overview of class syllabus

The Opportunity: Adoption of Electronic Health Records (EHR) has increased 9x in US since 2008



[Henry et al., ONC Data Brief, May 2016]

Large datasets

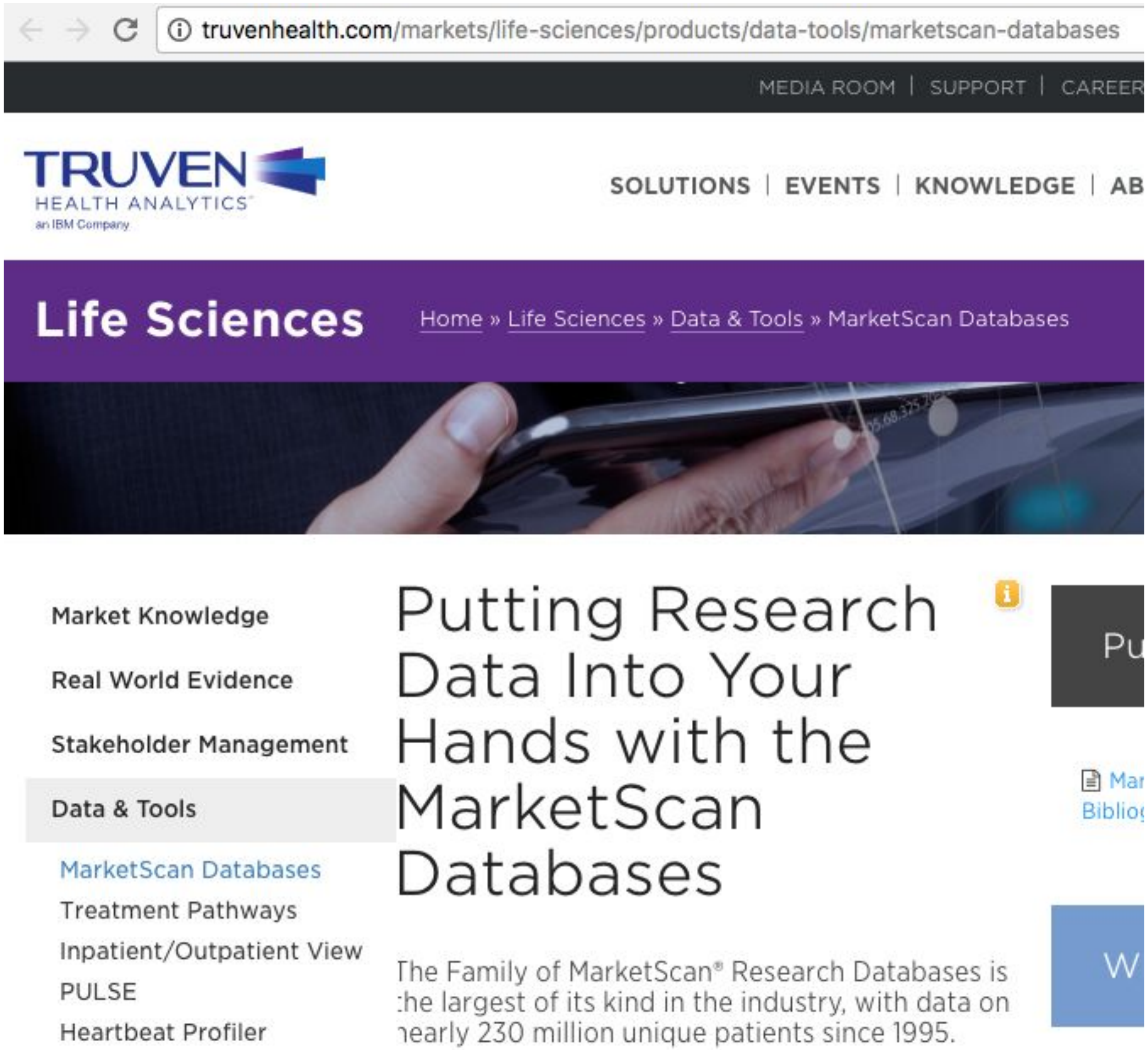
Massachusetts Institute of Technology

Laboratory for Computational Physiology

De-identified health data from ~40K critical care patients

Demographics, vital signs, laboratory tests, medications, notes, …

Large datasets



"Data on nearly 230 million unique patients since 1995"

$$$

Large datasets

President Obama's initiative to create a 1 million person research cohort

Core data set:

- Baseline health exam
- Clinical data derived from electronic health records (EHRs)
- Healthcare claims
- Laboratory data

[Precision Medicine Initiative (PMI) working Group Report, Sept. 17 2015]

Diversity of digital health data

Standardization

- Diagnosis codes: ICD-9 and ICD-10 (International Classification of Diseases)

…

ICD-9 codes 290–319: mental disorders

ICD-9 codes 320–359: diseases of the nervous system

ICD-9 codes 360–389: diseases of the sense organs

ICD-9 codes 390–459: diseases of the circulatory system

ICD-9 codes 460–519: diseases of the respiratory system

ICD-9 codes 520–579: diseases of the digestive system

ICD-9 codes 580–629: diseases of the genitourinary system

ICD-9 codes 630–679: complications of pregnancy, childbirth,

…

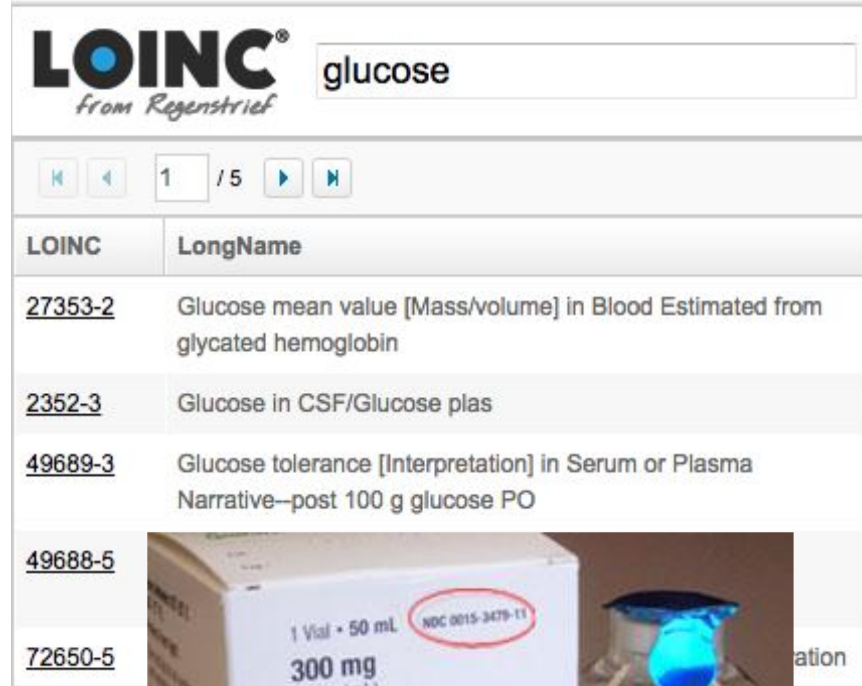
…

[https://en.wikipedia.org/wiki/List_of_ICD-9_codes]

[https://blog.curemd.com/the-most-bizarre-icd-10-codes-infographic/]

Standardization

- Diagnosis codes: ICD-9 and ICD-10 (International Classification of Diseases)
- Laboratory tests: LOINC codes
- Pharmacy: National Drug Codes (NDCs)
- Unified Medical Language System (UMLS): millions of medical concepts



[http://oplinc.com/newsletter/index_May08.htm]

Standardization

FHIR

Level 1 Basic framework on which the specification is built

Foundation	Base Documentation, XML, JSON, REST API + Search, Data Types, Extensions

Level 2 Supporting Implementation, and binding to external specifications

Implementer Support	**Security & Privacy**	**Conformance**	**Terminology**	**Linked Data**
Downloads, Common Use Cases, Testing	Security, Consent Provenance AuditEvent	StructureDefinition, CapabilityStatement, ImplementationGuide, Profiling	CodeSystem, ValueSet, ConceptMap, Terminology Svc	RDF

Level 3 Linking to real world concepts in the healthcare system

Administration	Patient, Practitioner, Device, Organization, Location, Healthcare Service

Level 4 Record-keeping and Data Exchange for the healthcare process

Clinical	**Diagnostics**	**Medications**	**Workflow**
Allergy, Problem, CarePlan, DetectedIssue, RiskAssessment, etc.	Observation, Report, Specimen, ImagingStudy,Genomics,etc	Order, Dispense, Administration, Statement, Immunization, etc.	Task, Appointment, Schedule, Referral, PlanDefinition, etc.

