Lecture 2: Electronic Health Record Data

BIOMEDICAL DATA SCIENCE

Parisa Rashidi

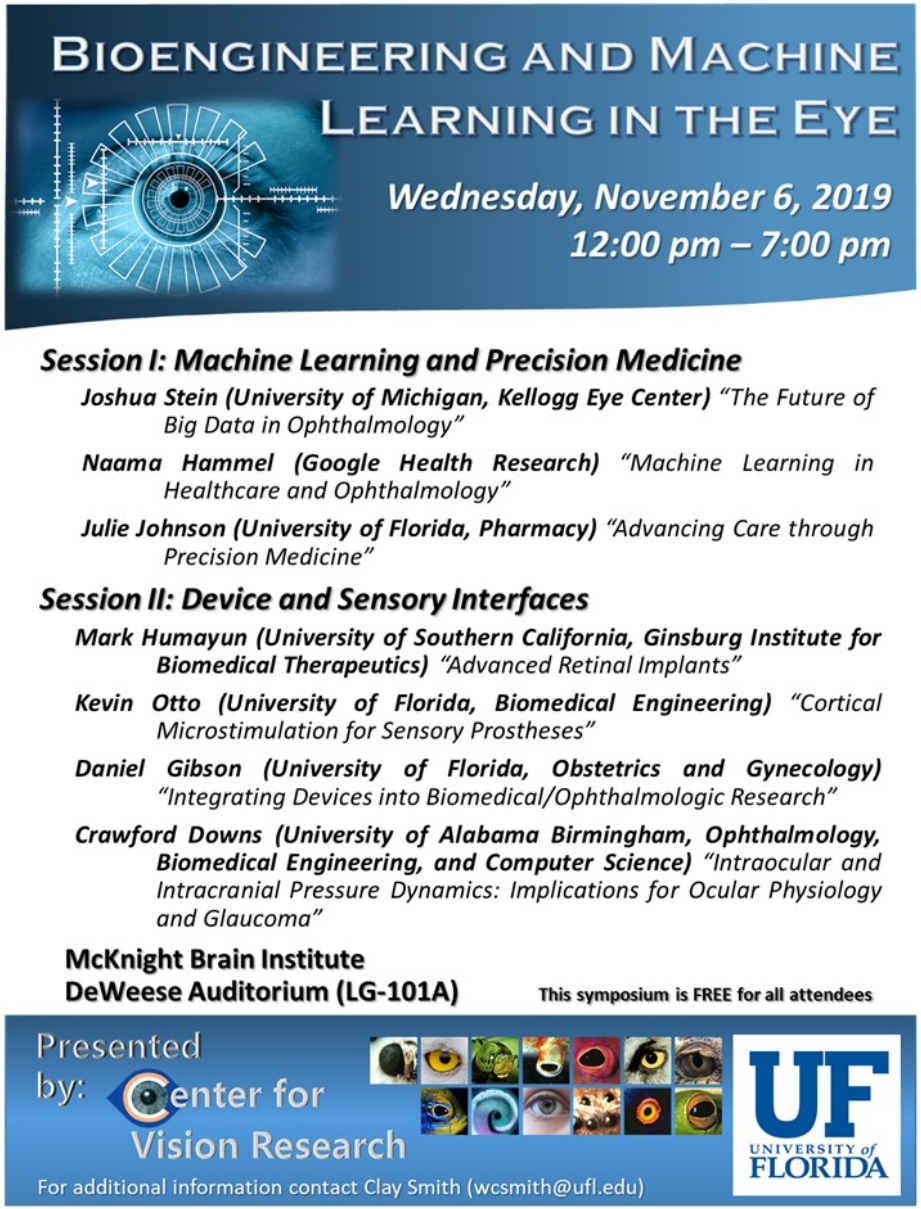
FALL 2019

Agenda

- Logistics
- Remaining Intro Slides
- EHR
- Python Programming

Seminars

- There are lots of interesting seminars hosted on campus, attend if you can.

