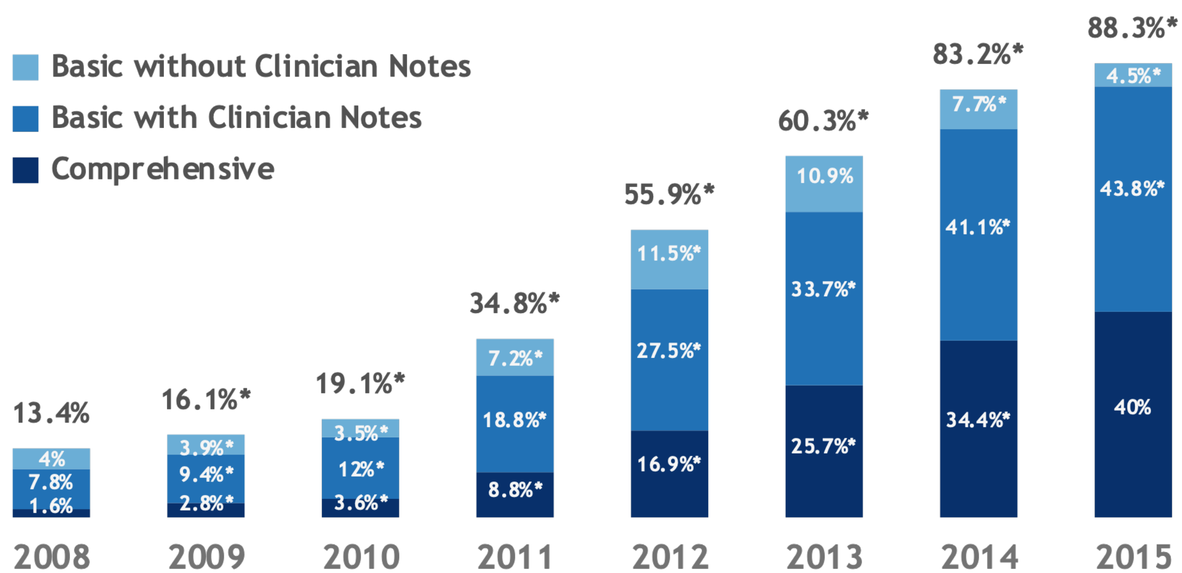
Topics in Machine Learning: Machine Learning for Healthcare

Monica Agrawal
PhD Student
MIT CSAIL

Shared slide credit with
Luke Murray, Divya Gopinath

Explosion of Electronic Health Records (EHRs)



Via https://dashboard.healthit.gov/evaluations/images/db-35-figure5.svg

Potential of EHRs

This rapid adoption has the potential to:

- Improve **clinical decision support** at the point-of-care
- Conduct **retrospective research** at an enormous scale
- **Empower patients** with their own data

Potential of EHRs

For example, data in EHRs could help answer:

- What clinical trials is my patient eligible for?
  (Clinical trial matching)
- What are likely diagnoses given my symptoms?
  (Differential diagnosis)

Potential of EHRs

For example, real-world evidence could help answer:

- What treatment would lead to the **best outcome** for **this patient**?
  (Heterogeneous treatment effect estimation, reinforcement learning)

- What is the patient's expected **disease trajectory**?
  (Disease progression modeling)

Potential Variables of Interest

Disease State

- Diagnoses
- Stage of Disease
- Symptoms

Treatment + Response

- Start/end of treatment
- Side effects

Confounders/Cohort Variables

- Pre-existing conditions
- Prior treatment (e.g. other hospital)

The caveat?

Many of these variables needed to tackle such use cases **responsibly** are not available in structured data, but trapped in **narrative, free-text** clinical notes

Variables of Interest

Disease State

- Progression/remission/stable
- Size of tumor mass(es)
- Sites of metastases

Treatment + Response

- Start/end of treatment
- Toxicity (symptoms)

Confounders/Cohort Variables

- Pre-existing conditions
- Prior treatment (e.g. other hospital)

So, how difficult can it be to interpret clinical text?

Pretty difficult!

“Pt given carbo ia for her TNBC. Will dc.”

Extraction involves two steps:

“Pt given carbo ia for her TNBC. Will dc.”

Extraction involves two steps:

1 **Entity Recognition:**
identification of
clinical concept spans

" Pt given carbo ia for her TNBC. Will dc. "

Extraction involves two steps:

1 Entity Recognition:
identification of
clinical concept spans

2 Entity Normalization:
mapping of spans
to clinical vocabularies (UMLS)

Patient? Prothrombin time? Carbodome? Intra-arterial? discontinue?

" Pt given carbo ia for her TNBC. Will dc. "

Physical therapist? Carboplatin? Intra-articular? D/C current? discharge? Doctor of Chiropractic?

Extraction involves two steps:

1 **Entity Recognition:** identification of clinical concept spans

2 **Entity Normalization:** mapping of spans to clinical vocabularies (UMLS)

“ Pt given carbo ia for her TNBC. Will dc. ”

Pt	given	carbo	ia	for	her	TNBC.	Will	dc.
Patient (C0030705)		Carboplatin (C0079083)	Intra-arterial (C1561451)			Triple-neg. breast cancer (C3539878)		Discontinue (C1706472)

This messiness in EHR data affects:

1. Patients, who cannot understand medical jargon.
2. Physicians, who have trouble retrospectively disambiguating between overloaded terms.
3. Learned algorithms that rely on structured data.

So, why are notes so messy?

Review > Acad Emerg Med. 2004 Nov;11(11):1127-34. doi: 10.1197/j.aem.2004.08.004.

Where's the beef? The promise and the reality of clinical documentation

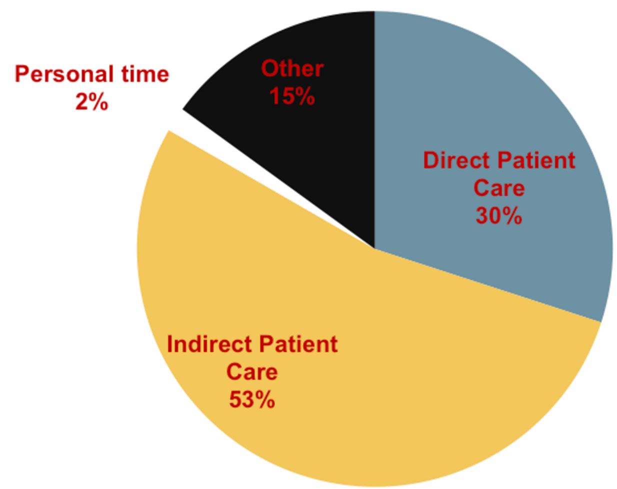
Steven J Davidson [1], Frank L Zwemer Jr, Larry A Nathanson, Kenneth N Sable, Abu N G A Khan

1. Recording of medical care and communication among providers
2. Payment for hospital and physician
3. Legal defense from medical negligence allegations
4. Symptom/disease surveillance, public health, and research functions

EHRs are not doctor-friendly

EHRs are not doctor-friendly

Emergency Physician Time



Chrisholm et. al. A Task Analysis of Emergency Physician Activities in Academic and Community Settings. Ann of Emerg Med. 2011.

Can we just create more labels?

Note Annotation is Difficult

Manually, note annotation is expensive and difficult to scale

- Requires domain expertise

- Not a natural byproduct of clinical practice

Automated methods face additional hurdles; there is often limited generalization due to dataset shift between settings and institutions and over time.

Note Annotation is Difficult

Manually, note annotation is expensive and difficult to scale:

- Requires domain expertise
- Difficult to share between institutions
- Not a natural byproduct of clinical practice
- Note bloat

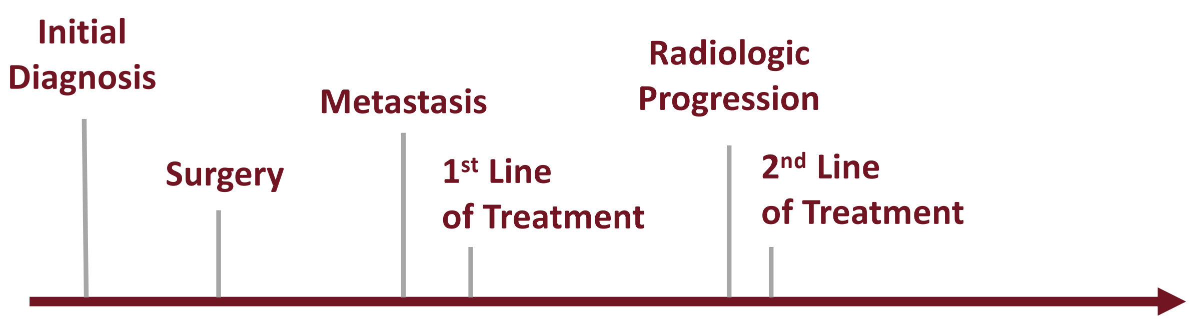
Note Annotation is Difficult

Automated methods face additional hurdles; there is often limited generalization due to dataset shift between settings and institutions and over time.