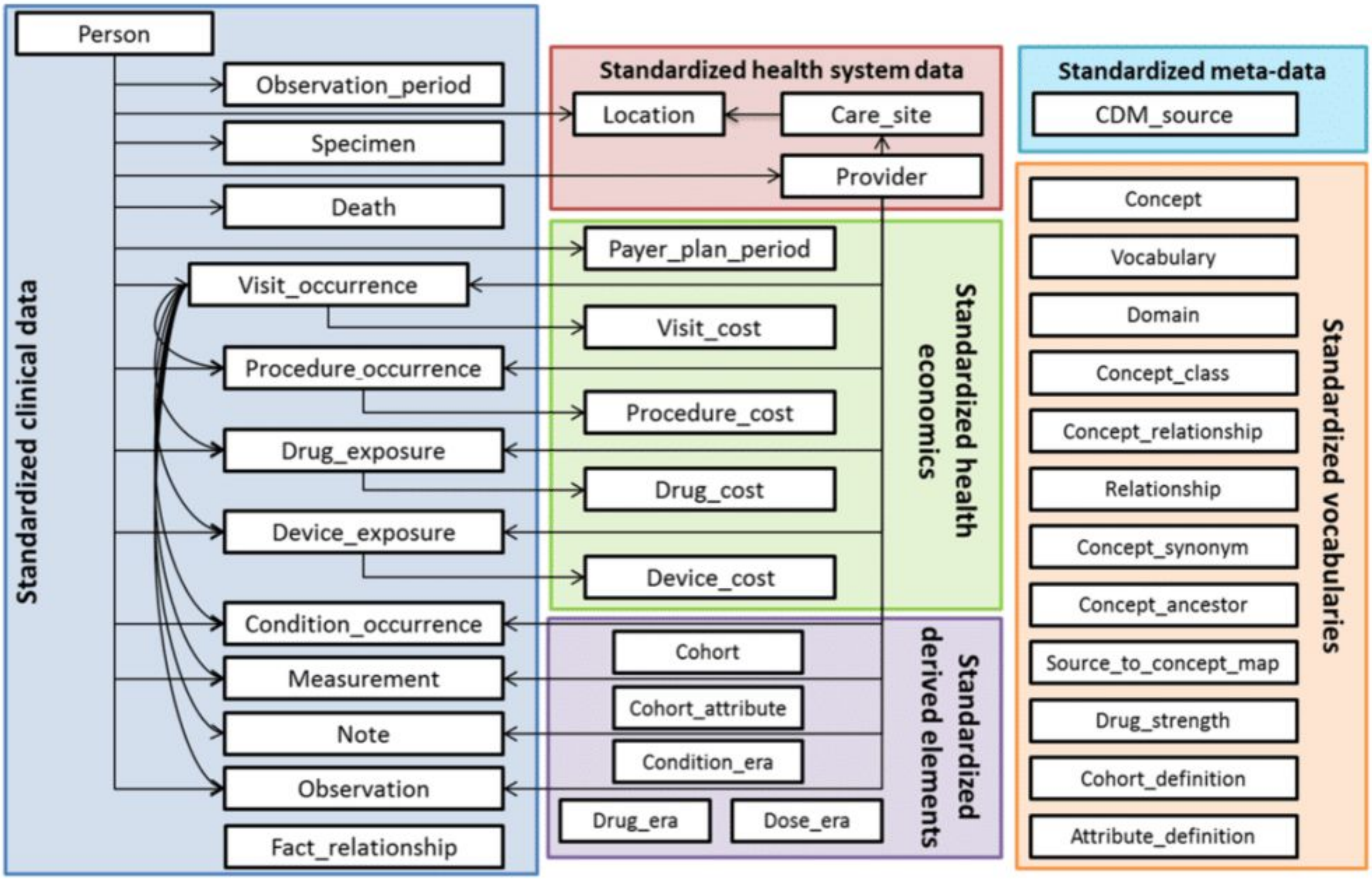
Financial
Claim, Account, Coverage, Claim, EligibilityRequest, ExplanationOfBenefit, etc.

Level 5 Providing the ability to reason about the healthcare process

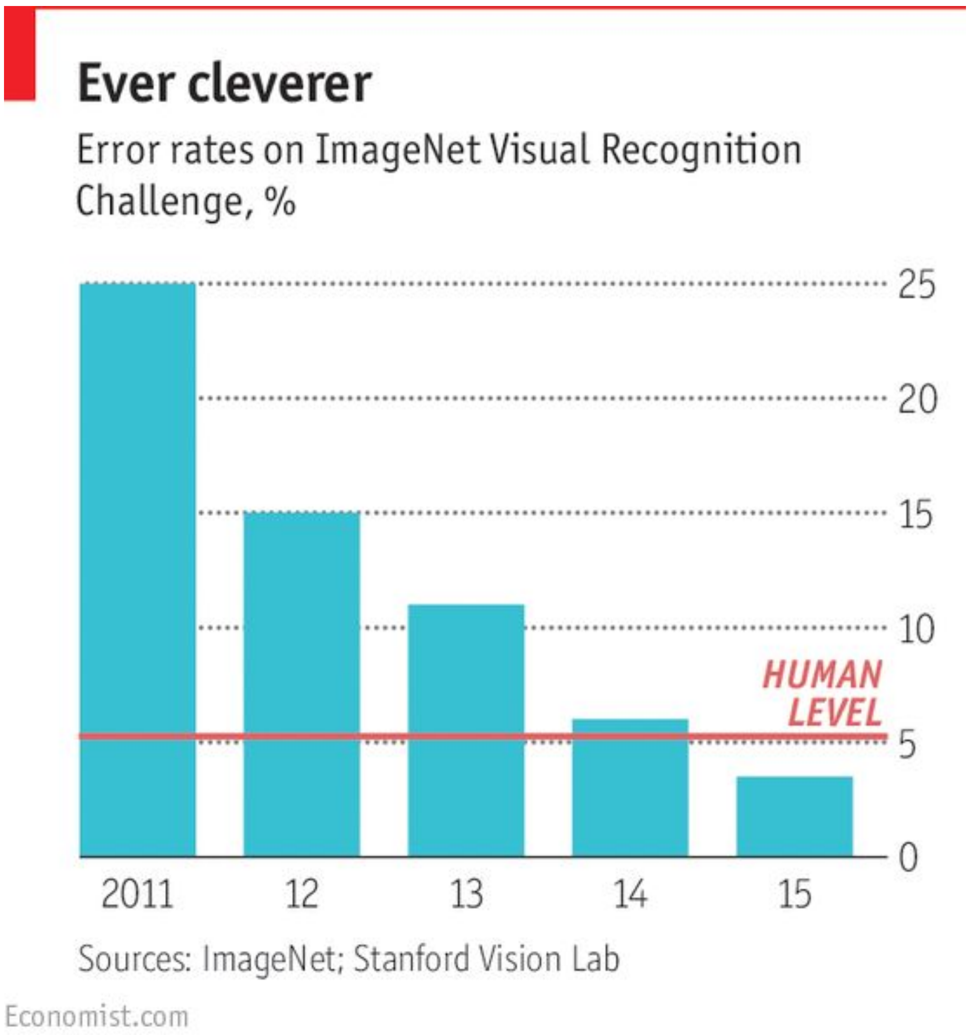
Clinical Reasoning	Library, ServiceDefinition & GuidanceResponse, Measure/MeasureReport, etc

