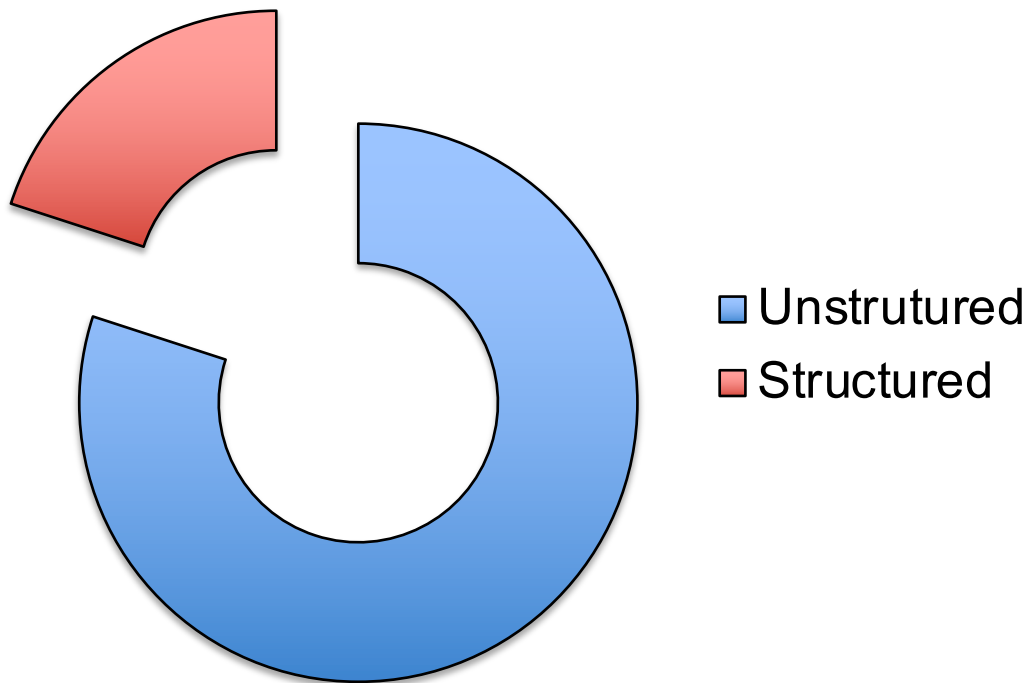


# **EMERSE: the Electronic Medical Record Search Engine**

David Hanauer  
University of Michigan  
ITCR Annual Meeting  
May 29, 2019

# 80% OF EHR DATA UNSTRUCTURED



<http://www-07.ibm.com/hk/watson/health/>

**Unstructured text = Free text**

“free” means clinicians are free to write anything they want, any way they want.

But the text is hardly “free” to use



# STRUCTURED VS UNSTRUCTURED

## Structured Data



**Lab results**



**Prescribed medications**



**Billing codes**



**Problem lists**

## Unstructured Data



**Radiology reports**



**Pathology reports**



**Clinical notes**

# FREE TEXT USES



**cohort identification**



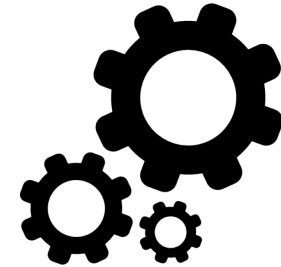
**eligibility determination**



**adverse event  
detection**



**data abstraction**



**much more!**



# NEED FOR FREE TEXT



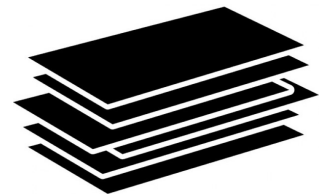
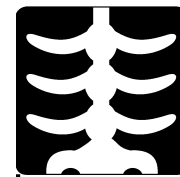
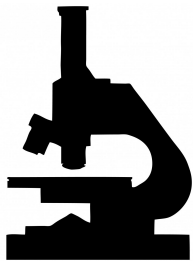
## eligibility determination

Free text needed for resolving...