Solution? Approaches that are less data-hungry

Example: Creation of a Timeline

Such retrospective studies often first require **constructing a timeline** of events, with many events only found in the unstructured text in EHRs



Hybrid human-ML teams

- Manual extraction is expensive, especially when sifting through a long history

Zhao and Agrawal et al., MLHC 2021

Hybrid human-ML teams

- Manual extraction is expensive, especially when sifting through a long history
- Automated extraction is error-prone, especially for complicated patient cases

Zhao and Agrawal et al., MLHC 2021

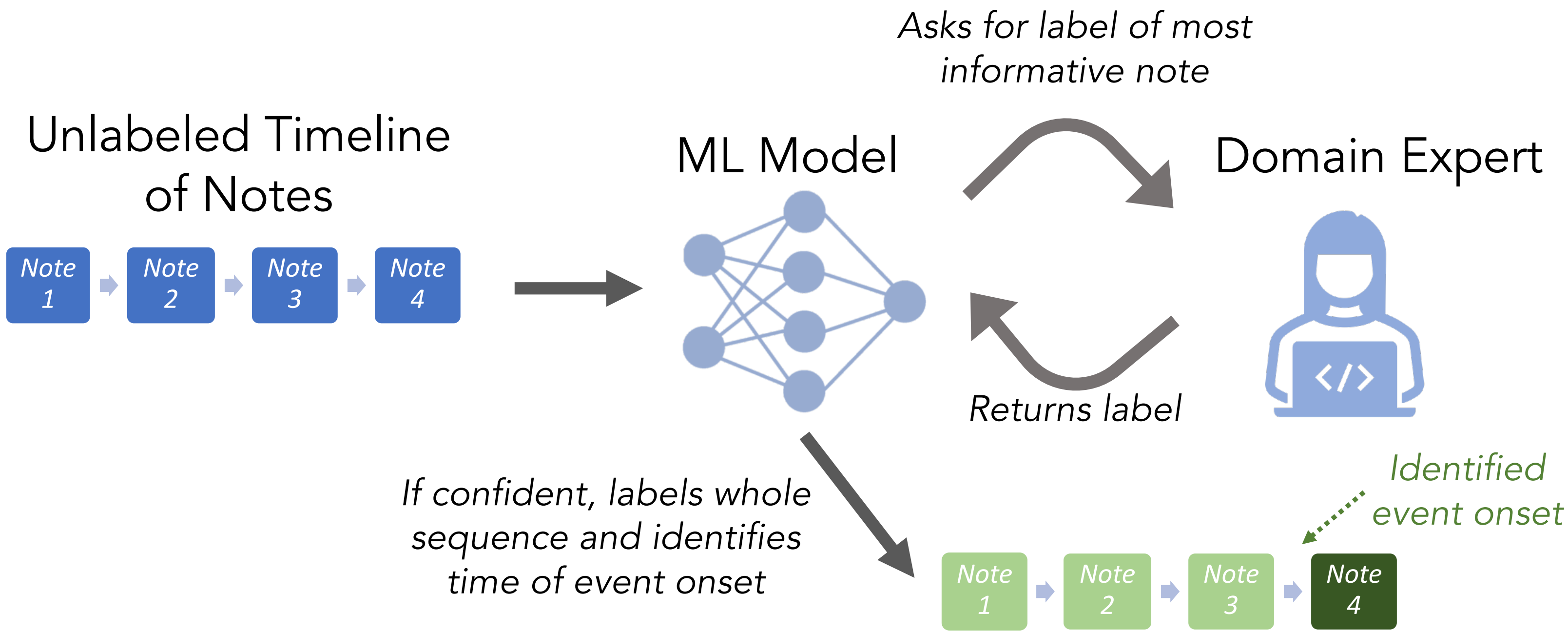
Hybrid human-ML teams

- Manual extraction is expensive, especially when sifting through a long history
- Automated extraction is error-prone, especially for complicated patient cases

Solution? Combine both! Use ML on simple cases, and defer to domain experts when needed

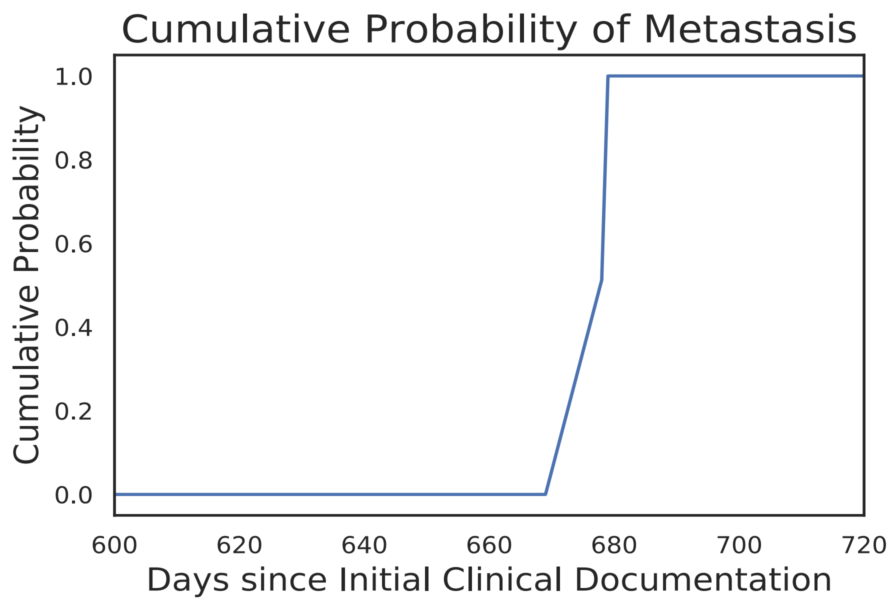
Zhao and Agrawal et al., MLHC 2021

Human-in-the-loop Framework

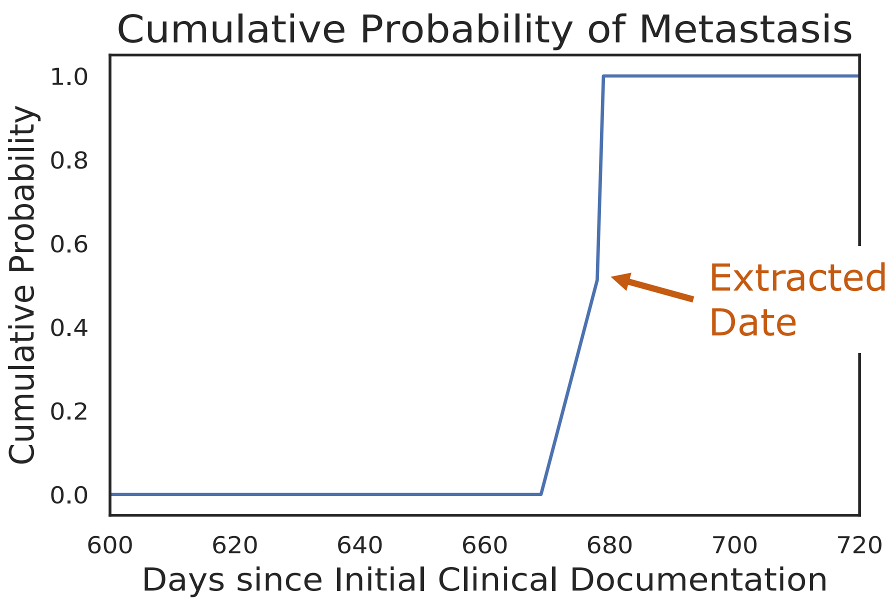


Zhao and Agrawal et al., MLHC 2021

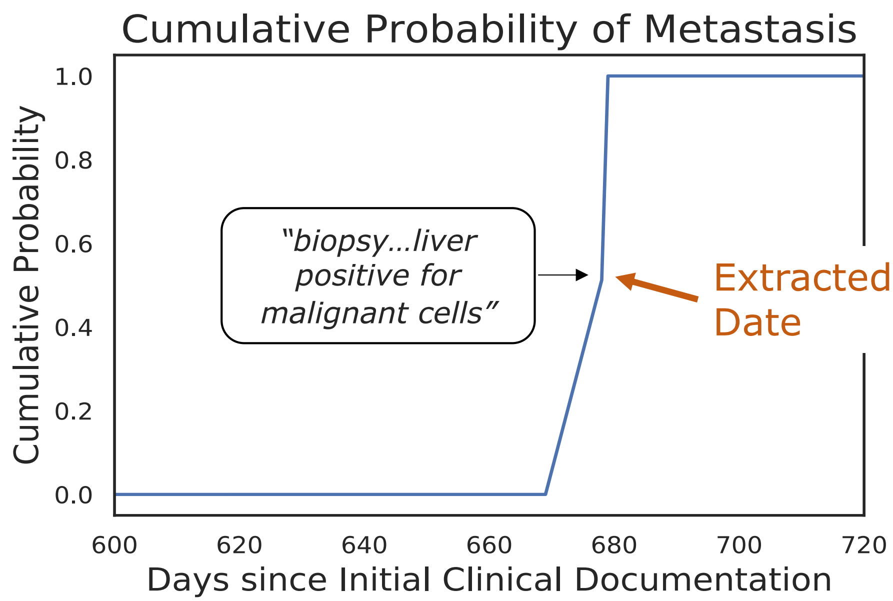
Metastasis Example #1: *No queries required*



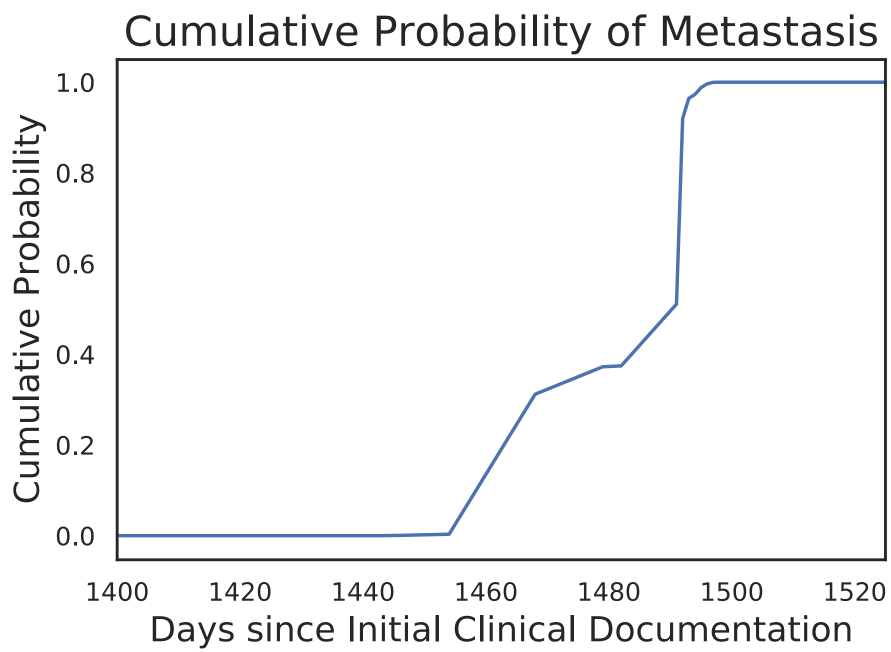
Metastasis Example #1: *No queries required*