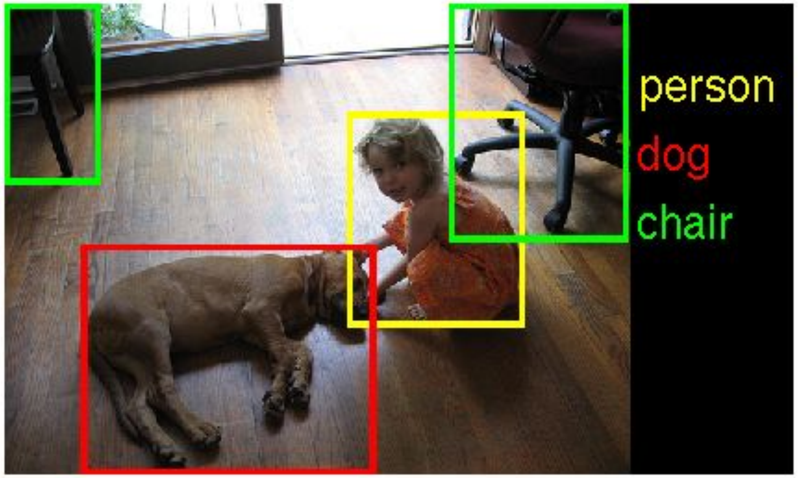
Standardization

OMOP Common Data Model v5.0



Breakthroughs in machine learning





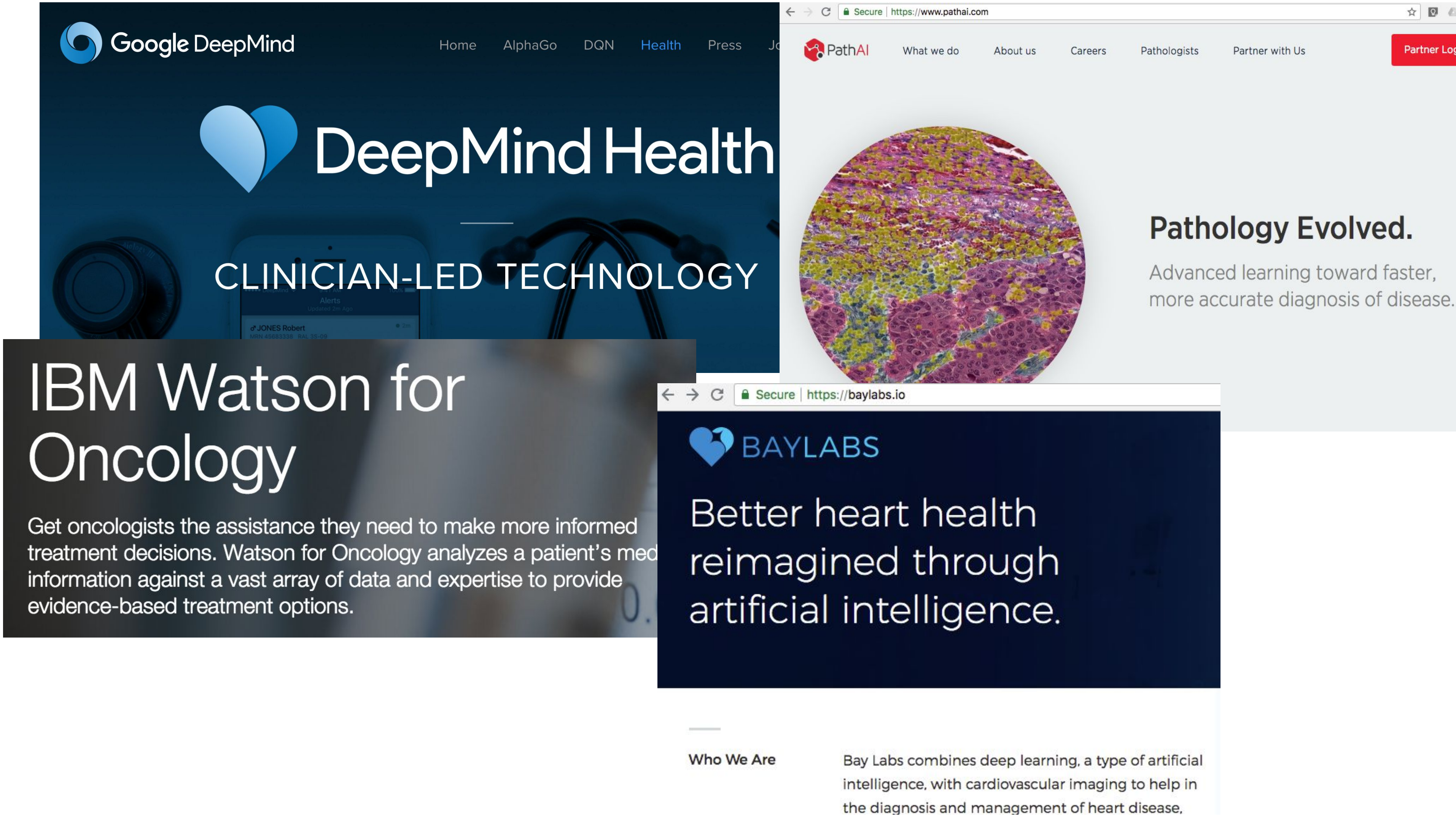
Why now?

- Big data
- Algorithmic advances
- Open-source software

Breakthroughs in machine learning

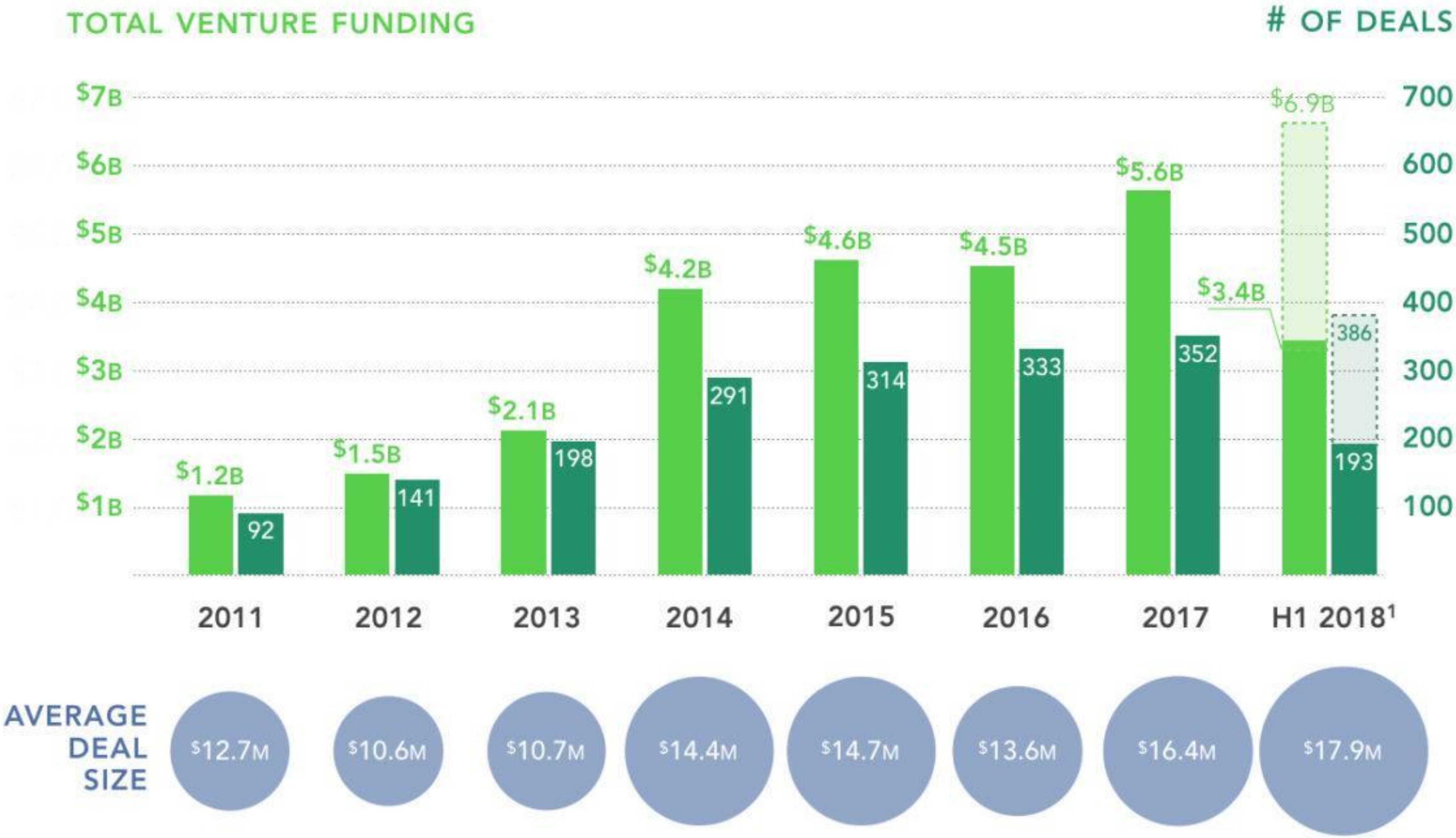
- Major advances in ML & AI
  - Learning with high-dimensional features (e.g., l1-regularization)
  - Semi-supervised and unsupervised learning
  - Modern deep learning techniques (e.g. convnets, variants of SGD)
- Democratization of machine learning
  - High quality open-source software, such as Python's scikit-learn, TensorFlow, Torch, Theano

Industry interest in ML & healthcare



DIGITAL HEALTH FUNDING

2011-H1 2018



Source: Rock Health Funding Database

1: Shadowed portion shows projections for entire year of 2018, assuming current funding pace continues.