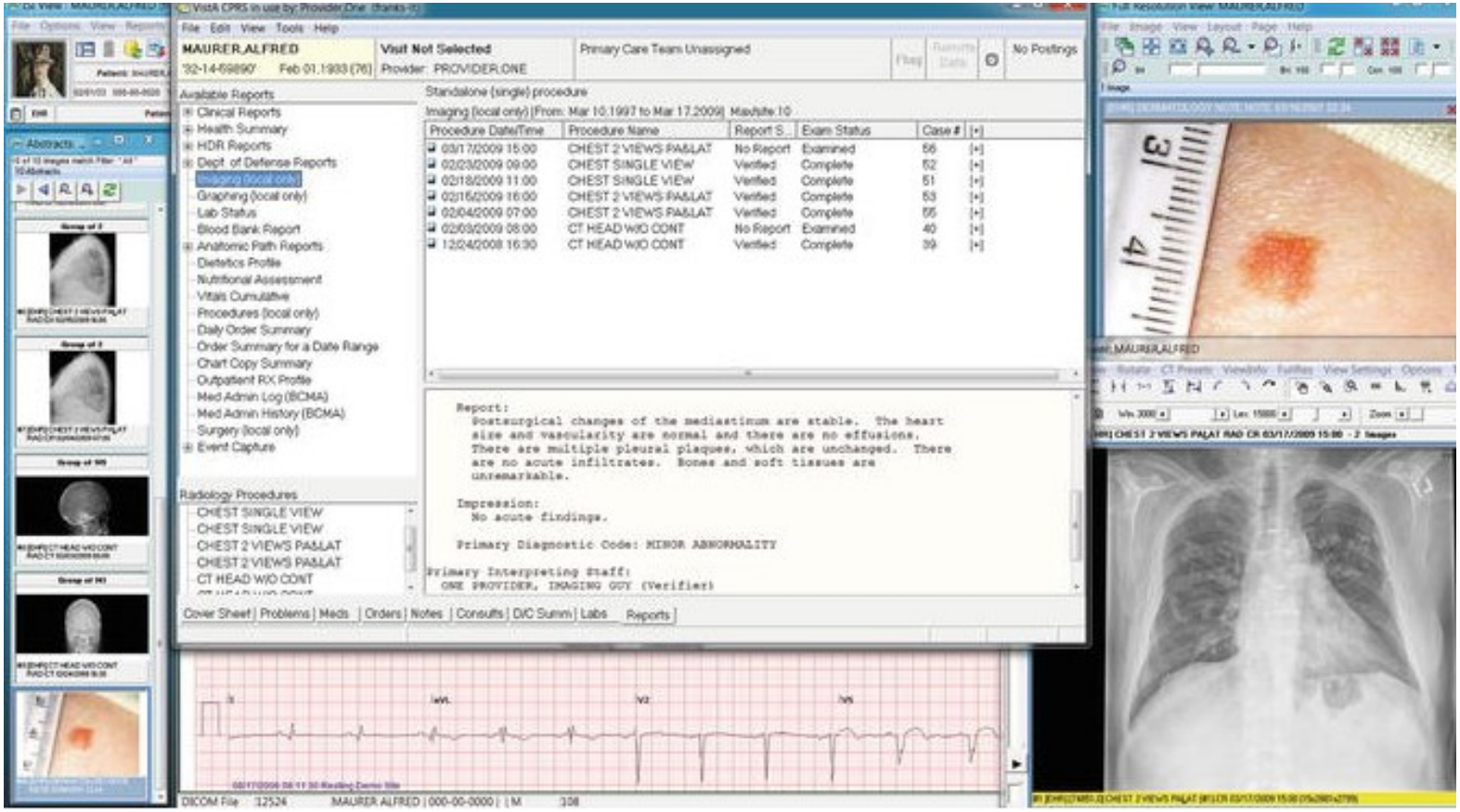


Moving Beyond the Paper Record

- Traditional paper-based medical record
- We are moving towards using "Electronic Health Records" (EHR)
  - Why?
    - Ease of information access
    - Supporting clinical trials
      - Learning about the safety and efficacy of new treatments
      - Gain insight into diseases
    - Ensure compliance

EHR

- Example patient record system:
  - WorldVistA Computer Based Patient Record System (CPRS) and ISI Imaging system (VA).