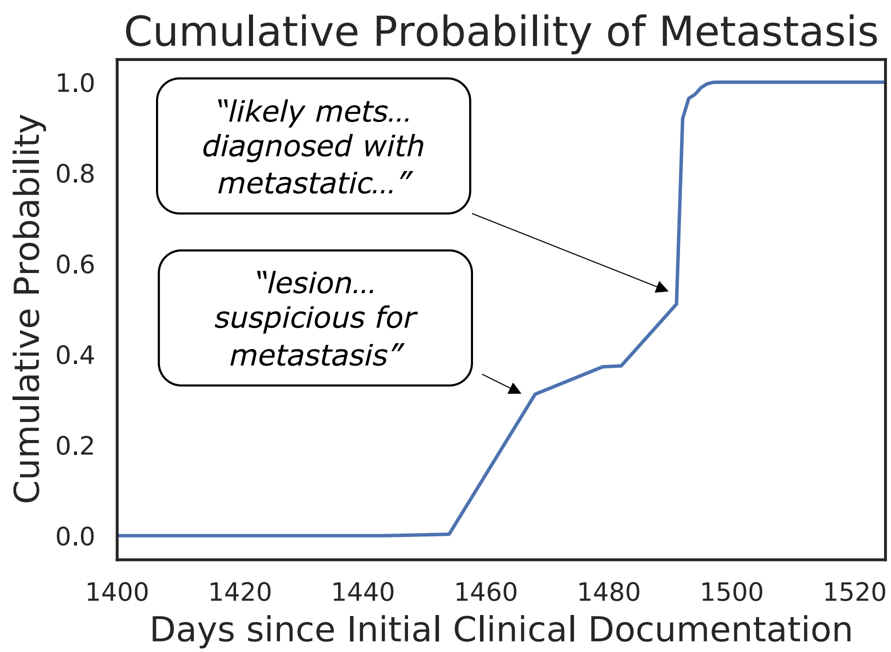
Metastasis Example #1: *No queries required*



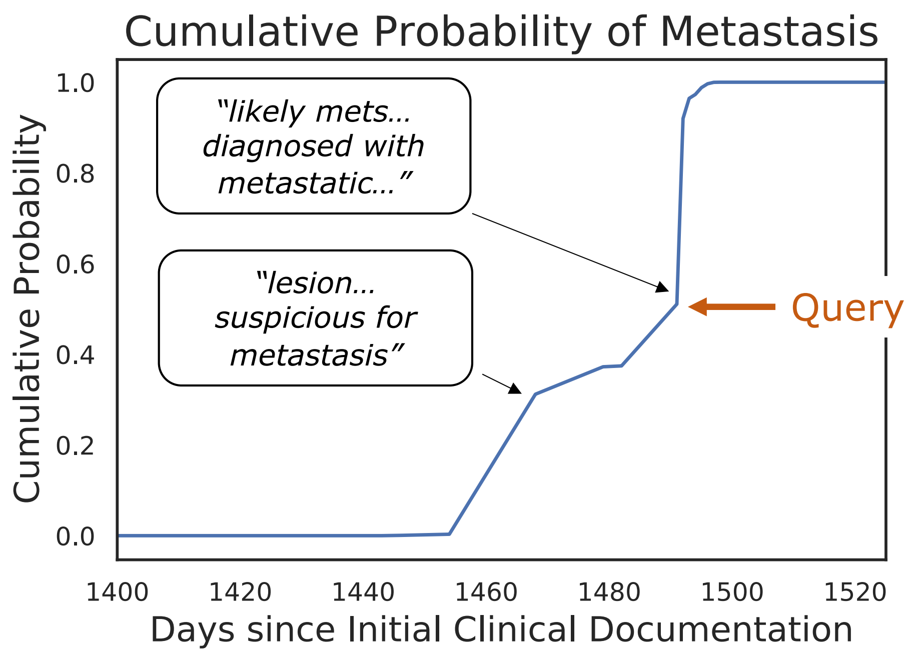
Metastasis Example #2: *Single query required*



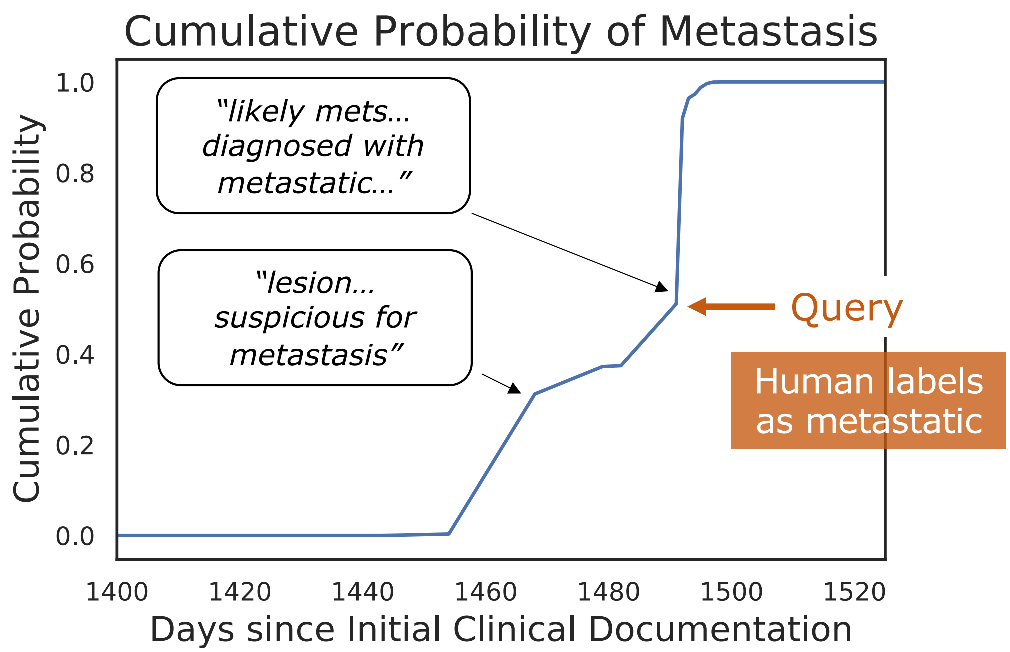
Metastasis Example #2: *Single query required*



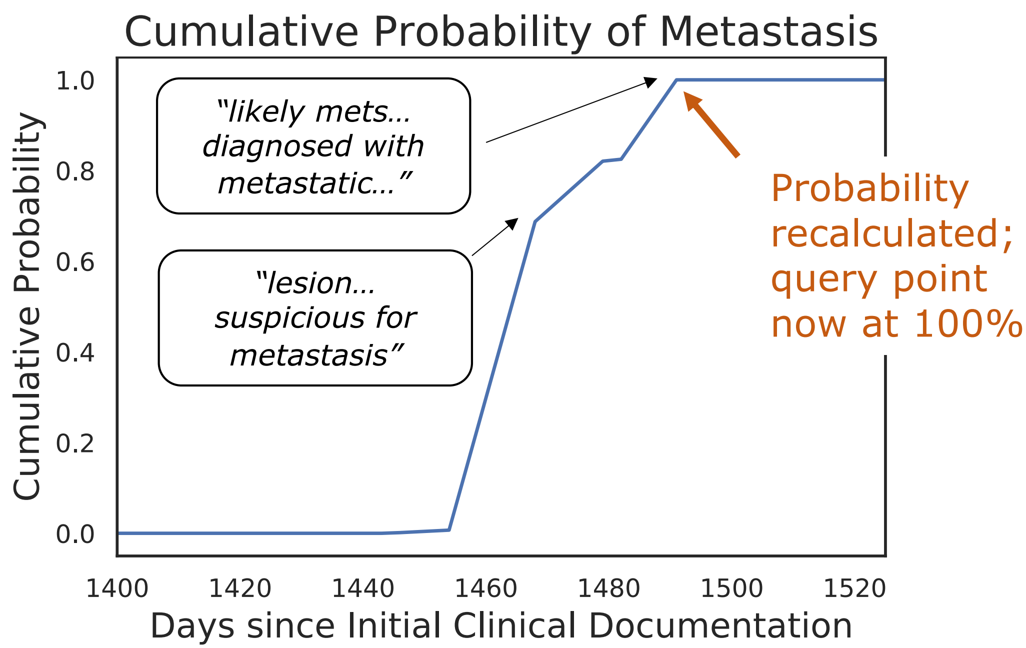
Metastasis Example #2: *Single query required*



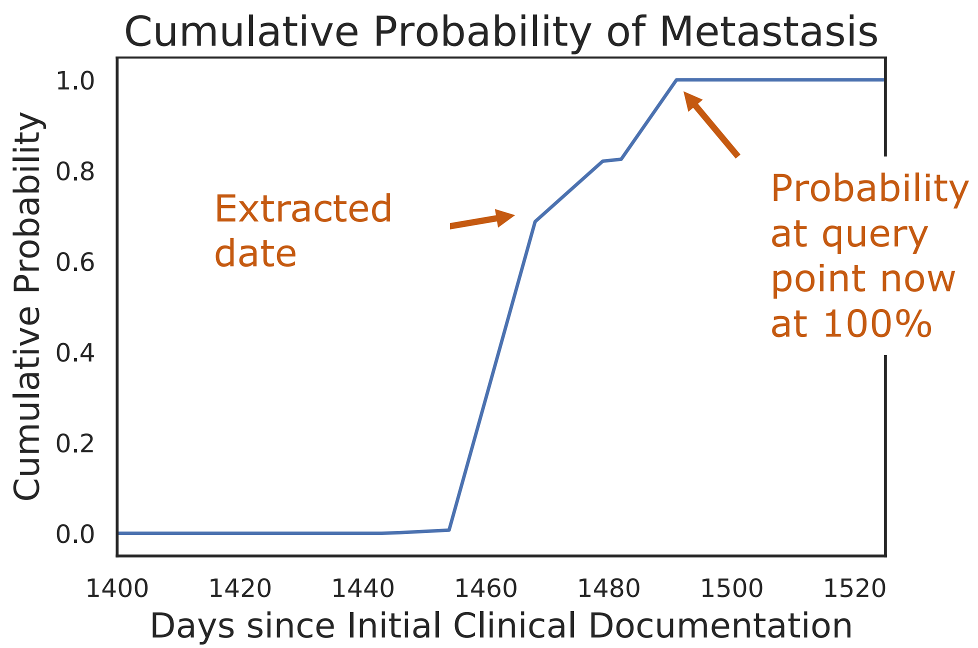
Metastasis Example #2: *Single query required*



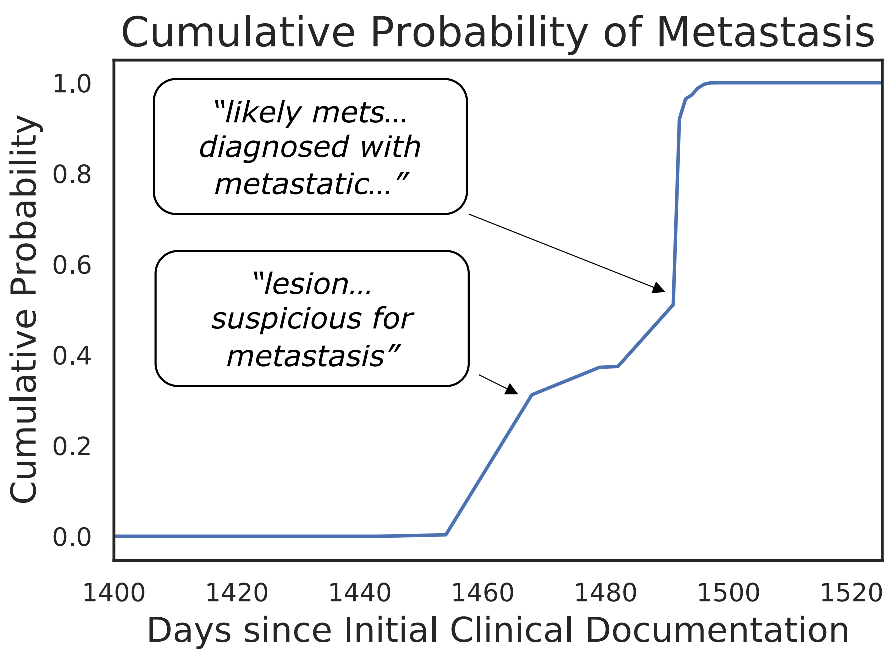
Metastasis Example #2: *Single query required*



Metastasis Example #2: *Single query required*



How to choose the most informative note?