Note: Only includes U.S. deals >$2M; data through June 30, 2018

106 STARTUPS TRANSFORMING HEALTHCARE WITH AI

Industry interest in ML & healthcare

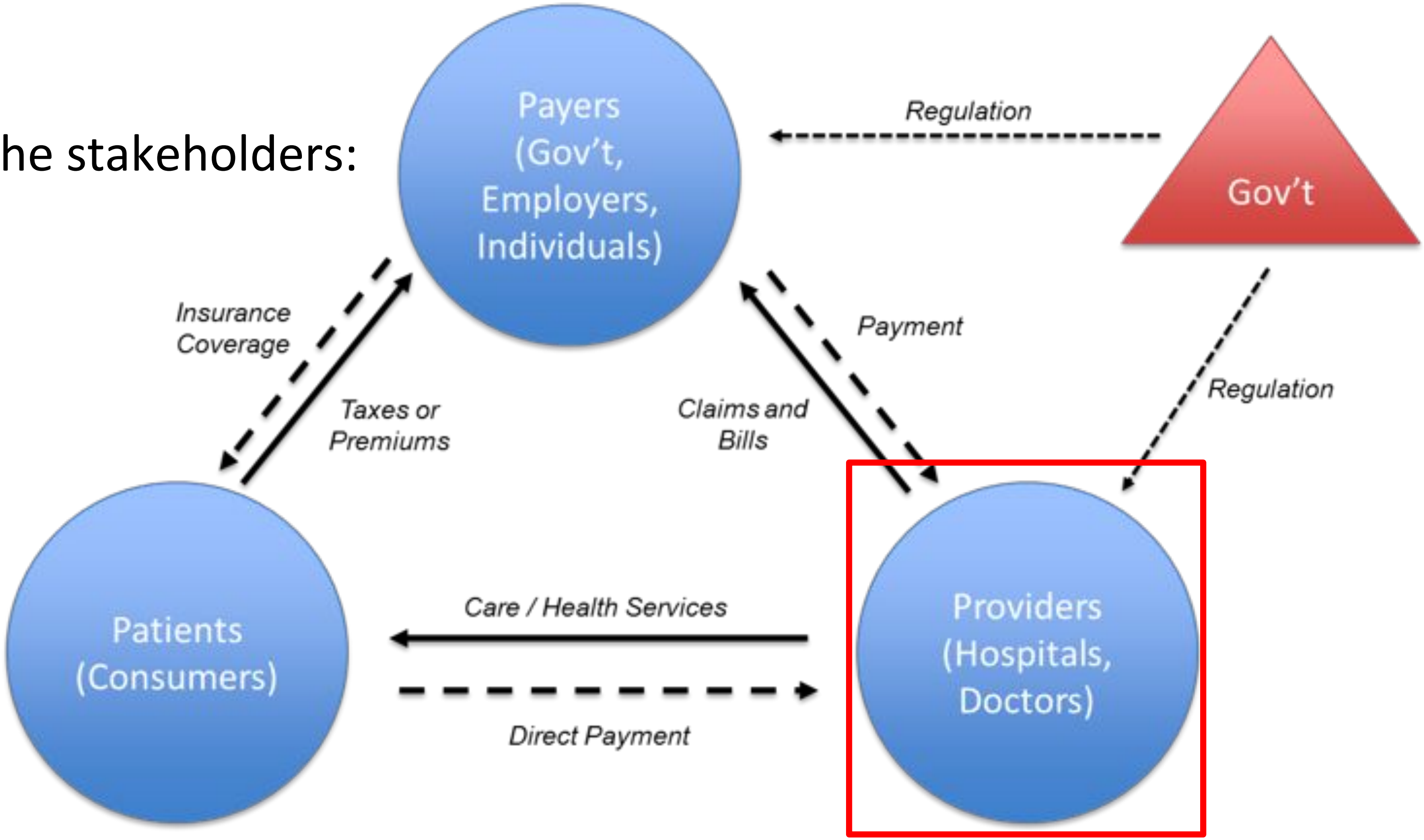
- Major acquisitions to get big data for ML:
  - Merge ($1 billion purchase by IBM, 2015)
    medical imaging
  - Truven Health Analytics ($2.6 billion purchase by IBM, 2016)
    health insurance claims
  - Flatiron Health ($1.9 billion purchase by Roche, 2018)
    electronic health records (oncology)

Outline for today's class

1. Brief history of AI and ML in healthcare
2. Why *now*?
3. **Examples of how ML will transform healthcare**
4. What is *unique* about ML in healthcare?
5. Overview of class syllabus

ML will transform every aspect of healthcare

The stakeholders:



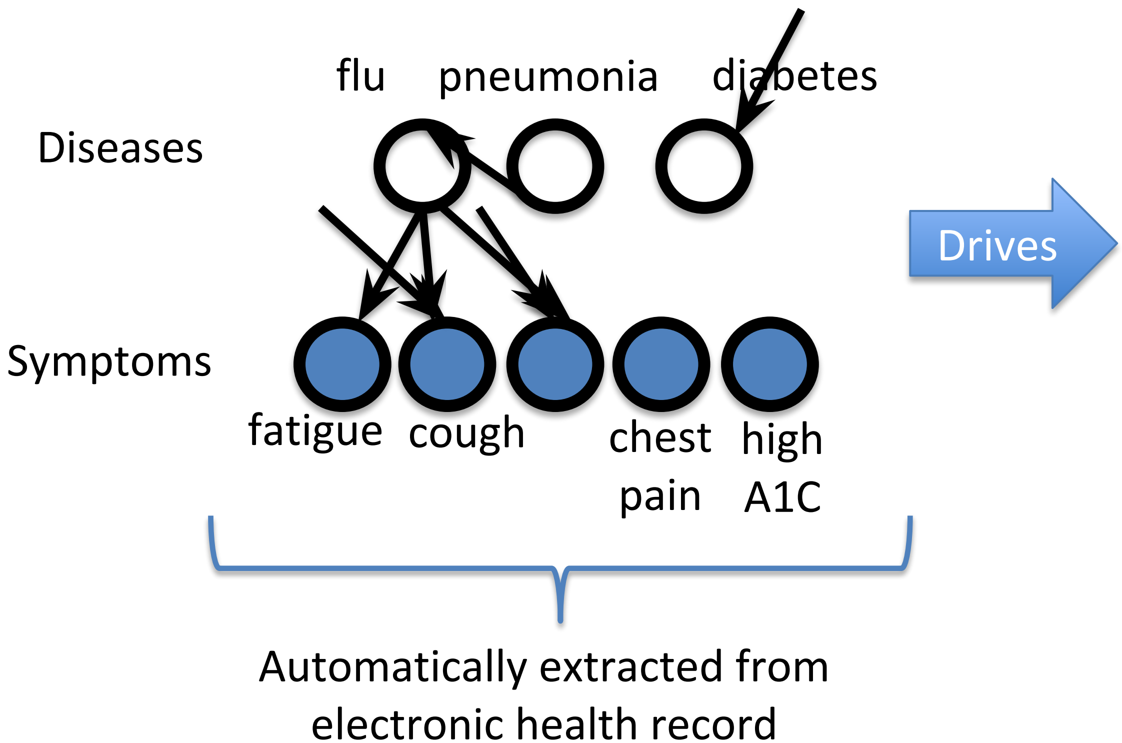
Source for figure:
http://www.mahesh-vc.com/blog/understanding-whos-paying-for-what-in-the-healthcare-industry

Emergency Department:

- **Limited resources**
- **Time sensitive**
- **Critical decisions**

What will the ER of the future be like?