Clinical Data

- Patient (ID)
- Parameter being observed (e.g. liver size, urine sugar,…)
- Value of the parameter
- Time of observation
- Method observation (self-report, lab results, …)
- …

Patient No.	Last name	First name	Sex	Date of birth	Ward No.
454	Smith	John	M	14.08.58	6
223	Jones	Peter	M	07.12.65	8
597	Brown	Brenda	F	17.06.61	3
234	Jenkins	Alan	M	29.01.67	7
244	Wells	Christopher	M	25.02.55	6

Ward No.	Ward name	Type	No. of Beds
3	Carey	Medical	8
6	Bracken	Medical	16
7	Brent	Surgical	12
8	Meavy	Surgical	10

*http://www.technologyuk.net/computing/sad/relational_data_analysis.shtml

Question

The main reason for creating EHR systems:

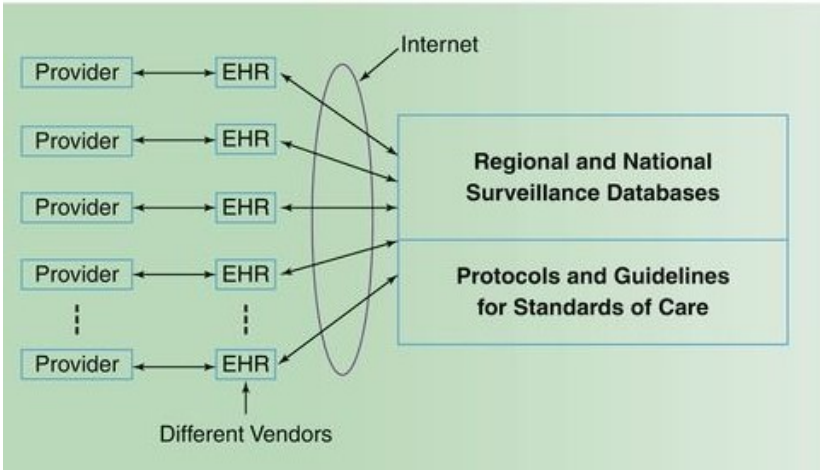
- ❑ Administrative tasks (insurance, billing, …)
- ❑ Research studies

Clinical Research

- Prospective studies
  - Clinical hypothesis is known in advance
  - Research protocol is designed to prevent bias
    - Randomized Controlled Trials (RCT)
- Retrospective studies
  - Hypothesis that was not identified at the time of data collection
  - Easy with EHR!

Health Data Integration

- Regional and national data integration
  - E.g. for disease surveillance



Question

We have a national EHR system, combining information of all regional/state EHR systems:

- ❑ True
- ❑ False

Question

We have state-wide EHR systems in all states:

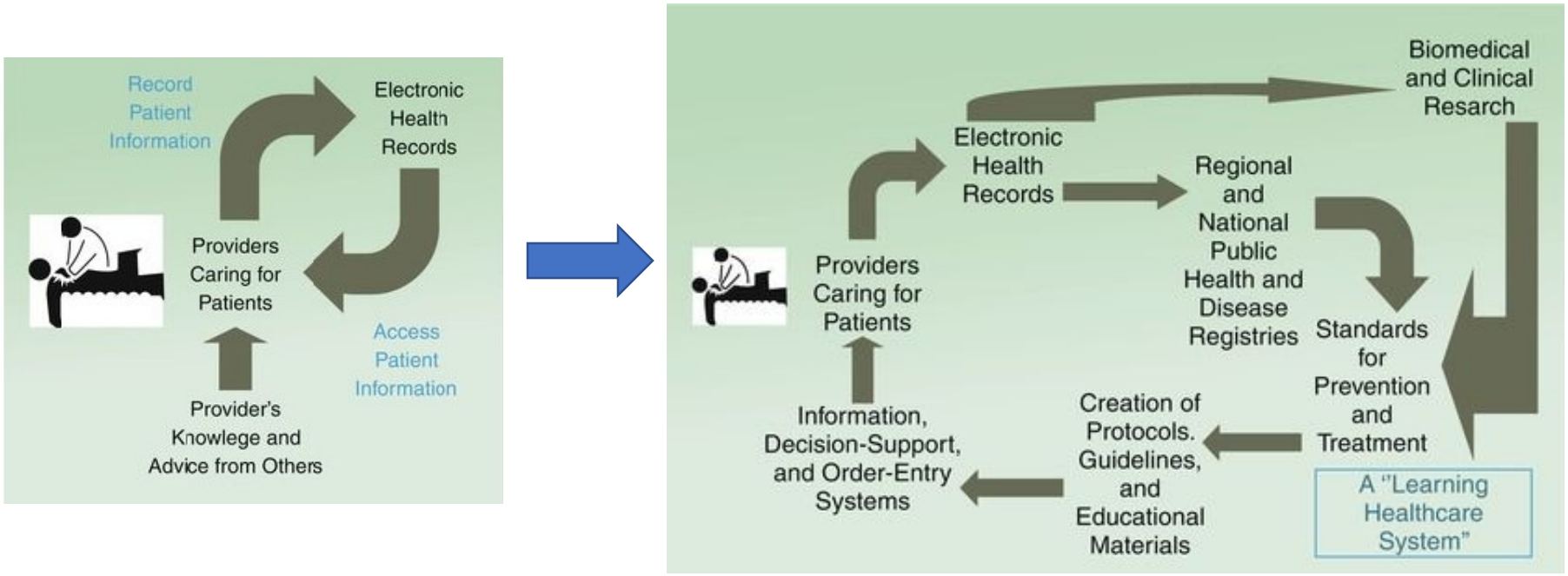
- ❑ True
- ❑ False

Regional & National EHR Integration

- Challenges
  - Mostly logistical, political, and financial rather than technical
    - HIPAA compliant policies
      - HIPAA= Health Insurance Portability and Accountability Act
    - Encryption of data
    - Standards for transmission and sharing
      - HL7= Health Level 7 (like a standard packet for transmitting data)
    - Standards for data definition
      - Content also needs to be standardized
    - Quality control and error checking
    - Role of state and federal government
      - Funding, creating, addressing political issues (individuals fearing social repercussion threating liberty, employment, …)

A Learning Health Care System

- A cycle of information flow in a distributed form



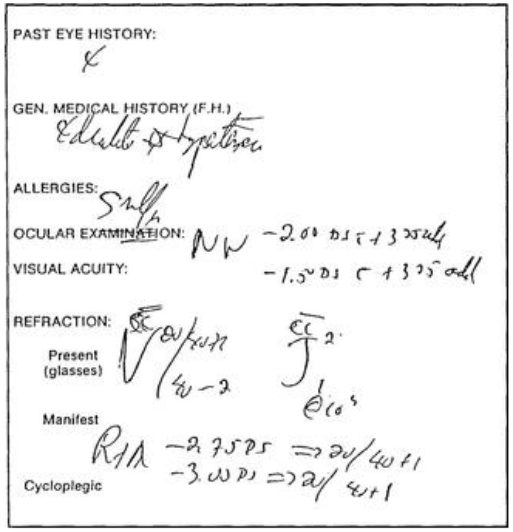
Complications with Clinical Data

- Frequency of data recording
  - Annual checkups versus continuous measurements of mean arterial blood pressure in cardiogenic shock
- Circumstances (context)
  - Was blood pressure taken in the leg or arm? Standing or sitting? What kind of device?
- Uncertainty
  - A radiologist looking at a shadow on a chest X-ray film is not sure whether it represents overlapping blood vessels or a lung tumor.
  - A confused patient is able to respond to simple questions about his or her illness, but under the circumstances the physician is uncertain how much of the history being reported is reliable.

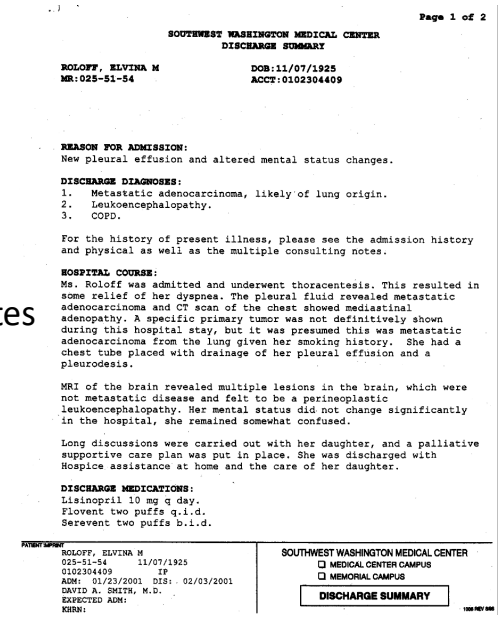
Narrative Clinical Data

- Lots of data is in narrative form
  - Patient description of illness, responses to physician's questions, physician's evaluations, pathologic examination, surgical procedures, …

An ophthalmologist's notes



Patient Discharge Notes



Page 1 of 2

SOUTHWEST WASHINGTON MEDICAL CENTER
DISCHARGE SUMMARY

ROLOFF, ELVINA M DOB:11/07/1925
MR:025-51-54 ACCT:0102304409

REASON FOR ADMISSION:
New pleural effusion and altered mental status changes.