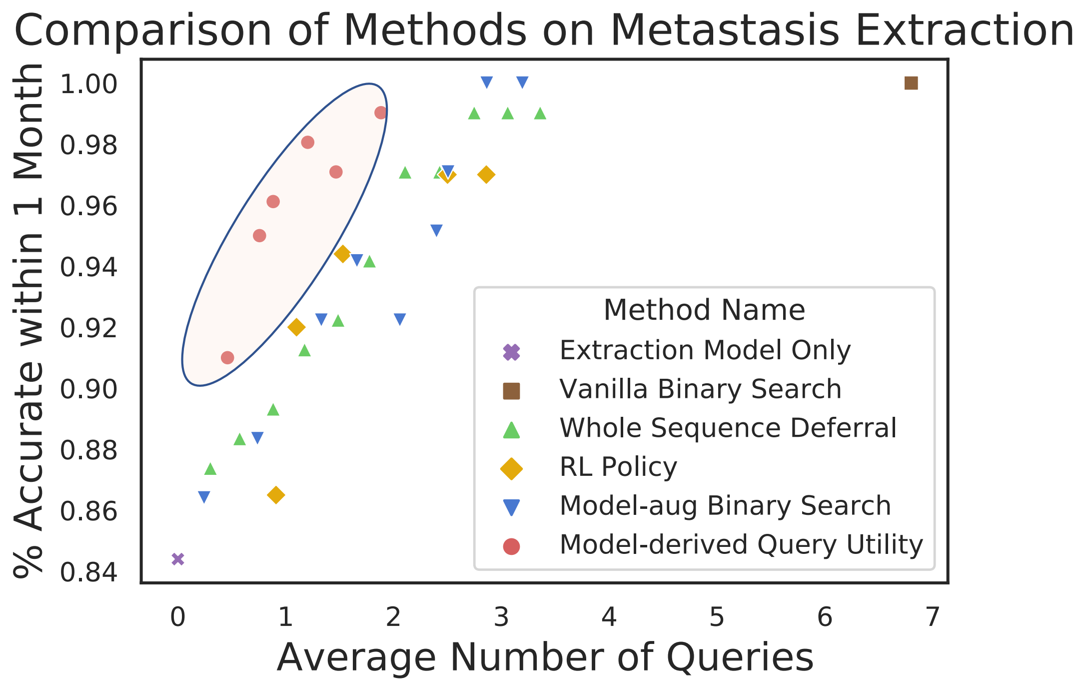


We develop a metric that chooses the note, which if labeled, would be **expected** to shift the estimate date by the **largest** number of days

Zhao and Agrawal et al., MLHC 2021

Experimental Results

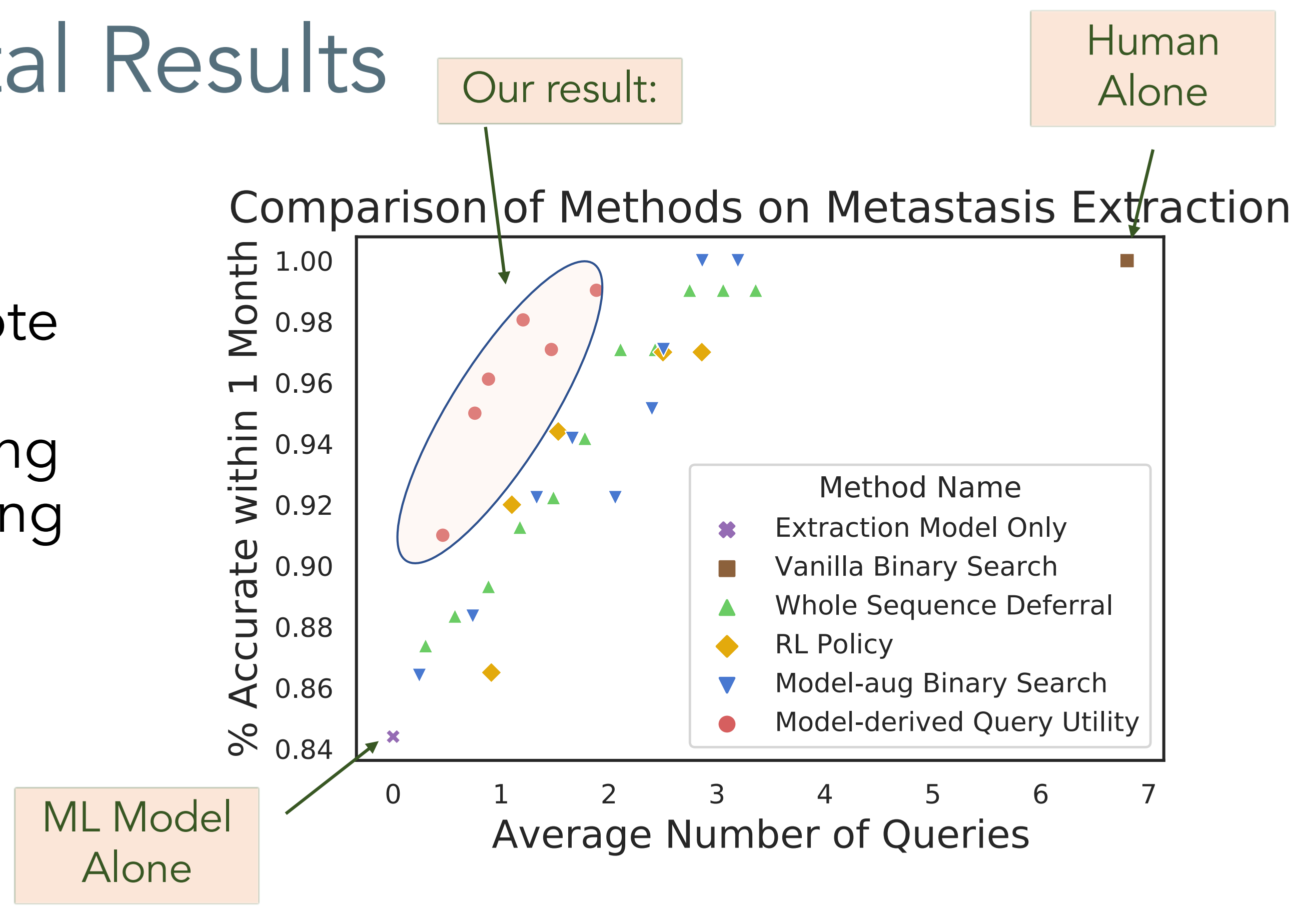
Our metric for note selection **outperforms** strong baselines, including an RL policy



Zhao and Agrawal et al., MLHC 2021

Experimental Results

Our metric for note selection **outperforms** strong baselines, including an RL policy



Zhao and Agrawal et al., MLHC 2021

This all seems like a lot of work

Why not change how we document?

A collaboration between MIT and the BIDMC Emergency Department

Gopinath et al MLHC 2020; Murray et al UIST 2021

David Sontag
Associate Professor
MIT

Steven Horng
MD, MMSc, FACEP
BIDMC

David Karger
Professor
MIT

Divya Gopinath
MEng Student
MIT

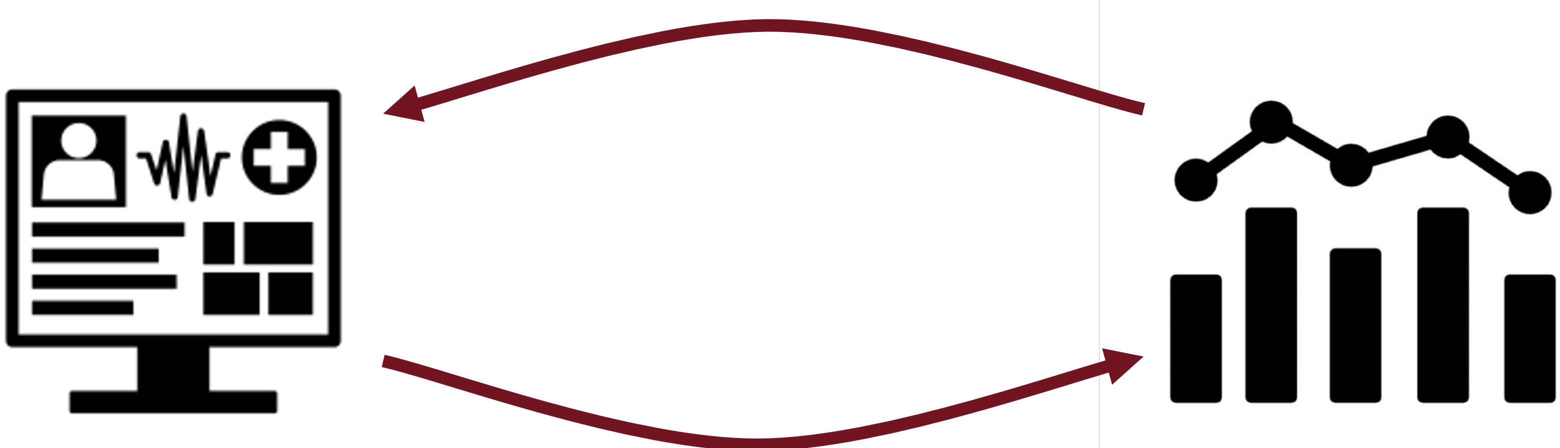
Luke Murray
PhD Student
MIT

Monica Agrawal
PhD Student
MIT

A collaboration between ML (Machine Learning) and HCI (Human Computer Interaction) Researchers and practicing Emergency Room Physicians

What we propose...