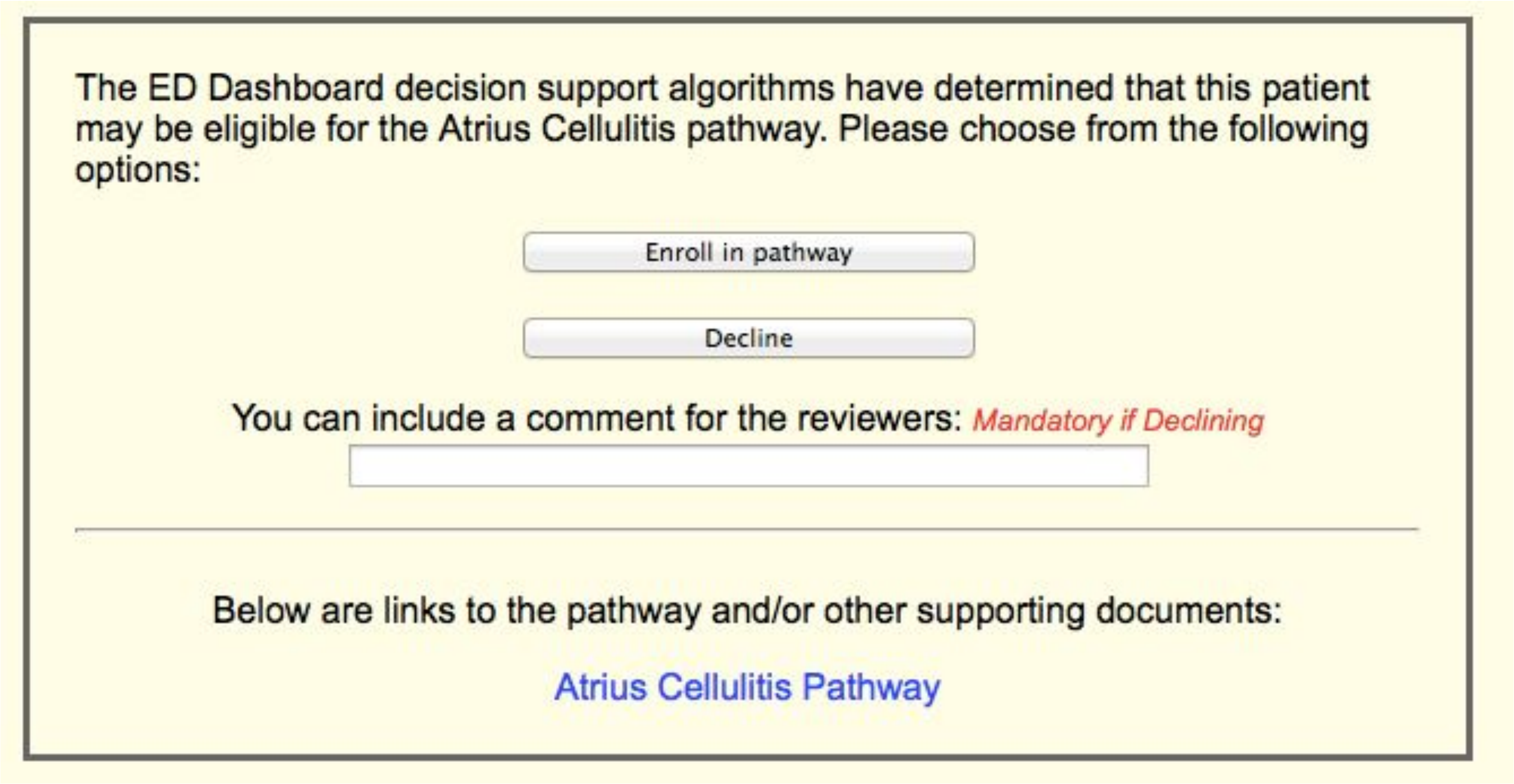
Behind-the-scenes reasoning about the patient's conditions (current and future)



- Better triage
- Faster diagnosis
- Early detection of adverse events
- Prevent medical errors

What will the ER of the future be like?

Propagating best practices



The ED Dashboard decision support algorithms have determined that this patient may be eligible for the Atrius Cellulitis pathway. Please choose from the following options:

Enroll in pathway

Decline

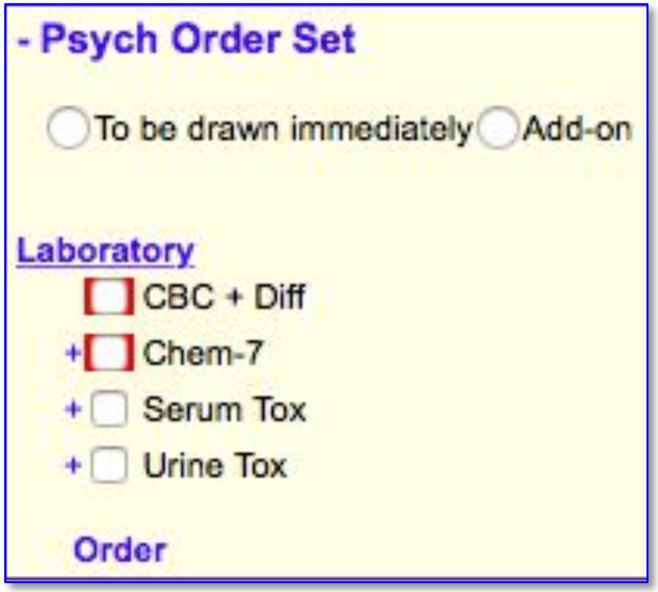
You can include a comment for the reviewers: *Mandatory if Declining*

Below are links to the pathway and/or other supporting documents:

Atrius Cellulitis Pathway

What will the ER of the future be like?

Anticipating the clinicians' needs


- Psych Order Set

To be drawn immediately Add-on

Laboratory

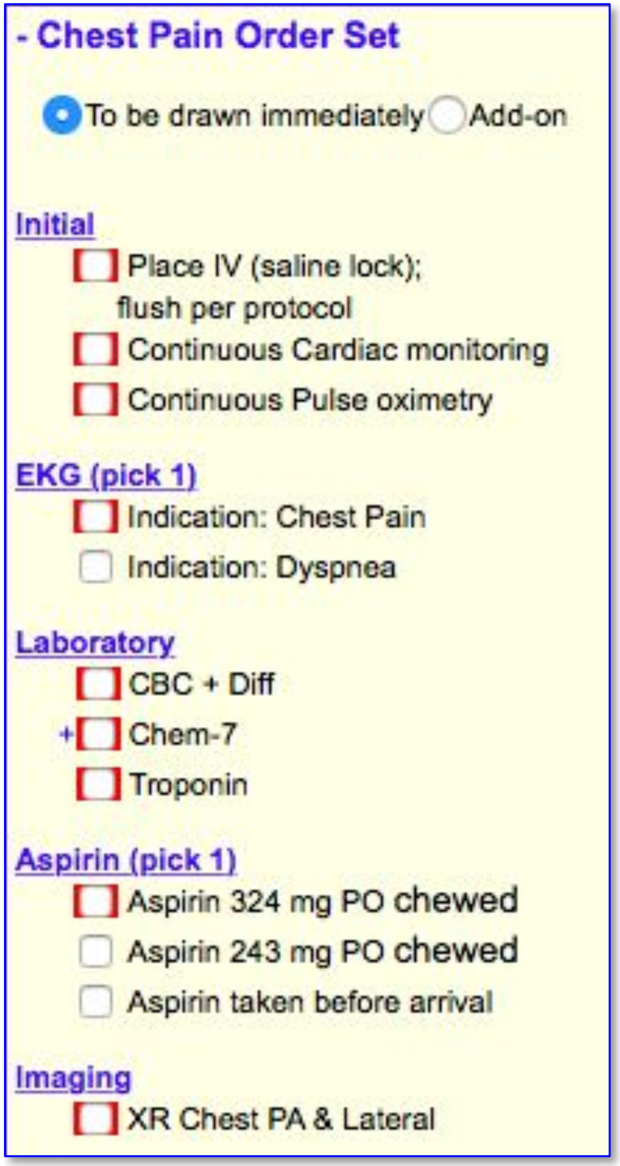
CBC + Diff

+ Chem-7

+ Serum Tox

+ Urine Tox

Order


- Chest Pain Order Set

To be drawn immediately Add-on

Initial

Place IV (saline lock); flush per protocol

Continuous Cardiac monitoring

Continuous Pulse oximetry

EKG (pick 1)

Indication: Chest Pain

Indication: Dyspnea

Laboratory

CBC + Diff

+ Chem-7

Troponin

Aspirin (pick 1)

Aspirin 324 mg PO chewed

Aspirin 243 mg PO chewed

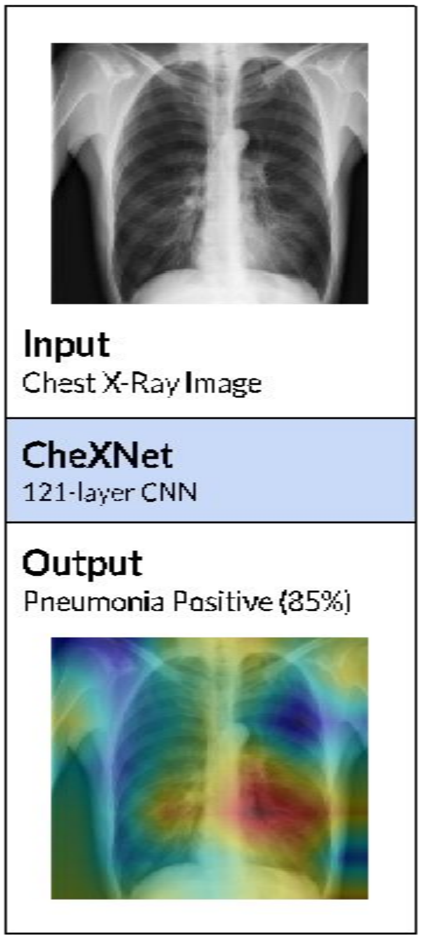
Aspirin taken before arrival

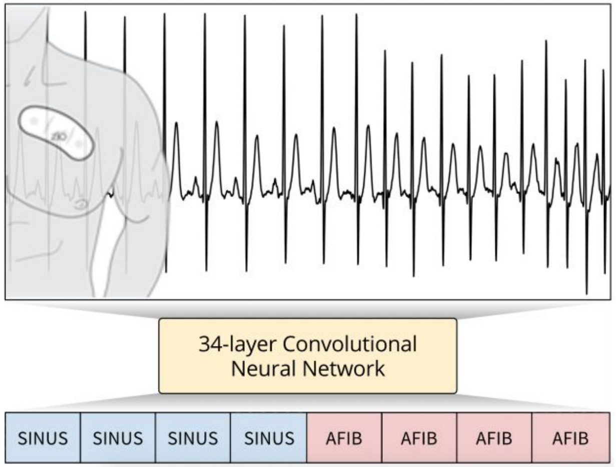
Imaging

XR Chest PA & Lateral

What will the ER of the future be like?