DISCHARGE DIAGNOSES:
1. Metastatic adenocarcinoma, likely of lung origin.
2. Leukoencephalopathy.
3. COPD.

For the history of present illness, please see the admission history and physical as well as the multiple consulting notes.

HOSPITAL COURSE:
Ms. Roloff was admitted and underwent thoracentesis. This resulted in some relief of her dyspnea. The pleural fluid revealed metastatic adenocarcinoma and CT scan of the chest showed mediastinal adenopathy. A specific primary tumor was not definitively shown during this hospital stay, but it was presumed this was metastatic adenocarcinoma from the lung given her smoking history. She had a chest tube placed with drainage of her pleural effusion and a pleurodesis.

MRI of the brain revealed multiple lesions in the brain, which were not metastatic disease and felt to be a perineoplastic leukoencephalopathy. Her mental status did not change significantly in the hospital, she remained somewhat confused.

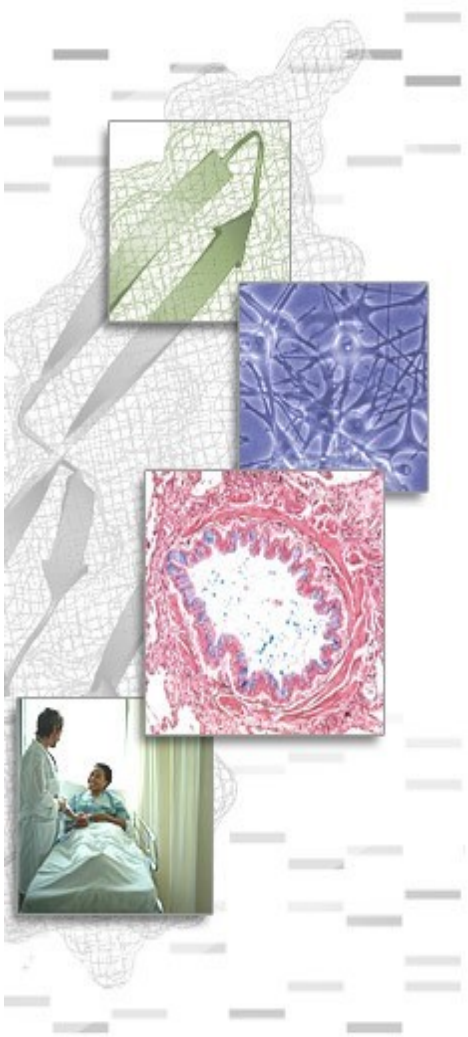
Long discussions were carried out with her daughter, and a palliative supportive care plan was put in place. She was discharged with Hospice assistance at home and the care of her daughter.

DISCHARGE MEDICATIONS:
Lisinopril 10 mg q day.
Flovent two puffs q.i.d.
Serevent two puffs b.i.d.

PATIENT IMPRINT
ROLOFF, ELVINA M
025-51-54 11/07/1925
0102304409 IP
ADM: 01/23/2001 DIS: 02/03/2001
DAVID A. SMITH, M.D.
EXPECTED ADM:
KHRN:

SOUTHWEST WASHINGTON MEDICAL CENTER
☐ MEDICAL CENTER CAMPUS
☐ MEMORIAL CAMPUS
DISCHARGE SUMMARY

Sample Clinical Documents



Pt recently hospitalized 7/19/06 for chf exacerbation (diastolic dysfunction) 2nd to dietary and medicine noncompliance (salty foods , stopped her HCTZ) and continued to smoke. Pt diuresed and sent home on new lasix 60qam 40qpm regimen. Pt noticed steady decline in functional status during the last 3 weeks because of SOB. at baseline should sat 85% on ra , 95% on 6L02NC at rest and ambulation. (on home o2) but now , can't ambulate , sating 83-89% on 6l at rest. also notes pnd , orthopnea. Pt notes intermif ent chest pain on and off lasting 5 minutes not associated with exertion or any other cardiac sx. 8/15 dobuta mibi-> ischemia in d1 territory. 11/19 :echo->ef 60% , Pa pressure 48 + RA. no valve dz. rv enlarged and hypokinetic. A/P: pump: decompesated CHF (diastolic dysfxn , cor pulmonale component) 2nd to diet/med non-compliance. uptitrate captopril , continue iv lasix 60 qd with goal net neg 2 liters , daily weights , strict Iand O. check cxray. Switched to po lasix 10/06 , back to lisinopril for d/c Fri. ischemia: has + mibi in past , but no further workup to d1 lesion. can't get ecasa 2nd to vWD. continue BB , will hold off on statin since not hyperlipidemic. rate:tele. …