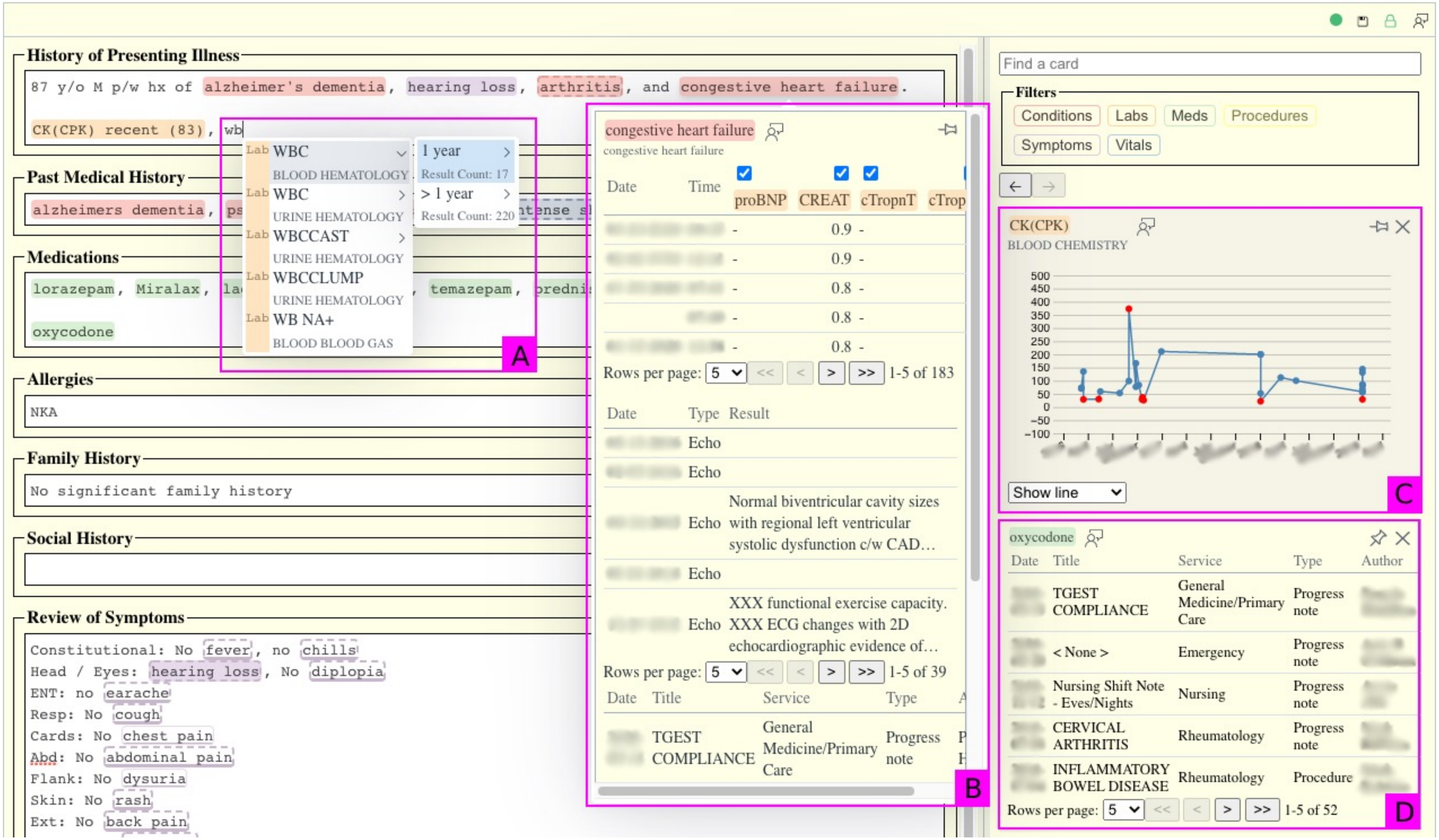
Incentivize cleaner clinical documentation via ML



Use cleaner data collected to build better ML models.

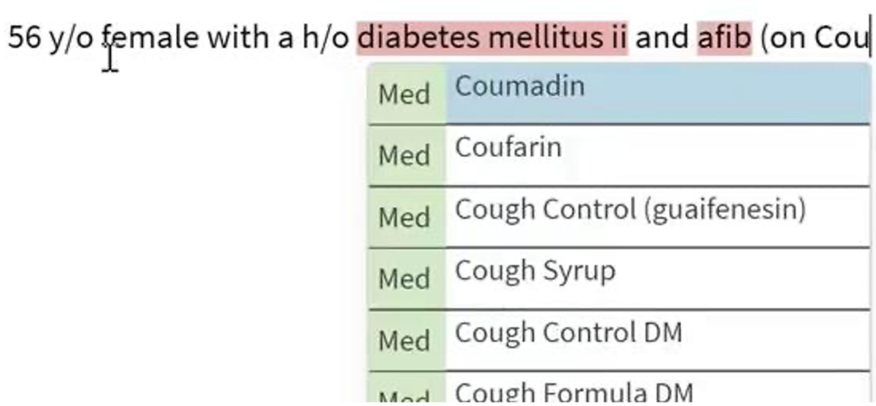
A demo:

https://www.youtube.com/watch?v=uqFOsPuRiS8



History of Presenting Illness
87 y/o M p/w hx of alzheimer's dementia, hearing loss, arthritis, and congestive heart failure.
CK(CPK) recent (83), wb
Lab WBC BLOOD HEMATOLOGY
1 year Result Count: 17
Lab WBC URINE HEMATOLOGY
> 1 year Result Count: 220
Lab WBCCAST URINE HEMATOLOGY
Lab WBCCLUMP URINE HEMATOLOGY
Lab WB NA+ BLOOD BLOOD GAS
A
Past Medical History
alzheimers dementia
Medications
lorazepam, Miralax, temazepam, oxycodone
Allergies
NKA
Family History
No significant family history
Social History
Review of Symptoms
Constitutional: No fever, no chills
Head / Eyes: hearing loss, No diplopia
ENT: no earache
Resp: No cough
Cards: No chest pain
Abd: No abdominal pain
Flank: No dysuria
Skin: No rash
Ext: No back pain
congestive heart failure
Date Time proBNP CREAT cTropnT cTrop
0.9
0.9
0.8
0.8
0.8
Rows per page: 5 1-5 of 183
Date Type Result
Echo
Echo
Echo Normal biventricular cavity sizes with regional left ventricular systolic dysfunction c/w CAD…
Echo
Echo XXX functional exercise capacity. XXX ECG changes with 2D echocardiographic evidence of…
Rows per page: 5 1-5 of 39
Date Title Service Type
TGEST COMPLIANCE General Medicine/Primary Care Progress note
B
Find a card
Filters
Conditions Labs Meds Procedures Symptoms Vitals
CK(CPK)
BLOOD CHEMISTRY
500 450 400 350 300 250 200 150 100 50 0 −50 −100
Show line
C
oxycodone
Date Title Service Type Author
TGEST COMPLIANCE General Medicine/Primary Care Progress note
< None > Emergency Progress note
Nursing Shift Note - Eves/Nights Nursing Progress note
CERVICAL ARTHRITIS Rheumatology Progress note
INFLAMMATORY BOWEL DISEASE Rheumatology Procedure
Rows per page: 5 1-5 of 52
D

Contextual autocomplete quickly captures clinical concepts at the point-of-care via learned, personalized suggestions.



Create semi-structured notes *as they are written.*

Decrease documentation burden for clinicians, who now type less.

Normalize concepts to clinical ontologies (UMLS, LOINC).