...**59%** of eligibility criteria for a chronic lymphocytic leukemia (CLL) clinical trial

...**77%** of eligibility criteria for a prostate cancer clinical trial

# RESEARCHERS NEED TOOLS



# EMERSE – INFORMATION RETRIEVAL

- Electronic Medical Record Search Engine

Information Retrieval (IR) tool

- Users familiar with Google, Bing, etc.
  - Flexible for many information needs
- 
- Similar to, but still different from, Natural Language Processing (NLP)
    - NLP can be powerful, but generally requires expertise (in short supply)
    - EMERSE meant for real-time, self-service data interactions
- \*\* EMERSE (IR) and NLP are not mutually exclusive

# EMERSE

- Developed at U of Michigan
- In use at Michigan since 2005
- 113,000 user sessions since 2012 (2,085 in last 30 days)
- 1,738 studies supported
- 209 (known) peer-reviewed publications (~12% citation rate)
- Now also available at: University of North Carolina, University of Cincinnati.
- Planning stages: Columbia Univ, Case Western Reserve Univ, Univ of Kentucky, UCSF, Dana Farber, Memorial Sloan Kettering

# RECENT PUBLICATIONS

- Tsao PA et al. **Cardiovascular and Metabolic Toxicity of Abiraterone in Castration-Resistant Prostate Cancer: Post-Marketing Experience.** Clinical Genitourinary Cancer. March 2019.
- Chappell G et al. **Maintenance sorafenib in FLT3-ITD AML following allogeneic HCT favorably impacts relapse and overall survival.** Bone Marrow Transplant. 2019 Feb 26.
- Manohar PM et al. **Prognostic Value of FDG-PET/CT Metabolic Parameters in Metastatic Radioiodine-Refractory Differentiated Thyroid Cancer.** Clin Nucl Med. 2018 Sep;43(9):641-647.
- Siontis BL et al. **Primary Cardiac Sarcoma: A Rare, Aggressive Malignancy with a High Propensity for Brain Metastases.** Sarcoma, vol. 2019, Article ID 1960593, 6 pages, 2019.

# WHAT WAS SEARCHED? (yesterday)

"no sign cancer"; recurrence; remission; "disease free"; lvef; "phosphaturic mesenchymal tumor"; trandolapril; azilsartan; cozaar; irbesartan; normodyne; nadolol; enema; oncotype; proctitis; ruxolitinib; laxative; risk; and; pulmozyme; "no cancer"; flutter; edema; "ace inhibitor"; olmesartan; candesartan; carvedilol; bystolic; zebeta; monocor; bisoprolol; metoprolol; "breast cancer"; sma; bleeding; initiative; jakafi; preterm; transsphenoidal; cushing's; coagu\*; "9520 like"; thromb\*; nyha; glucose; breath; "coronary angiography"; photosensitivity; pentasa; fai; fibrillation; "lv ef"; "ventricular ejection fraction"; moexipril; enalapril; accupril; quinapril; teveten; pindolol; timolol; betoptic; fracture; coagulopathy; mycobacterium; diarrhea; sts; "source radiology"; "gastric leiomyoma"; tibial; "congenital hypertrophy of the retinal pigment epithelium"; sweats; gender; **doxorubicin**; mets; bundle; ef; nexium; "low grade spindle cell neoplasm"; lotensin; diovan; betapace; lobeta; penbutolol; coreg; trandate; betaxolol; tenormin; propranolol; corgard;; hla; dulera; advair; shock; hemorrhage; bleed; "sinonasal melanoma"; "rectal inflammation"; immunocompromised; des; abscess; chf; pe; femoral; **metas\***; ned; canasa; "beta blocker"; vasotec; zestril; perindopril; carteolol; labetalol; nebivolol; atenolol; lopressor; cmv; tacrolimus; injured; "superior mesenteric"; "blood loss"; arthritis; ana; flovent; qvar; "stomach leiomyoma"; stent; hydrogen; conjunctivitis; **adriamycin**; manometry; "no residual"; "cancer free"; failure; arb; benazepril; altace; ramipril; valsartan; eprosartan; cilexetil; brevibloc; inderal; toprol; cardiogenic; injury; anticoag\*; complicate\*; hemperitoneum; pituitary; smith; asmanex; metal; dvt; fluid; test; coronary; angiography; flail; paraplegia; embol\*; "lynch syndrome"; mitoxantrone; afib; "left ventricular ejection fraction"; 2004; claudication; mavik; aceon; hyzaar; avapro; ocupress; visken; sotalol; acebutolol; pmt; hematoma; "distal colitis"; **genomics**; **lumpectomy**; pulmicort; score; echo; dvu; endocarditis; **epirubicin**; "no evidence"; "no tumor recurrence"; respiratory; mgi; branch; "ejection fraction"; "ventricular ejection"; pantoprazole; omeprazole; "sts score"; imuran; lisinopril; fosinopril; captopril; telmisartan; benicar; losartan; esmolol; sectral; donor; symbicort; 6mp; "chest pain"; **daunorubicin**; **idarubicin**;

User: hanauer

EMERSE

## 116 patients matched the search criteria

100 top-ranked document summaries are shown.