Issues with Narrative Data

- Physician-specific
  - Each physician writes reflecting his/her thought process
    - Shortness of breath vs. dyspnea
- Domain-specific shorthand conventions
  - In eye examination: PERRLA (“Pupils are Equal in size, Round, and Reactive to Light and Accommodation)
- Context-specific
  - MI: Mitral Insufficiency (leakage in one of the heart’s valves)
  - MI: Myocardial infarction (the medical term for what is commonly called a heart attack).
- Sentences changed into phrases
  - Pain relieved by antiacid

Biomedical & Clinical Standards

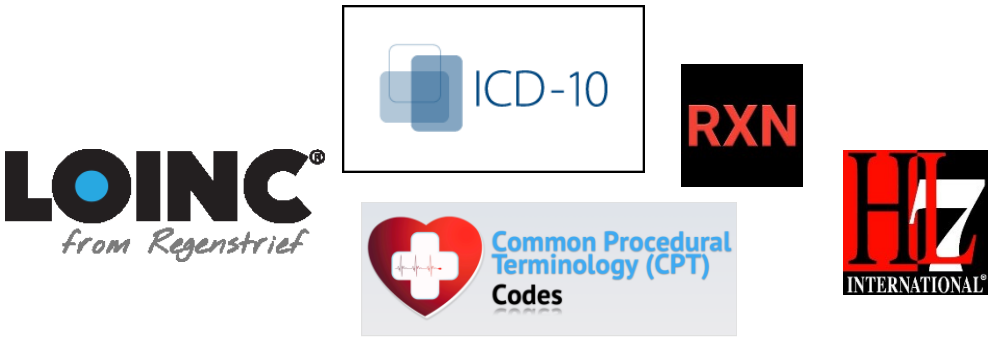
The Need for Standards

- Excessive diversity in communication among many units
  - Primary physician
  - Specialists
  - Reimbursement & Insurance
  - …
- Pooling data together for research

- Early interest in standards were driven by the need for exchanging data between clinical laboratories and clinical systems.

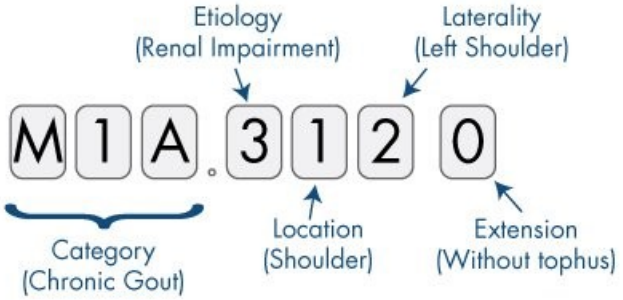
Some Important Standards

- Diagnosis: ICD-10
- Procedures: CPT
- Laboratory tests: LOINC
- Drugs: National Drug Codes (NDC), RxNorm
- Data Exchange: HL7
- Unified Medical Language System: UMLS



Examples of Standards

- There are many clinical standards



ICD-10 Example

Schema	Number of Codes	Examples
ICD-10 *(Diagnosis)*	68,000	- J9600: Acute respiratory failure - I509: Heart failure - I5020: Systolic heart failure
CPT *(Procedures)*	9,641	- 72146: MRI Thoracic Spine - 67810: Eyelid skin biopsy - 19301: Partial mastectomy
LOINC *(Laboratory)*	80,868	- 4024-6: Salicylate, Serum - 56478-1: Ethanol, Blood - 3414-0: Buprenorphine Screen
RxNorm *(Medications)*	116,075	- 161: Acetaminophen - 7052: Morphine - 1819: Buprenorphine

Shickel, Benjamin, Patrick James Tighe, Azra Bihorac, and Parisa Rashidi. "Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis." IEEE journal of biomedical and health informatics 22, no. 5 (2018): 1589-1604.