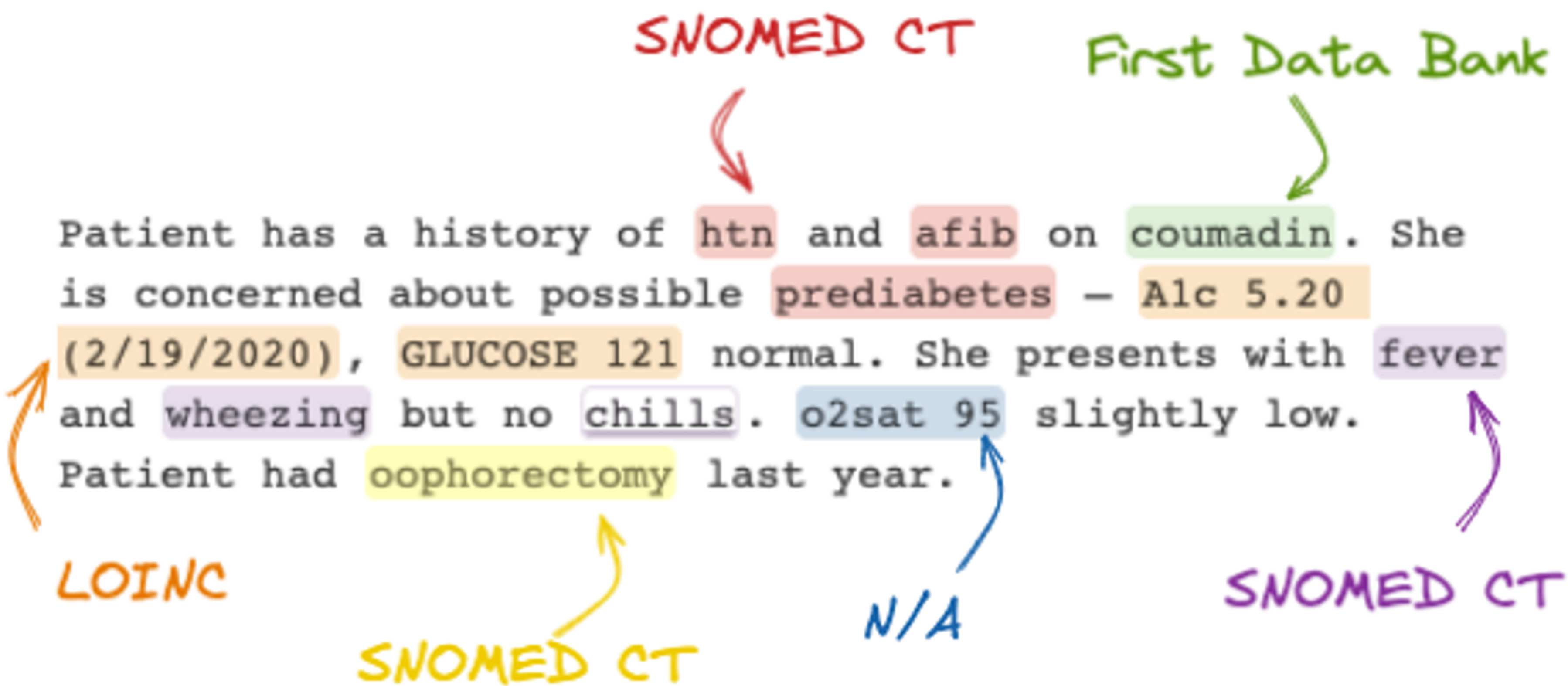
Chips

Condition Med

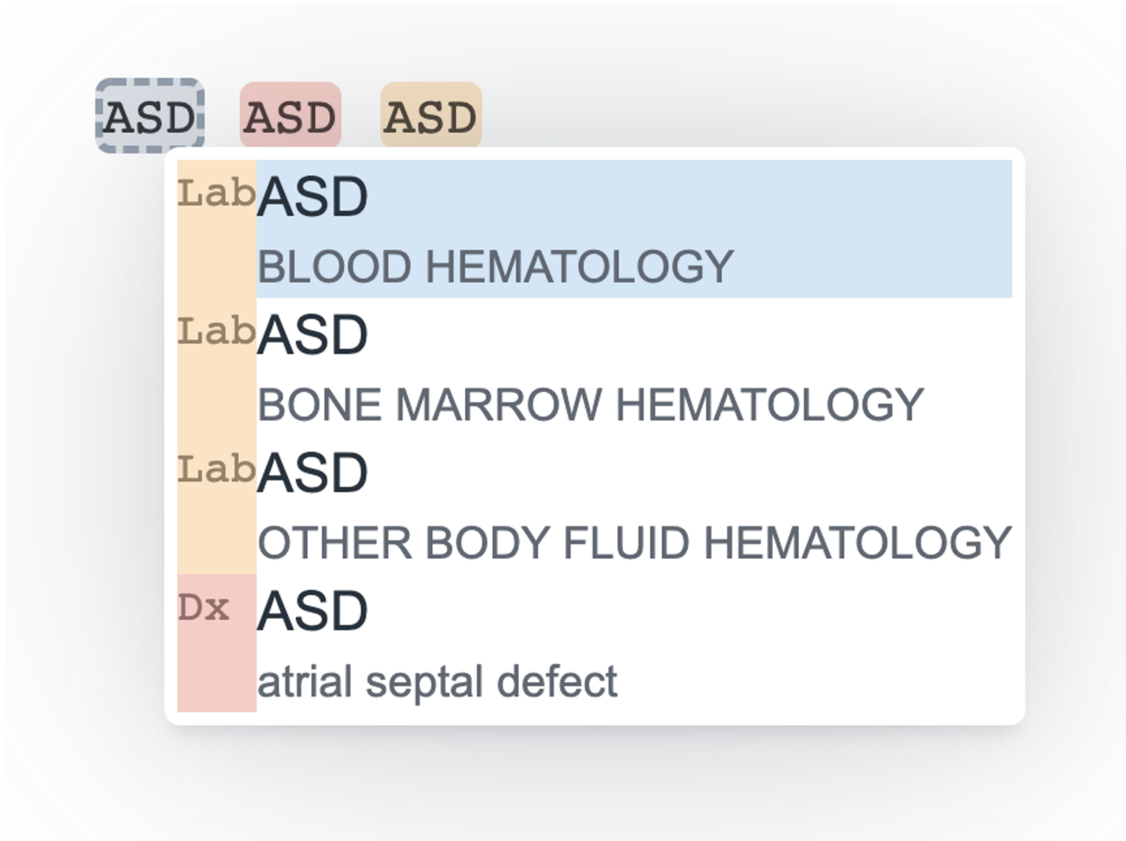
Patient has a history of htn and afib on coumadin. She is concerned about possible prediabetes — A1c 5.20 (2/19/2020), GLUCOSE 121 normal. She presents with fever and wheezing but no chills. o2sat 95 slightly low. Patient had oophorectomy last year.

Lab Procedure Vital Symptom

Link to Ontologies

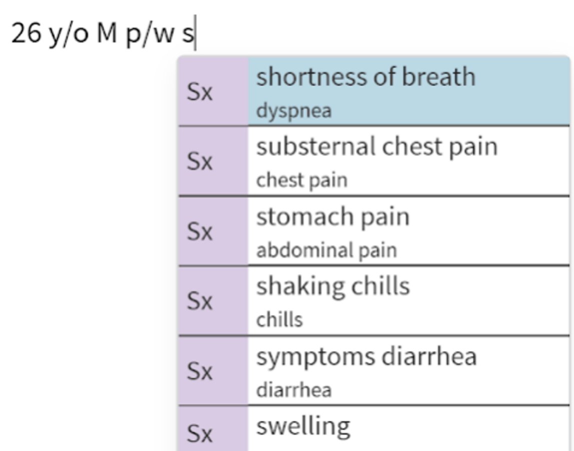


Disambiguation



Sources of information in contextual autocomplete:

Use available information from a given patient to predict concepts that will be documented in a clinical note.



(0) Prior notes (EHR)

(1) Triage assessment

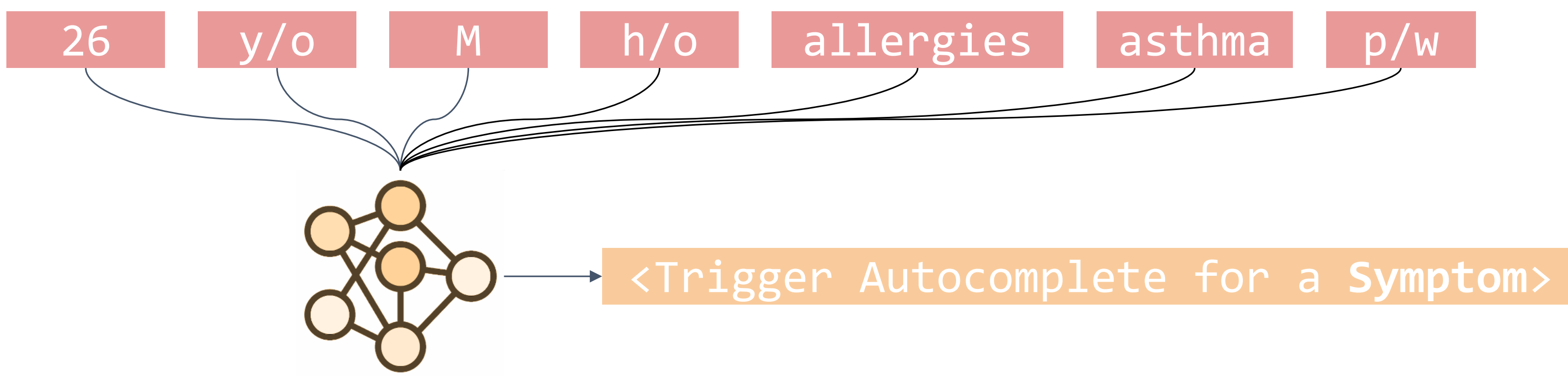
(2) Chief complaint

(3) Nurse's Notes

(4) Doctor's Notes (our focus)