Reducing the need for specialist consults





Arrhythmia?

Figure sources: Rajpurkar et al., arXiv:1711.05225 ’17
Rajpurkar et al., arXiv:1707.01836, '17

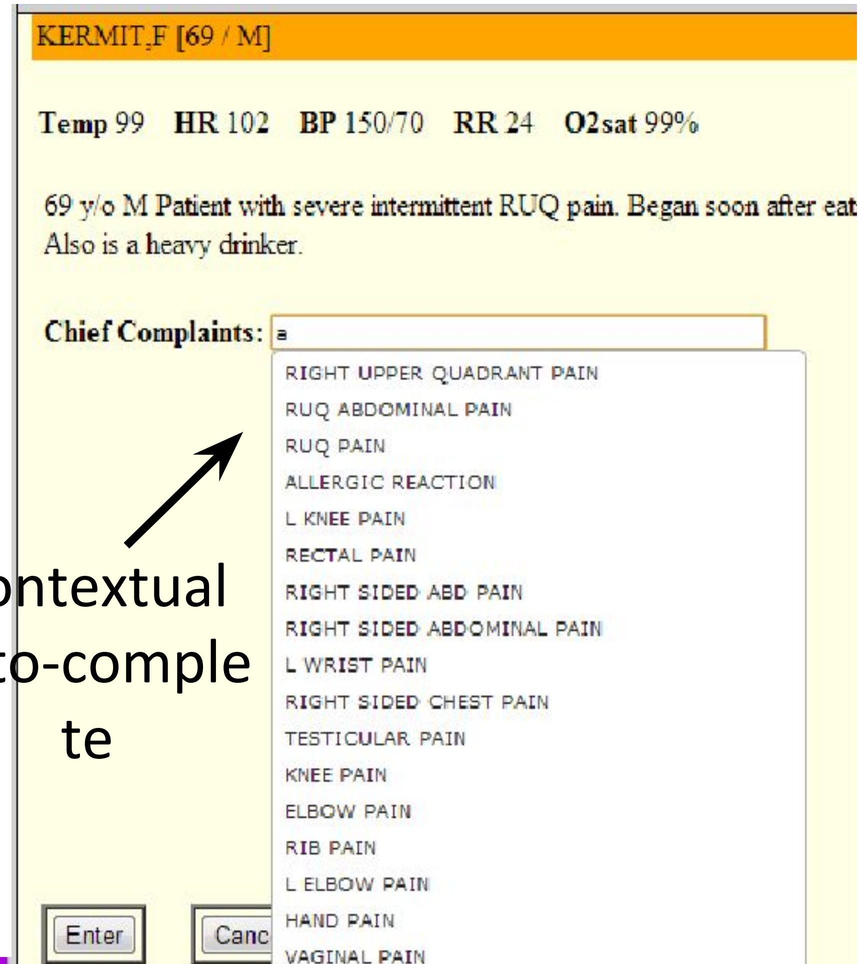
What will the ER of the future be like?

Automated documentation and billing

Triage note

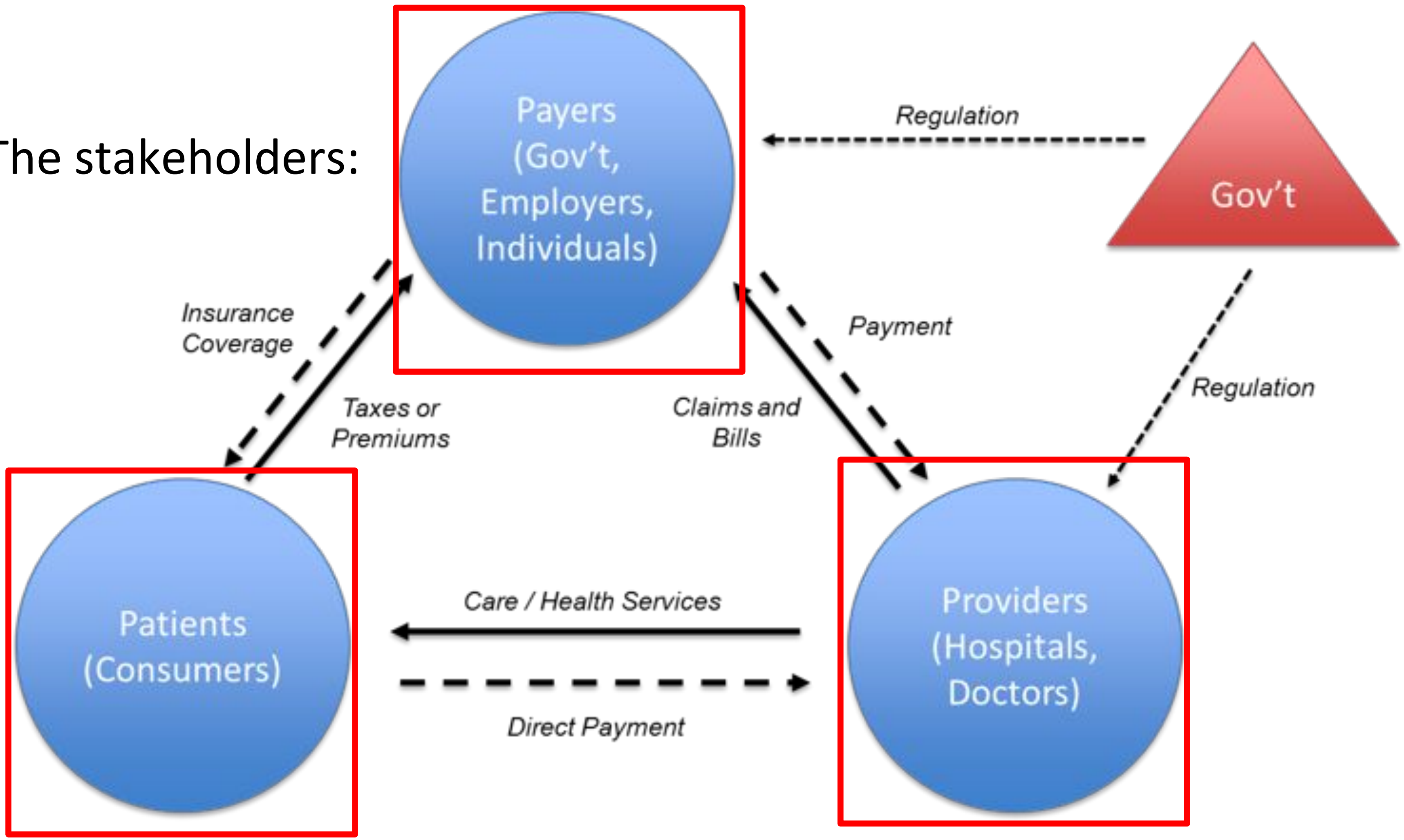
Predicted chief complaints

Contextual auto-complete



ML will transform every aspect of healthcare

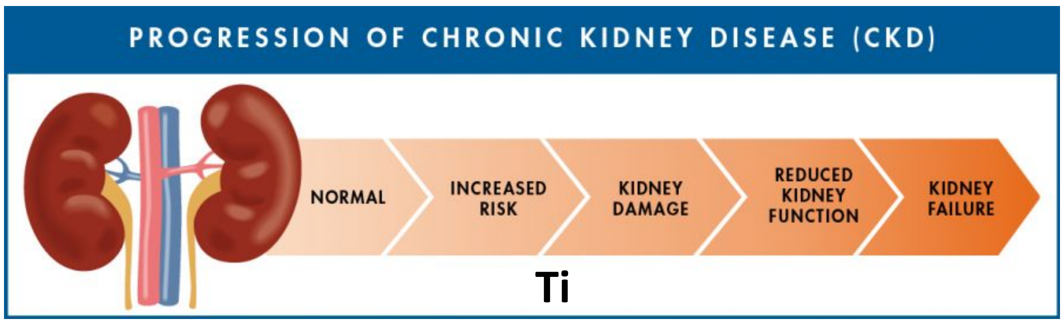
The stakeholders:



Source for figure:
http://www.mahesh-vc.com/blog/understanding-whos-paying-for-what-in-the-healthcare-industry

What is the future of how we treat chronic disease?