ICD-10

- Codes can be up to seven symbols.
- The first one is always a letter.
- The second is always numeric.
- The remaining five can be any combination.

Example ICD-10 Codes

A01 Typhoid and paratyphoid fevers
- A01.0 Typhoid Fever
  - A01.03 Typhoid Pneumonia *

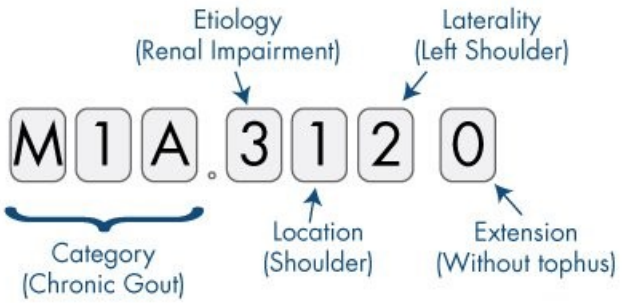
A02 Other salmonella infection
- A02.2 Localized salmonella infections
  - A02.22 Salmonella pneumonia *

A20 Plague
- A20.2 Pneumonic plague

A22 Anthrax
- A22.1 Pulmonary anthrax

A37 Whooping cough
- A37.0 Whooping cough due to *Bordetella pertussis*
  - A37.01 Whooping cough due to *Bordetella pertussis* with pneumonia *
- A37.1 Whooping cough due to *Bordetella parapertussis*
  - A37.11 Whooping cough due to *Bordetella parapertussis* with pneumonia *
- A37.8 Whooping cough due to other Bordetella species
  - A37.81 Whooping cough due to other Bordetella species with pneumonia *
- A37.9 Whooping cough, unspecified
  - A37.91 Whooping cough, unspecified species with pneumonia *

A50 Congenital syphilis
- A50.0 Early congenital syphilis, symptomatic
  - A50.04 Early congenital syphilitic pneumonia *



Diagnosis Codes: ICD

- World Health Organization (WHO) publishes diagnostic coding schema called the International Classification of Diseases (ICD)
- Developed for epidemiologic reporting
  - ICD-10 used in much of the world,
  - ICD-11 released in May 2018
- ICD-10 has roughly 68,000 available codes (with flexibility for adding new ones) in comparison to ICD-9's 13,000.

Question

We have a unique patient identifier across US EHR systems:

- ❑ True
- ❑ False

Patient Identifier

- Should include a check digit to ensure accuracy
- Mechanism for issuing such identifiers
  - National Provider Identifier by Center for Medicare and Medicaid Services (CMS)
    - 9 digits + check digit
  - Individual Identifier
    - Law passed in 1996, but pushed back by privacy advocates
    - US is among the few developed countries without such an identifier

Procedures: CPT

- Current Procedural Terminology (CPT)
- Widely used in producing bills and reimbursement
  - Specifies information that differentiates the codes based on the cost
    - E.g., there are different codes for pacemaker insertions, depending on whether the leads are “epicardial, by thoracotomy” (33200), “epicardial, by xiphoid approach” (33201) …
  - Also provides information about the reasons for a procedure

CPT: Example

CPT Code	CPT Code Descriptor	Medicare Physician Fee Schedule - National Average*		
		Global Payment	Professional Payment	Technical Payment
76536	Ultrasound of soft tissues of head and neck (e.g., thyroid, parathyroid, parotid), real time with image documentation	$125.54	$27.22	$98.33
76705	Ultrasound, abdominal, real time with image documentation); limited (e.g., single organ, quadrant, follow-up)	$111.26	$28.58	$82.68
76815	Ultrasound, pregnant uterus, real time with image documentation, limited (eg, fetal heart beat, placental location, fetal position and/or qualitative amniotic fluid volume), one or more fetuses	$92.20	$30.96	$61.24
76817	Ultrasound, pregnant uterus, real time with image documentation, transvaginal	$101.31‡	$36.74	$64.57‡
76818	Fetal biophysical profile; with non-stress testing	$125.89	$51.71	$74.17
76881	Ultrasound, extremity, nonvascular, real-time with image documentation; complete	$124.52	$30.96	$93.56