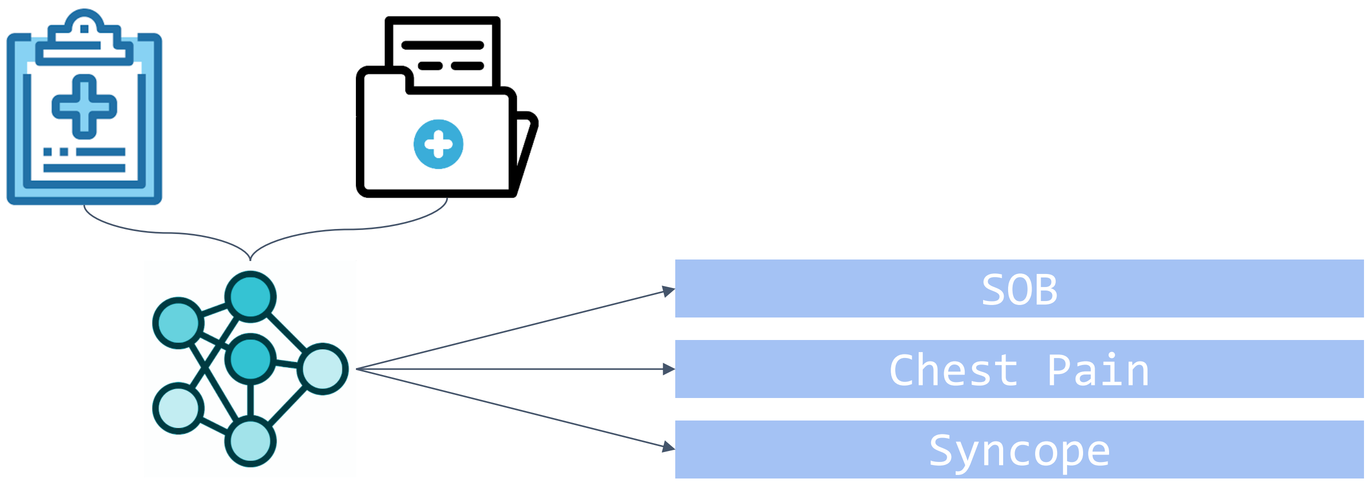
Deconstructing the Language Model

1 Use **local context** to predict **when** to autocomplete and **what concept type** is needed

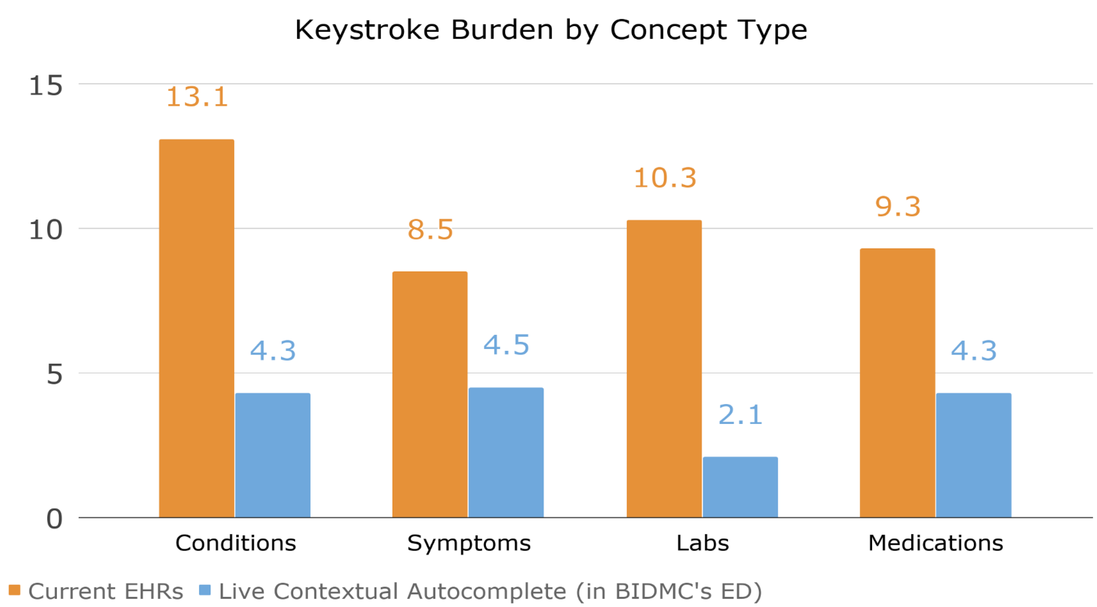


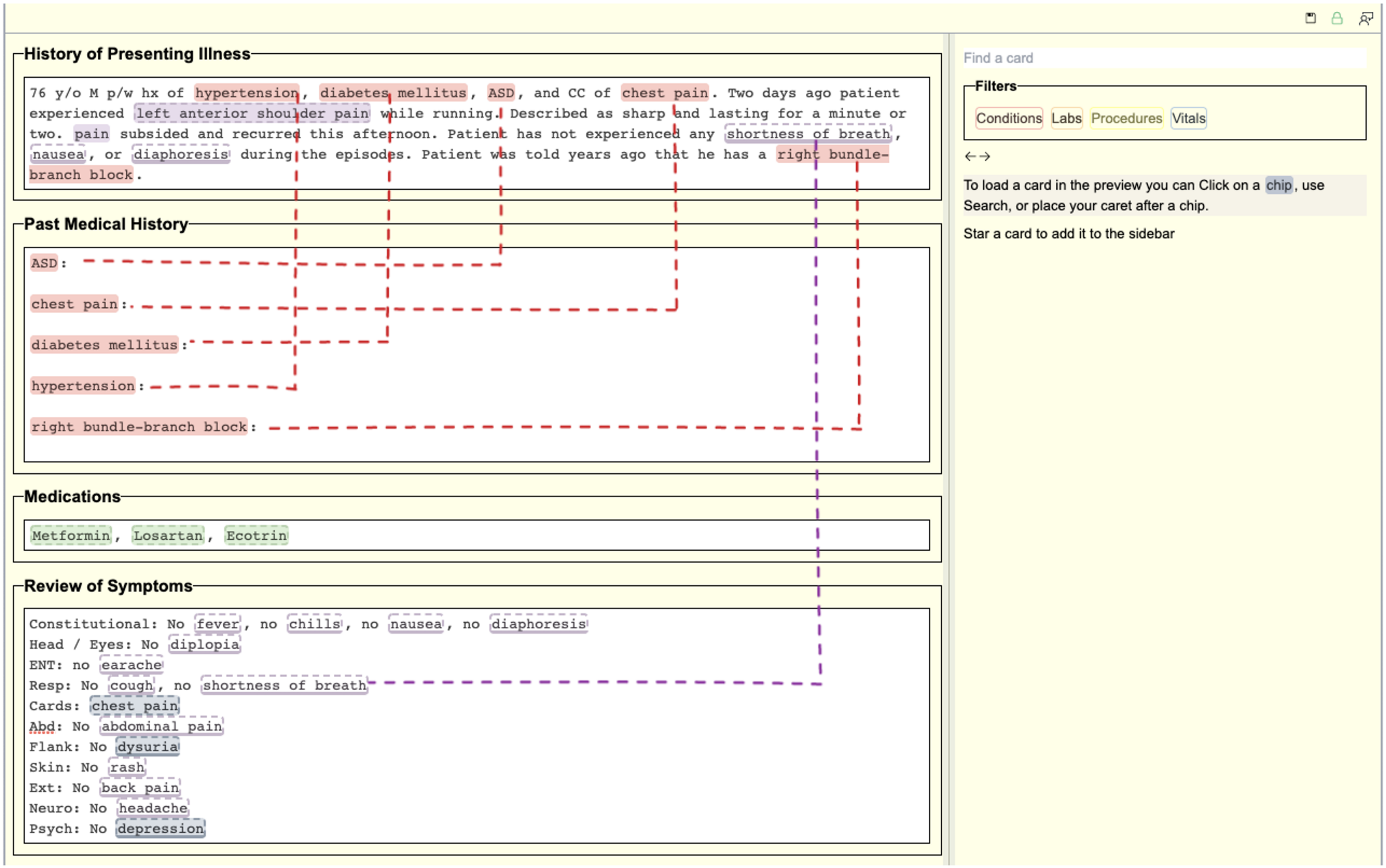
Deconstructing the Language Model

1. Use **local context** only to predict **when** to autocomplete and **what concept type** is needed
2. Use **past data** (triage note, medical history) to **rank suggested terms** to autocomplete for each concept type

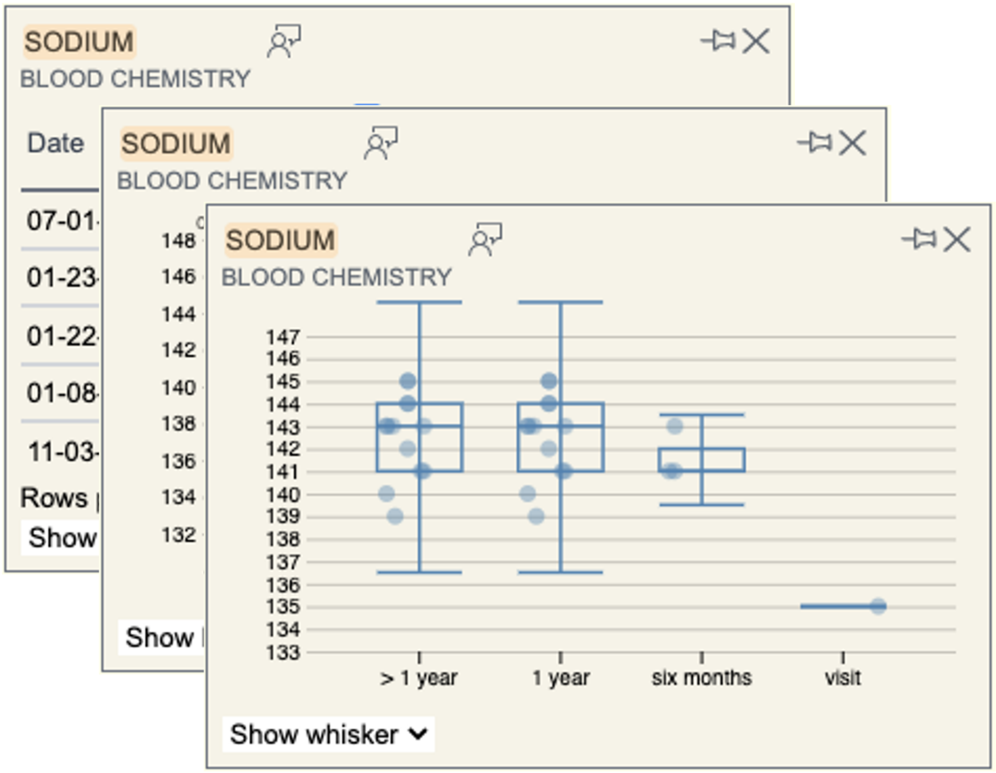


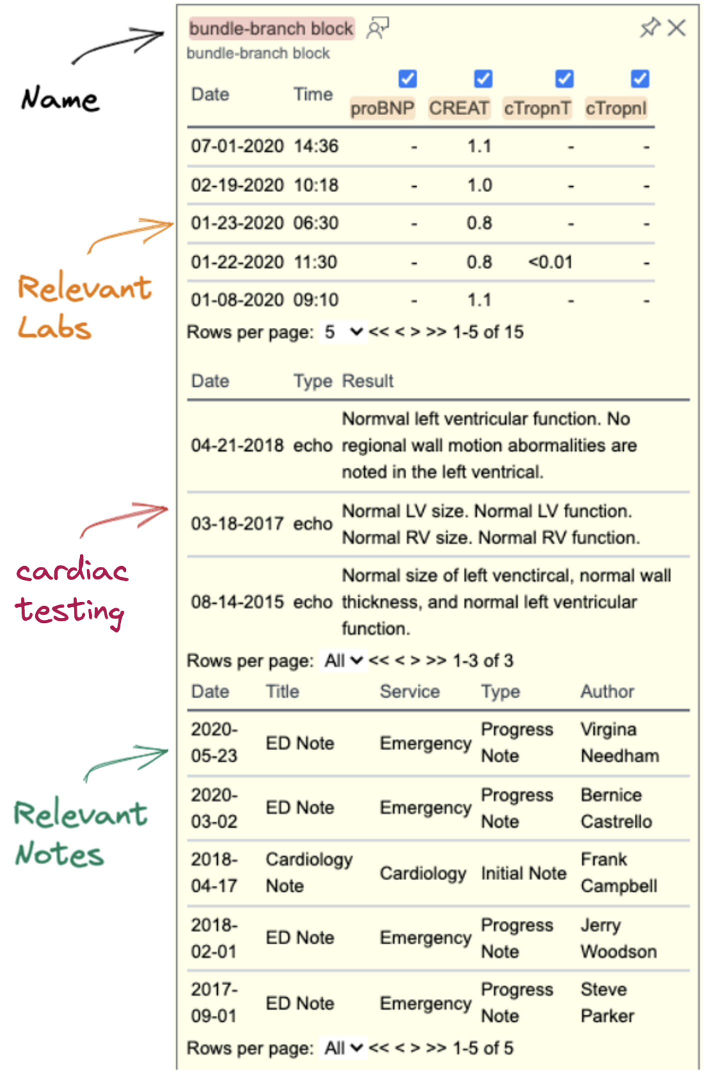
We dramatically reduced the **keystroke burden** of data entry in a **live setting.**