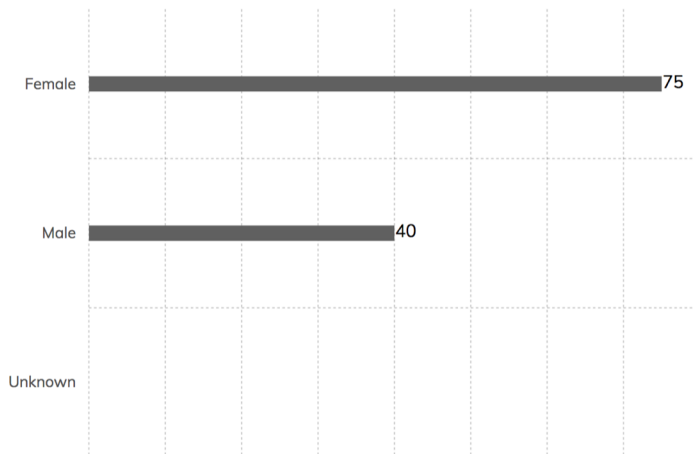
To review these patients in more detail, move them to a Temporary Patient List and then run the search again.

Move patients to Temporary Patient List

Revise Terms

☒ Sort charts by Demographics ☐ Sort charts by Frequency

### Gender



### Summary

...suggestive of epithelioid hemangioendothelioma. See COMMENT...  
...includes epithelioid hemangioendothelioma, poorly differentiated...  
...considerations, epithelioid hemangioendothelioma is our primary...

...History of epithelioid hemangioendothelioma. About 30...  
...with known epithelioid hemangioendothelioma. This scan...  
...history of epithelioid hemangioendothelioma. Please note...

...HISTORY: Epithelioid hemangioendothelioma. COMPARISON:...  
...diagnosis of epithelioid hemangioendothelioma. There is...  
...diagnosis of epithelioid hemangioendothelioma....

...COMPLAINT: Epithelioid hemangioendothelioma of the liver...  
...angiosarcoma versus epithelioid hemangioendothelioma. He was  
evaluated...  
...consistent with epithelioid hemangioendothelioma. The latest...  
...diagnosed epithelioid hemangioendothelioma of the liver...  
...liver biopsy) Epithelioid hemangioendothelioma. CD31 highlights...  
...man with epithelioid hemangioendothelioma of the liver...  
...liver. Epithelioid hemangioendothelioma in the liver...  
... Since epithelioid hemangioendothelioma is indolent...

...hepatic epithelioid hemangioendothelioma and history...  
...needle biopsy: Epithelioid hemangioendothelioma. ...

...results/plan for her epithelioid hemangioendothelioma. Labs from...  
...metastatic epithelioid hemangioendothelioma. CTs have...

...): Epithelioid hemangioendothelioma. Immunohistochemical...  
...): Epithelioid hemangioendothelioma. Immunohistochemical...

...of case: Epithelioid hemangioendothelioma. GROSS:...  
...): Epithelioid hemangioendothelioma. COMMENT:...

...metastatic epithelioid hemangioendothelioma with enlarged...  
...Metastatic epithelioid hemangioendothelioma. ...

...consistent with epithelioid hemangioendothelioma seen in subsequent...  
...({): Epithelioid hemangioendothelioma (CD31, factor...

Patients

Demo List (new) (12 Patients)

User: emerse

EMERSE

Dates

All Dates: 02/15/2008 through 09/28/2011

Terms

abdomen

therapy

lungs

"cancer of the brain"

"cancer of brain"

"tumor of brain"

"pontine glioma"

glioma

craniopharyngioma

"meninges cancer"

"malignancy of brain"

"CNA ca"

"CNS malignancy"

"tumor o

More...