*http://www.idealmed.com/blog/billing-information-primary-care/

CPT Codes

TABLE 15–4 Examples of CPT Codes

CPT Code	Description
Level 1: 80000–89398	Pathology and laboratory tests
Level 2: 81000–81099	Urinalysis procedures
81015	Urinalysis; microscopic only
81005	Urinalysis; qualitative or semiquantitative, except immunoassays
Level 2: 80100–80103	Drug testing
80100	Drug screen, qualitative; multiple drug classes chromatographic method, each procedure

http://www.csun.edu/~dn58412/IS531/IS531_Lecture-5.ppt

Drugs: NDC & RxNorm

- Drug Codes
  - US: National Drug Codes (NDC): no uniform class hierarchy
  - US: RxNorm: to address the above problems, part of UMLS

Example National Drug Code (NDC)

Different Coding Systems

- Each has idiosyncrasies and limitations
  - ICD9-CM: more than 500 separate codes for Tuberculosis
- None can be completely satisfactory
  - Yet, if there is no structure to data and physicians treat EHR as a blank page, there is little use!
- Tension between the need for a system
  - General enough to cover many different patients
  - Precise and unique enough to cover a specific patient
- Is this variety a problem?

UMLS

- Researchers have worked for two decades to develop a unified language
  - Unified Medical Language System (UMLS)
    - Metathesaurus: contains over 8.9 million terms from 160 sources, has semantic relations tying concepts

http://www.nlm.nih.gov/research/umls/

UMLS

- Example

Bacterial pneumonia

Source:	CSP93/PT/2596-5280; DOR27/DT/U000523; ICD91/PT/482.9; ICD91/IT/482.9
Parent:	Bacterial Infections; Pneumonia; Influenza with Pneumonia
Child:	Pneumonia, Mycoplasma
Narrower:	Pneumonia, Lobar; Pneumonia, Rickettsial; Pneumonia, Staphylococcal; Pneumonia due to *Klebsiella pneumoniae;* Pneumonia due to Pseudomonas; Pneumonia due to *Hemophilus influenzae*
Other:	*Klebsiella pneumoniae, Streptococcus pneumoniae*

Pneumonia, Lobar

Source:	ICD91/IT/481; MSH94/PM/D011018; MSH94/MH/D011018; SNM2/RT/M-40000; ICD91/PT/481; SNM2/PT/D-0164; DXP92/PT/U000473; MSH94/EP/D011018; INS94/MH/D011018;INS94/SY/D011018
Synonym:	Pneumonia, diplococcal
Parent:	Bacterial Infections; Influenza with Pneumonia
Broader:	Bacterial Pneumonia; Inflammation
Other:	*Streptococcus pneumoniae*
Semantic:	inverse-is-a: *Pneumonia*
	has-result: *Pneumococcal Infections*

Interchange Standards: HL7

- For transfer of clinical and administrative data between hospital information systems
- Version 2: XML capabilities makes it Web-enabled(latest version 3.0)

Interchange Standards: HL7

- Example

```
MSH|^~&\|DHIS|OR|TMR|SICU|199212071425|password|ADT|16603529|P|2.1<cr>

EVN|A02|199212071425||<cr>

PID|||Z99999^5^M11||GUNCH^MODINE^SUE|RILEY|19430704 |F||C|RT. 1, BOX
97^ZIRCONIA^NC^27401 |HEND|(704)982-1234|(704)983-1822||S|C||245-33-
9999<cr>

PV1|1|I|N22^2204|||OR^03|0940^DOCTOR^HOSPITAL^A||| SUR|||||A3<cr>

OBR|7|||93000^EKG REPORT|R|199401111000|199401111330|||RMT||||19940111
11330|?|P030|||||199401120930|||||88-126666|A111|VIRANYI^ANDREW<cr>

OBX|1|ST|93000.1^VENTRICULAR RATE(EKG)||91|/MIN|60-100<cr>

OBX|2|ST|93000.2^ATRIAL RATE(EKG)||150|/MIN|60-100<cr>

. . .

OBX|8|ST|93000&IMP^EKG DIAGNOSIS|1|^ATRIAL FIBRILATION<cr>
```

An example of an HL7 ADT transaction message. This message includes the Message Heading segment, the EVN trigger definition segment, the PID patient-identification segment, the PV1 patient-visit segment, the OBR general-order segment, and several OBX results segments.

Interchange Standards: IEEE 1073

- Standard for medical device communication
  - Bedside devices in intensive care unit, operating room, and emergency room