- Predicting a patient’s future disease progression



Ti
m
e

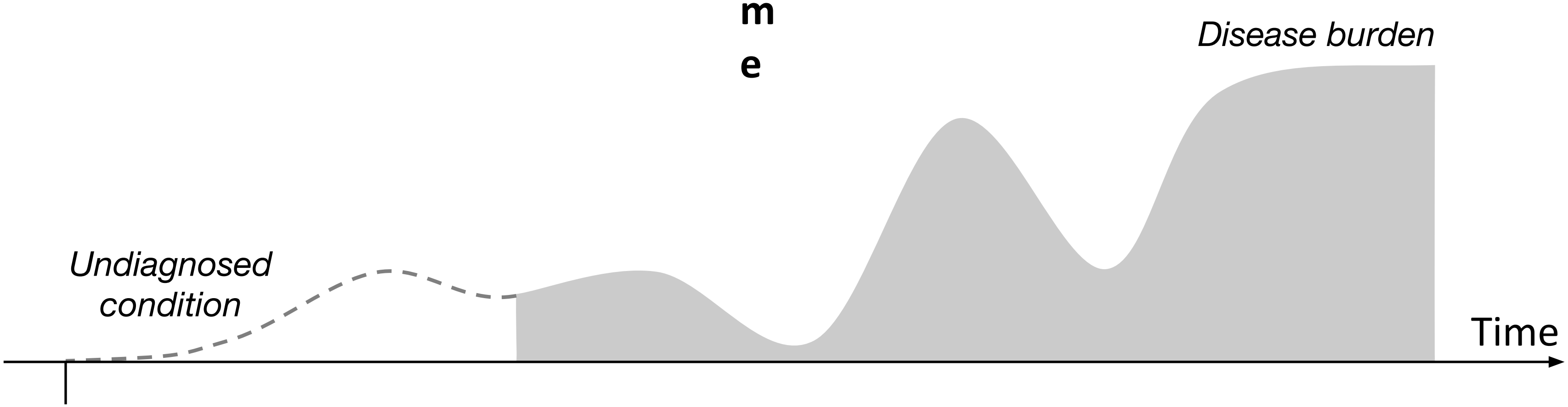


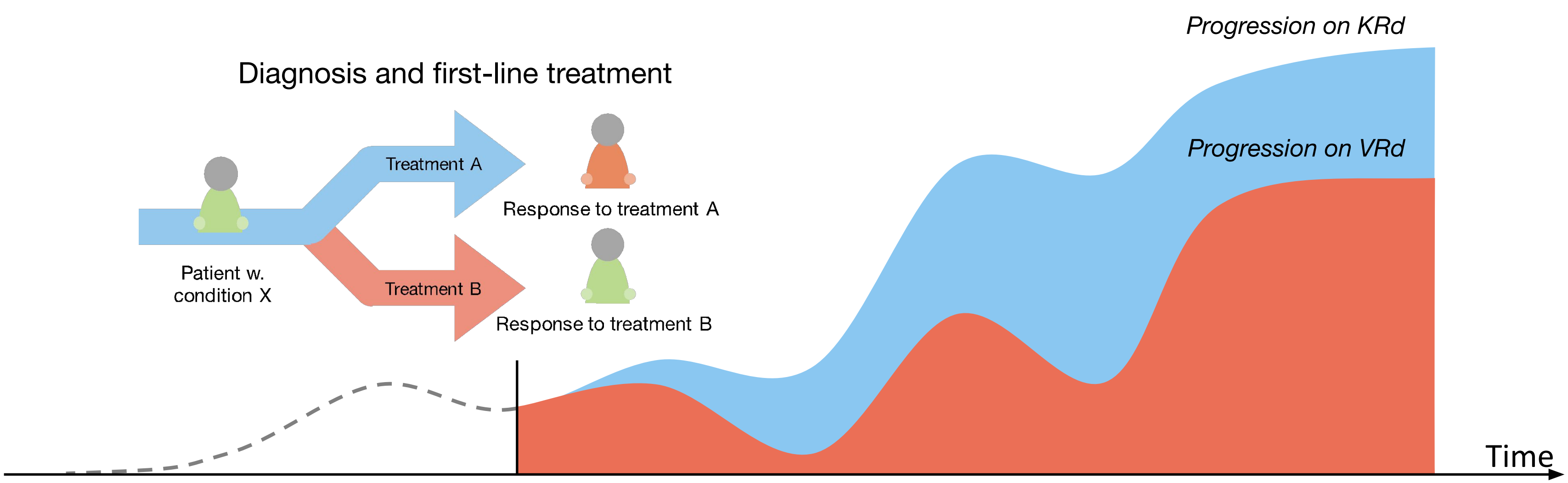
Figure credit: https://www.cdc.gov/kidneydisease/prevention-risk.html

What is the future of how we treat chronic disease?

- Predicting a patient's future disease progression
- Precision medicine

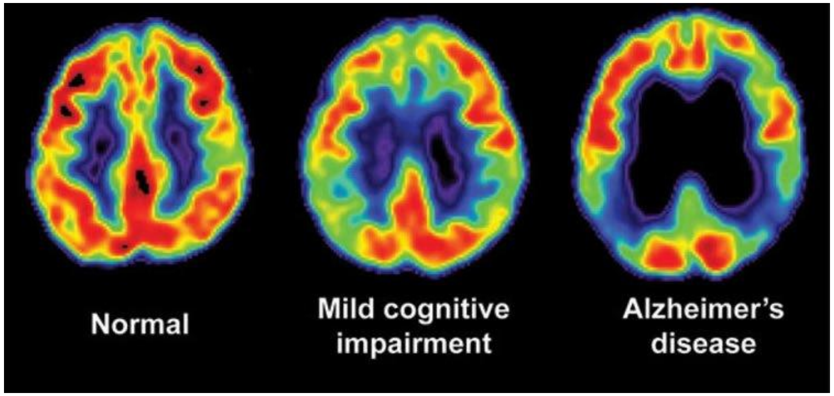
Choosing first line therapy in multiple myeloma

A) KRd: carfilzomib-lenalidomide-dexamethasone, **B) VRd:** bortezomib-lenalidomide-dexamethasone



What is the future of how we treat chronic disease?

- Early diagnosis, e.g. of diabetes, Alzheimer's, cancer



Liquid biopsy

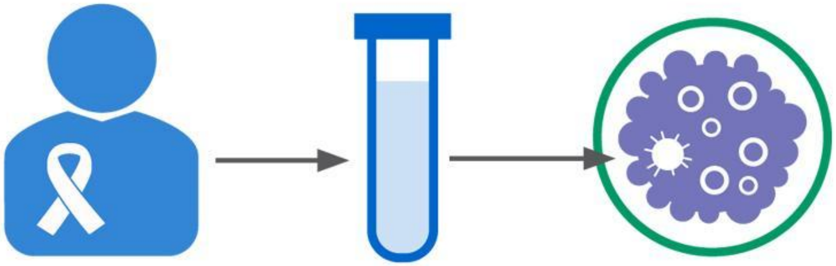


Figure sources: NIH, https://www.roche.com/research_and_development/what_we_are_working_on/oncology/liquid-biopsy.htm

What is the future of how we treat chronic disease?