Overview

Overview

Sorted By:

Insert Order

Ascending

Numbers

Grayscale

Mosaic

MRN	Patient Name	Careweb	Radiology	Pathology
1000000049	Bloom, Harrison	7 of 20	6 of 30	15 of 20
1000000047	Patel, Joshua	9 of 10	5 of 20	3 of 30
1000000036	SCOUTTEN, MARILYN	21 of 25	1 of 10	1 of 5
1000000073	Errazuriz, Alberto	2 of 10	1 of 10	13 of 15
1000000048	Fay, Pat	9 of 10	1 of 10	2 of 10
1000000040	Chen, John	9 of 10		4 of 20
1000000035	LUCCHESI, VINCENZO		15 of 15	2 of 10
1000000075	Sarchand, Nandita	1 of 5	1 of 5	9 of 10
1000000065	Ladwig, Renske		2 of 10	5 of 5
1000000050	Jacobs, Kelly		2 of 10	5 of 5
1000000072	Walsh, Alan		1 of 10	
1000000044	Taylor, Jonathon	3 of 5		

(fake names)



Patients

Demo List (new) (12 Patients)

User: emerse

EMERSE

Dates

All Dates: 02/15/2008 through 09/28/2011

Terms

cd20

cd5

cd34

cd10

cd3

cd45

"marginal zone lymphoma"

"BCL 2"

BEXXAR

rituximab

cytoxan

"R CHOP"

anthracycline

"autologous stem cell transplant"

"diffuse large cell lymphoma"

"salvage therapy"

More...

Overview

Overview

Sorted By:












Insert Order

Ascending

Numbers

Grayscale

Mosaic

MRN	Patient Name	Careweb	Radiology	Pathology
1000000049	Bloom, Harrison			
1000000047	Patel, Joshua			
1000000036	SCOUTTEN, MARILYN			
1000000073	Errazuriz, Alberto			
1000000048	Fay, Pat			
1000000040	Chen, John			
1000000035	LUCCHESI, VINCENZO			

Patients

Demo List (new) (12 Patients)

User: emerseEMERSE

Dates

All Dates: 02/15/2008 through 09/28/2011

Terms

cd20

cd5

cd34

cd10

cd3

cd45

"marginal zone lymphoma"

"BCL 2"

BEXXAR

rituximab

cytoxin

"R CHOP"

anthracycline

"autologous stem cell transplant"

"diffuse large cell lymphoma"

"salvage therapy"

More...

Overview

Summaries

Patel, Joshua

1000000047

Comment

Tag

Patient 2 of 12

Document 4 of 10

Page Top

Page Bottom

Print

ADMITTING DIAGNOSIS: Relapsed Diffuse large cell lymphoma

TRANSPLANT TYPE: Autologous

PROTOCOL: CTN 0401 (Bexxar – BEAM)

TRANSPLANT DATE: \*\*DATE

VIRAL TITERS: CMV positive, HSV positive, VZV positive

HISTORY OF PRESENT ILLNESS: Mr. \*\*NAME is a \*\*AGE year old male who is being admitted for an autologous stem cell transplant per BMT CTN ; randomized to Bexxar-BEAM therapy. Mr. \*\*NAME is currently admitted to undergo the BEAM portion of his conditioning regimen. Autologous stem cell transplant is scheduled for \*\*DATE.

PAST MEDICAL HISTORY: The patient presented in \*\*DATE, with proptosis of the right eye and sinus congestion. A CT scan at that time revealed a 3.4 cm x 5.8 cm right orbital mass, and 6.7 cm x 3.7 cm right maxillary mass. The patient had developed significant exophthalmos of the right eye with visual changes along with coexistent bilateral cervical adenopathy. A biopsy of the mass was obtained on \*\*DATE, initially consistent with a CD20 positive, CD45 positive, BCL-2 positive, marginal zone lymphoma. The patient was subsequently treated with 5 cycles of R-CHOP between \*\*DATE, and \*\*DATE.

A second-look surgery performed on \*\*DATE, of the right ethmoid and right maxillary sinus mass revealed that the lesion had transformed to a diffuse large cell lymphoma, again CD20 positive, CD45 positive, BCL-2 positive. The patient subsequently received 4 additional cycles of R-CHOP between \*\*DATE, through \*\*DATED, YYYY>. He has now received a total anthracycline dose of 240 mg/sq. m. A CT scan obtained at the end of induction therapy revealed a CR status. No XRT was given.

In \*\*DATE, the patient developed recurrent nasal congestion with the lacrimal duct obstruction. A CT scan on \*\*DATED, YYYY>, revealed a new left maxillary mass, 4.5 cm in maximal diameter. The mass extended from the midline of the oropharynx laterally to the point of the lateral pterygoid plate on the left side. Thus, the patient recurred contralateral to the initial side of disease. There was no other adenopathy in the chest or abdomen on screen at that time. The patient started salvage therapy with RICE \*\*DATE. He has received 2 cycles of RICE to date, the 2nd cycle given \*\*DATE. He enrolled on CTN in \*\*DATE, receiving Rituximab , and Cytosin 4 gm/m2 . He underwent autologous stem cell collection \*\*DATED>--\*\*DATE, with 4.4 x 10(6) CD34 cells/kg collected.