History of Presenting Illness
76 y/o M p/w hx of hypertension, diabetes mellitus, ASD, and CC of chest pain. Two days ago patient experienced left anterior shoulder pain while running. Described as sharp and lasting for a minute or two. pain subsided and recurred this afternoon. Patient has not experienced any shortness of breath, nausea, or diaphoresis during the episodes. Patient was told years ago that he has a right bundle-branch block.
Past Medical History
ASD:
chest pain:
diabetes mellitus:
hypertension:
right bundle-branch block:
Medications
Metformin, Losartan, Ecotrin
Review of Symptoms
Constitutional: No fever, no chills, no nausea, no diaphoresis
Head / Eyes: No diplopia
ENT: no earache
Resp: No cough, no shortness of breath
Cards: chest pain
Abd: No abdominal pain
Flank: No dysuria
Skin: No rash
Ext: No back pain
Neuro: No headache
Psych: No depression
Find a card
Filters
Conditions
Labs
Procedures
Vitals
← →
To load a card in the preview you can Click on a chip, use Search, or place your caret after a chip.
Star a card to add it to the sidebar

Concept Oriented Views





Date	Time	proBNP	CREAT	cTropnT	cTropnI
07-01-2020	14:36	-	1.1	-	-
02-19-2020	10:18	-	1.0	-	-
01-23-2020	06:30	-	0.8	-	-
01-22-2020	11:30	-	0.8	<0.01	-
01-08-2020	09:10	-	1.1	-	-

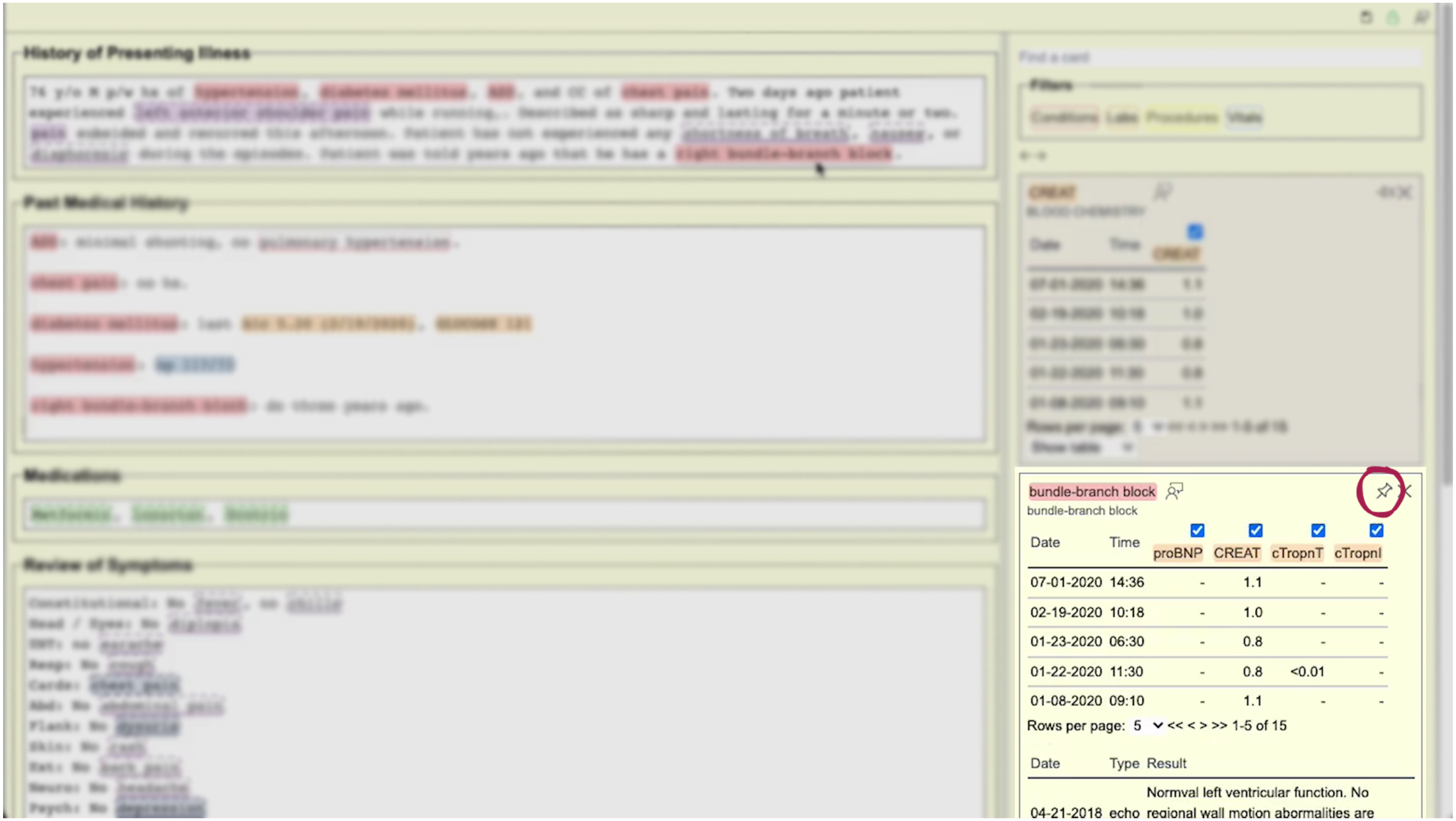
Rows per page: 5 << < > >> 1-5 of 15

Date	Type	Result
04-21-2018	echo	Normval left ventricular function. No regional wall motion abormalities are noted in the left ventrical.
03-18-2017	echo	Normal LV size. Normal LV function. Normal RV size. Normal RV function.
08-14-2015	echo	Normal size of left venctircal, normal wall thickness, and normal left ventricular function.

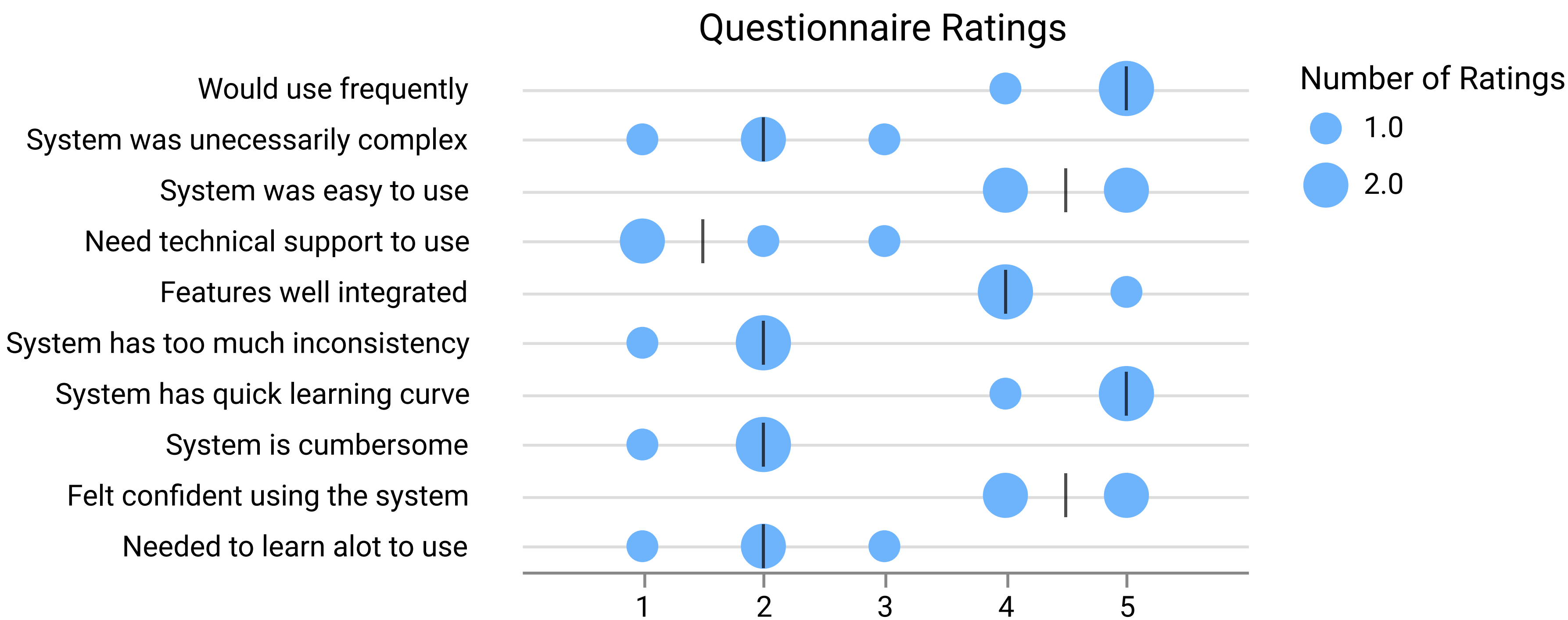
Rows per page: All << < > >> 1-3 of 3

Date	Title	Service	Type	Author
2020-05-23	ED Note	Emergency	Progress Note	Virgina Needham
2020-03-02	ED Note	Emergency	Progress Note	Bernice Castrello
2018-04-17	Cardiology Note	Cardiology	Initial Note	Frank Campbell
2018-02-01	ED Note	Emergency	Progress Note	Jerry Woodson
2017-09-01	ED Note	Emergency	Progress Note	Steve Parker

Rows per page: All << < > >> 1-5 of 5



bundle-branch block
bundle-branch block
Date
Time
proBNP
CREAT
cTropnT
cTropnI
07-01-2020 14:36 - 1.1 - -
02-19-2020 10:18 - 1.0 - -
01-23-2020 06:30 - 0.8 - -
01-22-2020 11:30 - 0.8 <0.01 -
01-08-2020 09:10 - 1.1 - -
Rows per page: 5 << < > >> 1-5 of 15
Date Type Result
04-21-2018 echo Normval left ventricular function. No regional wall motion abormalities are


Questionnaire Ratings
Would use frequently
System was unecessarily complex
System was easy to use
Need technical support to use
Features well integrated
System has too much inconsistency
System has quick learning curve
System is cumbersome
Felt confident using the system
Needed to learn alot to use
1
2
3
4
5
Number of Ratings
1.0
2.0

If you want to explore notes yourself…

Clinical NLP Datasets

- Physionet
  - MIMIC-III includes 2 million+ notes, including 65k discharge summaries; common tasks include readmission prediction
  - MIMIC-CXR contains notes to accompany chest x-rays
- Clinical Abbreviation Sense Inventory (CASI)
  - Consists of ~40k examples of medical abbreviations in a short context
  - Contains 440 abbreviations/acronyms with 949 sense

Clinical NLP Datasets

- n2c2 NLP Research Data Sets
  - Tasks include deidentification, entity recognition, adverse drug event extraction, etc.
- TREC Biomedical Track
  - Annual competition with open-ended tasks such as matching patients to trial, identifying relevant studies
  - Submissions are manually evaluated by experts

Feel free to contact me
with any questions!

magrawal@mit.edu