- Continuous monitoring and coaching, e.g. for the elderly, diabetes, psychiatric disease

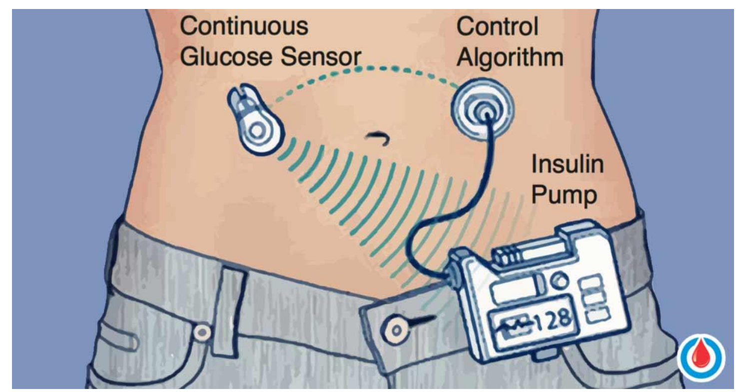
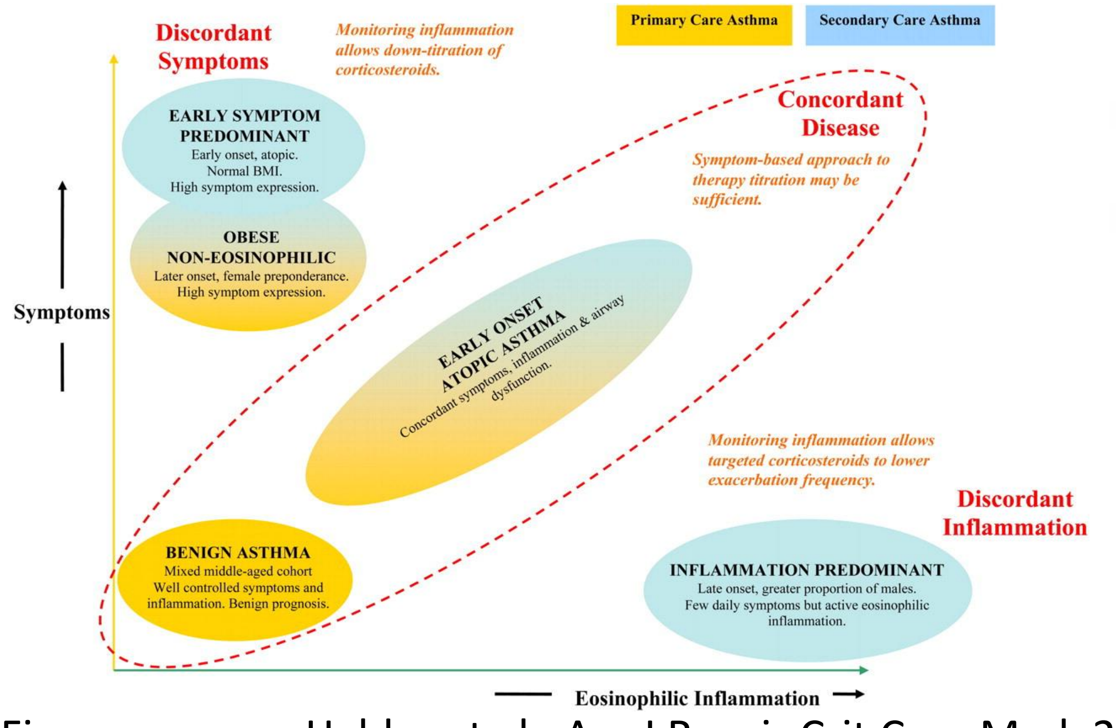


Figure source (left): http://www.emeraldforhome.com/

What is the future of how we treat chronic disease?

- Discovery of new disease subtypes; design of new drugs; better targeted clinical trials



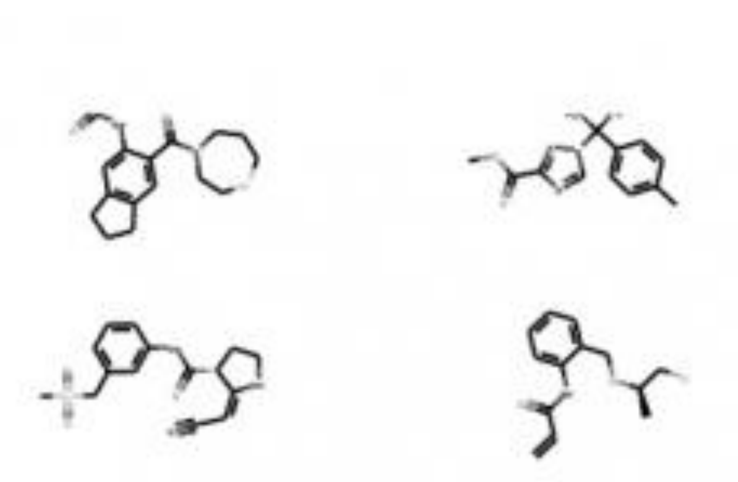


Figure sources: Haldar et al., Am J Respir Crit Care Med, 2008
http://news.mit.edu/2018/automating-molecule-design-speed-drug-development-0706

Outline for today's class

1. Brief history of AI and ML in healthcare
2. Why *now*?
3. Examples of how ML will transform healthcare
4. **What is *unique* about ML in healthcare?**

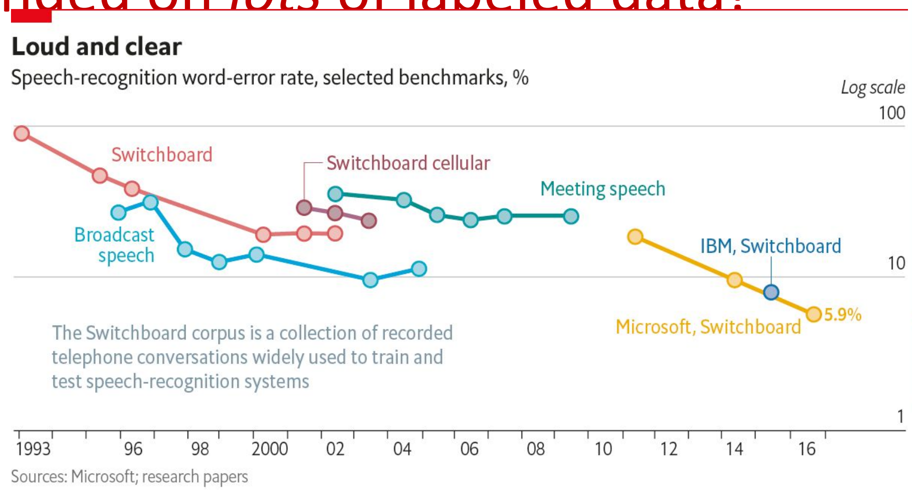
What makes healthcare different?

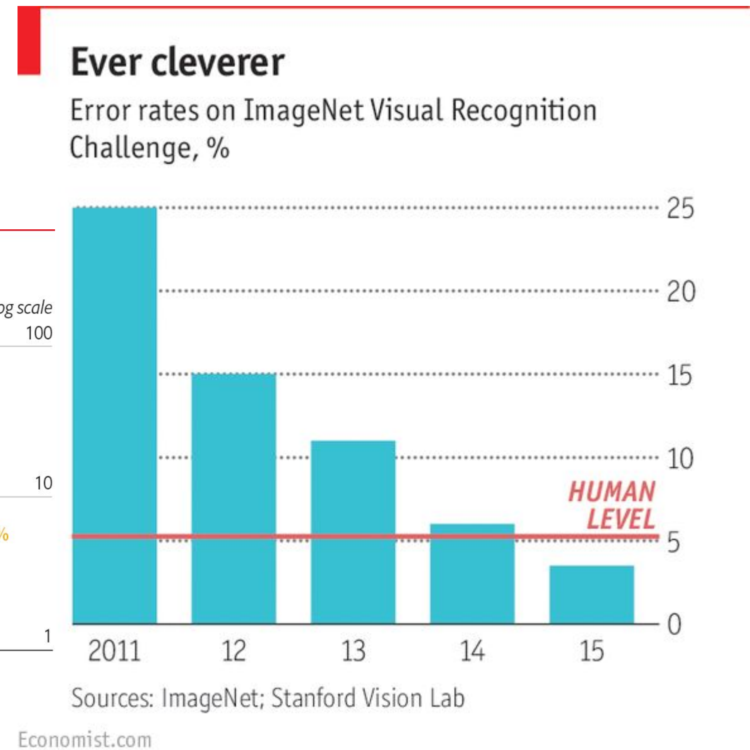
- Life or death decisions
  - Need **robust** algorithms
  - Checks and balances built into ML deployment
  - (Also arises in other applications of AI such as autonomous driving)
  - Need **fair** and **accountable** algorithms
- Many questions are about unsupervised learning
  - Discovering disease subtypes, or answering question such as “characterize the types of people that are highly likely to be readmitted to the hospital”?
- Many of the questions we want to answer are *causal*
  - Naïve use of supervised machine learning is insufficient

What makes healthcare different?

- Very little labeled data

Recent breakthroughs in AI depended on *lots* of labeled data!





What makes healthcare different?

- Very little labeled data
  - Motivates semi-supervised learning algorithms
- Sometimes small numbers of samples (e.g., a rare disease)
  - Learn as much as possible from other data (e.g. healthy patients)
  - Model the problem carefully
- Lots of missing data, varying time intervals, censored labels

What makes healthcare different?

- Difficulty of de-identifying data
  - Need for data sharing agreements and sensitivity
- Difficulty of deploying ML
  - Commercial electronic health record software is difficult to modify
  - Data is often in silos; everyone recognizes need for interoperability, but slow progress
  - Careful testing and iteration is needed