In preparation for his Bexxar dosimetry on \*\*DATE, the patient

PatientsAll (2,351,932)

DatesAll Dates: 01/01/1900 through 05/18/2018

Termsemesis

Quick Terms

Term Bundles

Advanced Terms

Name/Description

Manage Terms

View Terms

Sharing

Clear/Delete

The current settings will enable you to search through All Patients. Terms with different colors will be searched using the AND operator. Patients identified will be those that have at least one document. With a Quick Terms search, EMERSE assigns the colors to the terms. The current screen looks similar to a Term Bundle screen, but be able to adjust the colors of the terms yourself. For more control over the colors assigned to terms, use a Term Bundle. For more control over your search criteria, such as combining Boolean operators, use the Advanced Terms screen.

Search Options

Terms to include

Click on the term to edit

emesis

Synonyms

Click individual terms to highlight or de-highlight

Vertical View

X

Synonyms (119)

barfbarfedbarferbarfingbarfsblowing chunkschunderchunderedchunderingchundersdry heavedry heaveddry heavesdry heavingemesesemeticemeticsemetogenicemetogenicityheaveheavedheavesheavinghurlhurledhurlerhurlinghurlsdry heavingemesesemeticemeticsemetogenicemetogenicityheaveheavedheavesheavinghurlhurledhurlerhurlinghurlsinduce vomitinginduced vomitinginduces vomitinginducing vomitingN & VN and VN & VN/Vnauseanausea & vomitingnausea / vomitingnausea and vomitingnausea or vomitingnausea vomitingnausea/vomitingnauseatenauseatednauseatesnauseatingnauseousnauseousnessNBNB emesisnon-bloody, non-bilious emesisnon-bloody, nonbilious emesisnonbloody, non-bilious emesisnonbloody, nonbilious emesisNVpost-anesthesia emesispost-anesthesia vomitingpost-anesthetic emesispost-anesthetic vomitingpost-tussive emesispostanesthesia emesispostanesthesia vomitingpostanesthetic emesispostanesthetic vomitingposttussive emesispukepukedpukerpukespukingpurgepurgedpurgerpurgespurgingqueasinessqueasyralphralphsregurgitateregurgitatedregurgitatesregurgitatingretchretchedretchesretchingreverse peristalsisthrew it all upthrew it upthrew upthrow it all upthrow it upthrowing chunksthrowing it all upthrowing it upthrowing upthrows it all upthrows it upthrows uptossing cookiesup-chuckup-chuckedup-chuckingup-chucksupchuckupchuckedupchuckingupchucksvomitivomitedvomitervomitersvomitingvomitivevomitoryvomitsvomitus

Spelling Alternatives (9)

emeisemeisisemeissemisemismeticmeticsnause and vomitingqueaziness

Highlight All

De-Highlight All

Add Highlighted Terms

Patients

Demo List (new) (12 Patients)

User: emerse

EMERSE

Dates

All Dates: 02/15/2008 through 09/28/2011

Terms

emesis

Overview

Overview

Sorted By: 

Insert Order

Ascending

Numbers

Grayscale

Mosaic

MRN	Patient Name	Careweb	Radiology	Pathology
100000049	Bloom, Harrison			
100000047	Patel, Joshua			
100000036	SCOUTTEN, MARILYN			
100000073	Errazuriz, Alberto			
100000048	Fay, Pat			
100000040	Chen, John			
100000035	LUCCHESI, VINCENZO			

Overview

Overview

Sorted By:

Insert Order

Ascending

Numbers

Grayscale

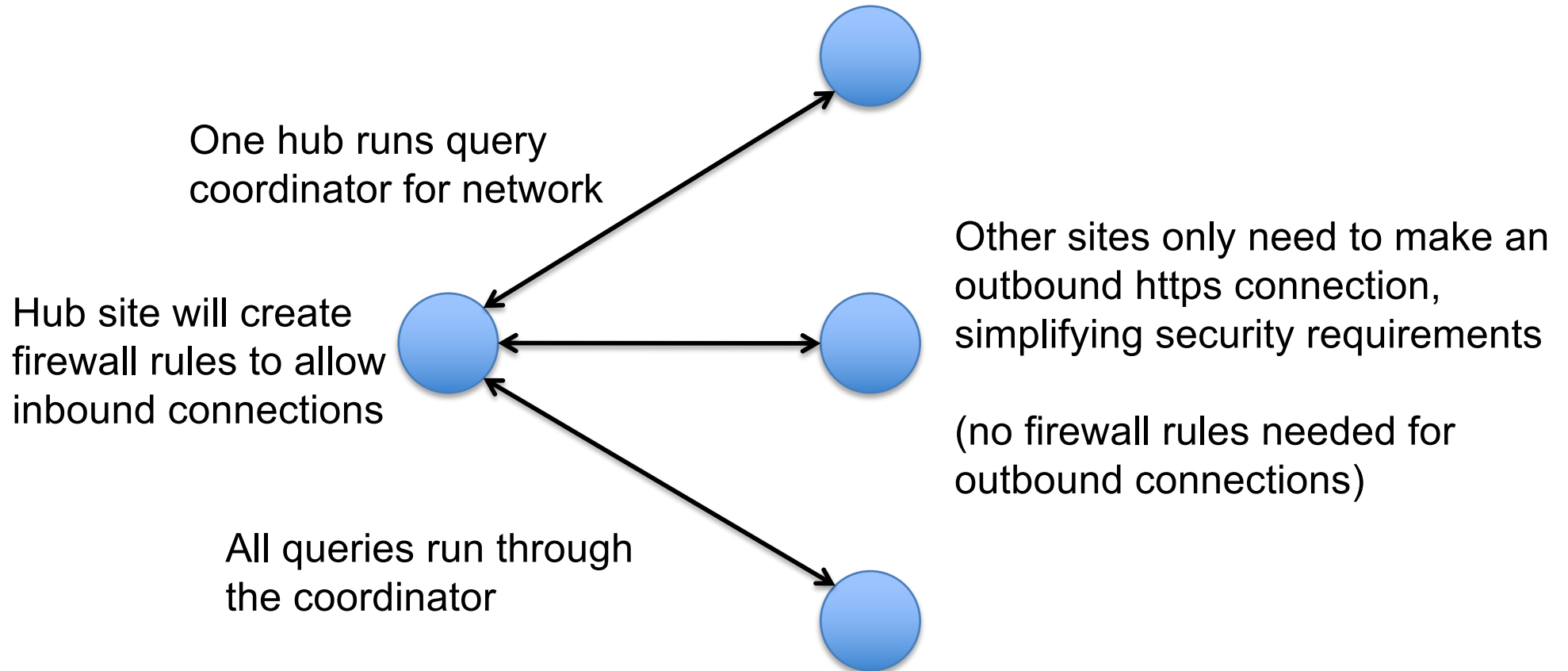
Mosaic

MRN	Patient Name	Careweb	Radiology	Pathology
1000000049	Bloom, Harrison			
1000000047	Patel, Joshua			
1000000036	SCOUTTEN, MARILYN			
1000000073	Errazuriz, Alberto			
1000000048	Fay, Pat			
1000000040	Chen, John			
1000000035	LUCCHESI, VINCENZO			

# NETWORKING

- Search across sites for counts
- Similar concept to other networks for structured data (PCORnet, i2b2)
- Many security considerations
- Architecture designed for simplicity at other sites

# NETWORKING ARCHITECTURE



# NETWORKING - SECURITY

- Obfuscated counts only – obfuscation customizable
- Customizable thresholds (e.g.,  $< 20$ ) to prevent pinpointing exact count of individual
- Throttle/randomize return time for queries
- Throttle # queries per user per time
- Audit logs of all users/queries across the network
- Meet all requirements by our *Information Assurance* team (the “internal hackers”)



# EMERSE TRENDS

- Similar idea to Google Trends
  - Google Trends – search terms over time
  - EMERSE Trends – patients over time
- A good way to explore data for patterns over time

# GOOGLE TRENDS

● "car t cell"  
Search term

+ Compare

United States ▼

2004 - present ▼

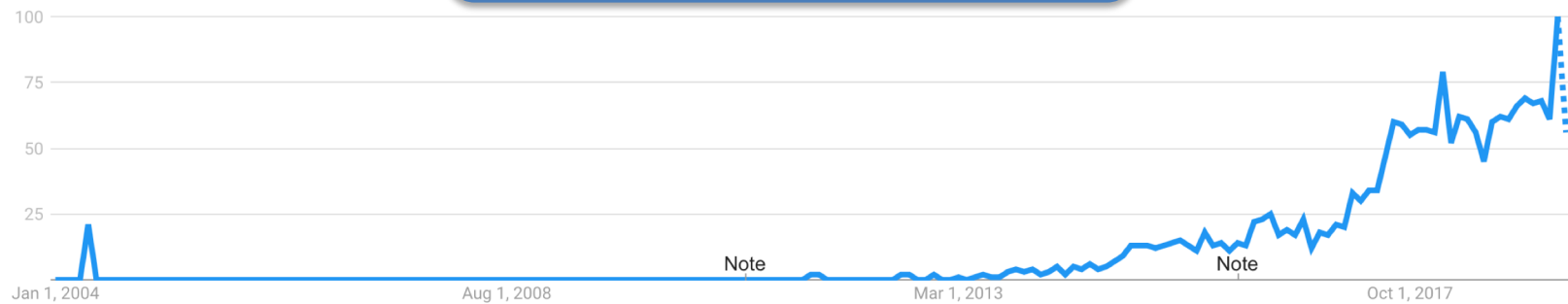
All categories ▼

Web Search ▼

Interest over time ⓘ



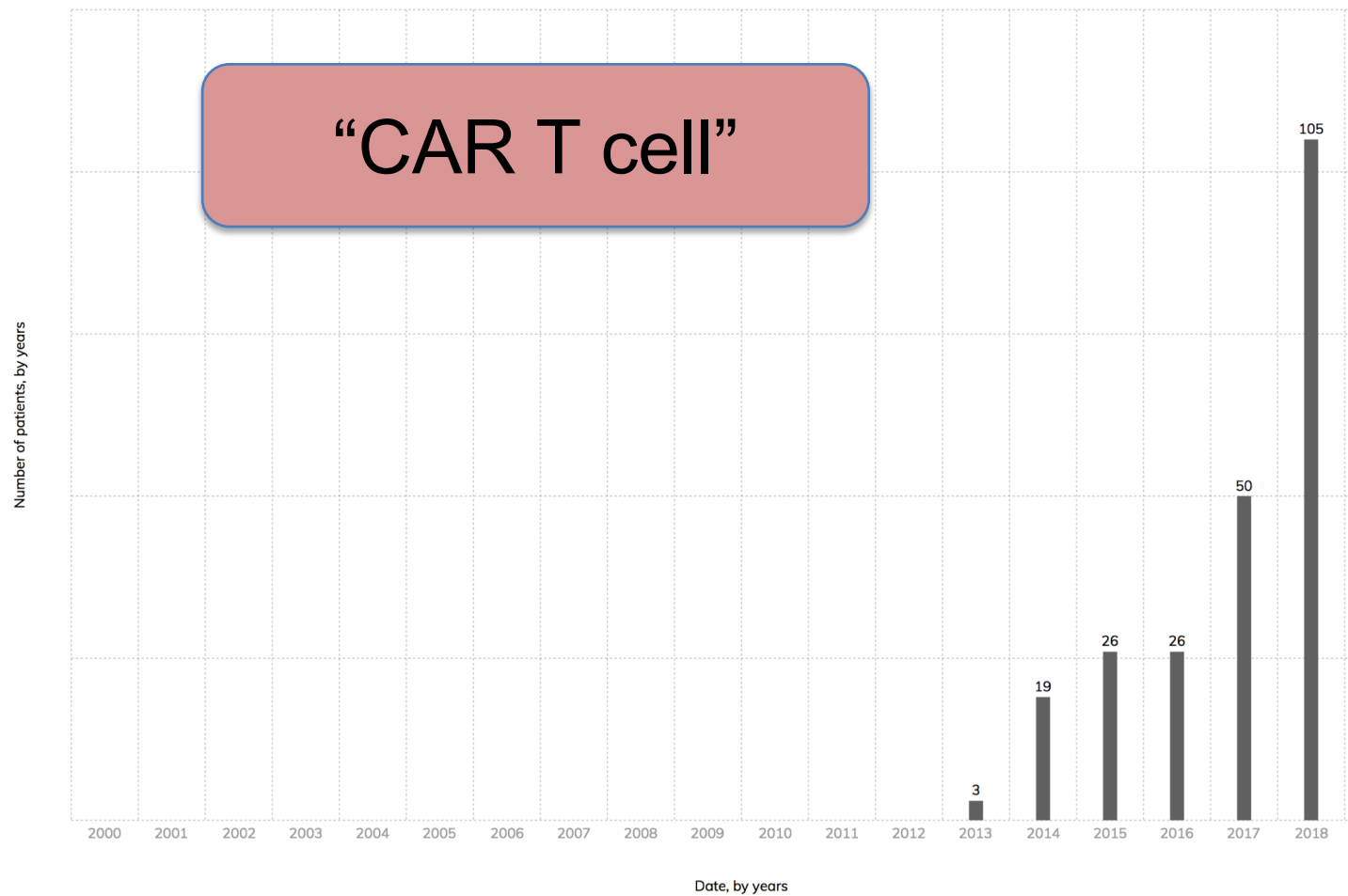
“CAR T cell”



Terms

"car t cell"

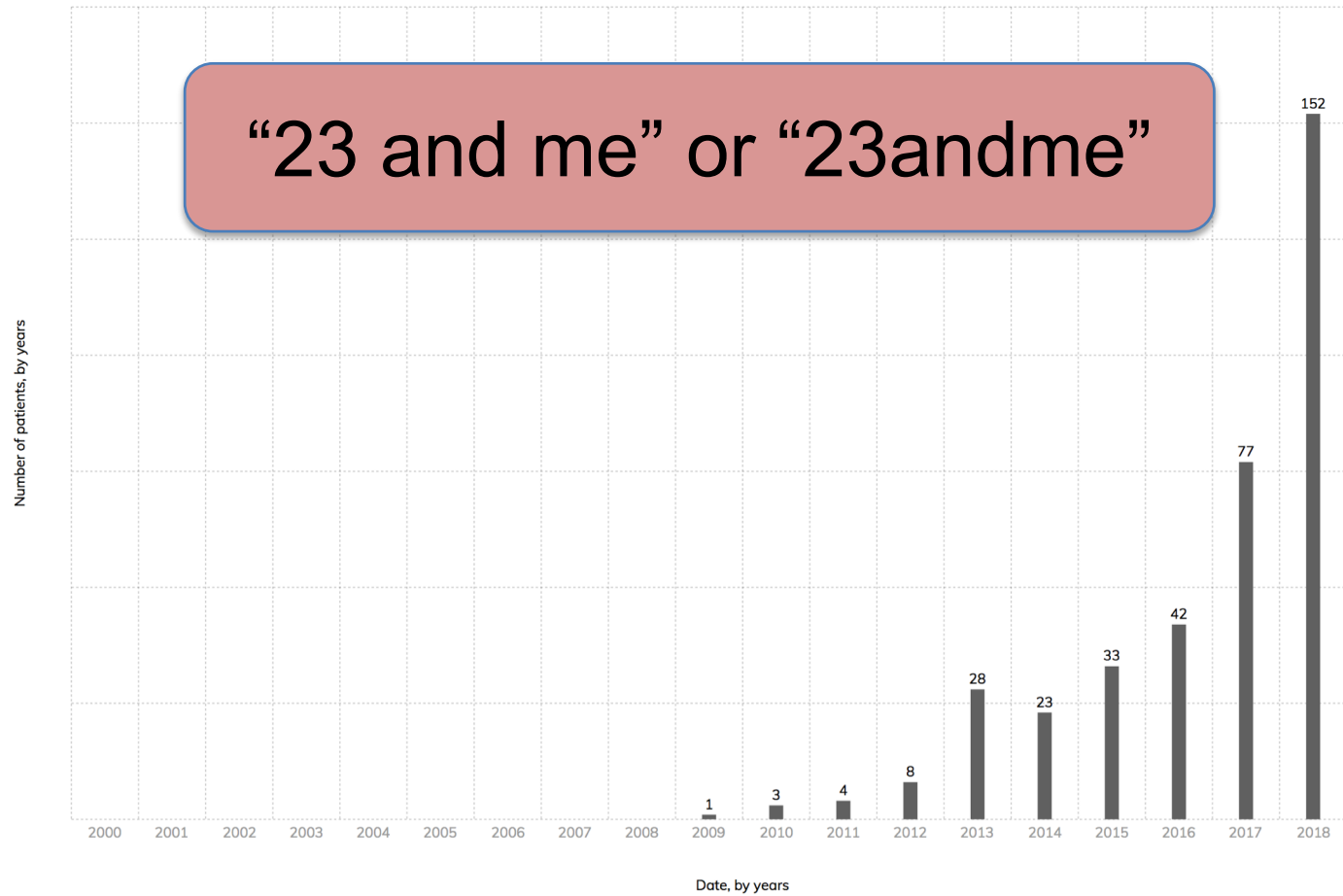
## Patient Trends



Terms

"23 and me" OR "23andme"

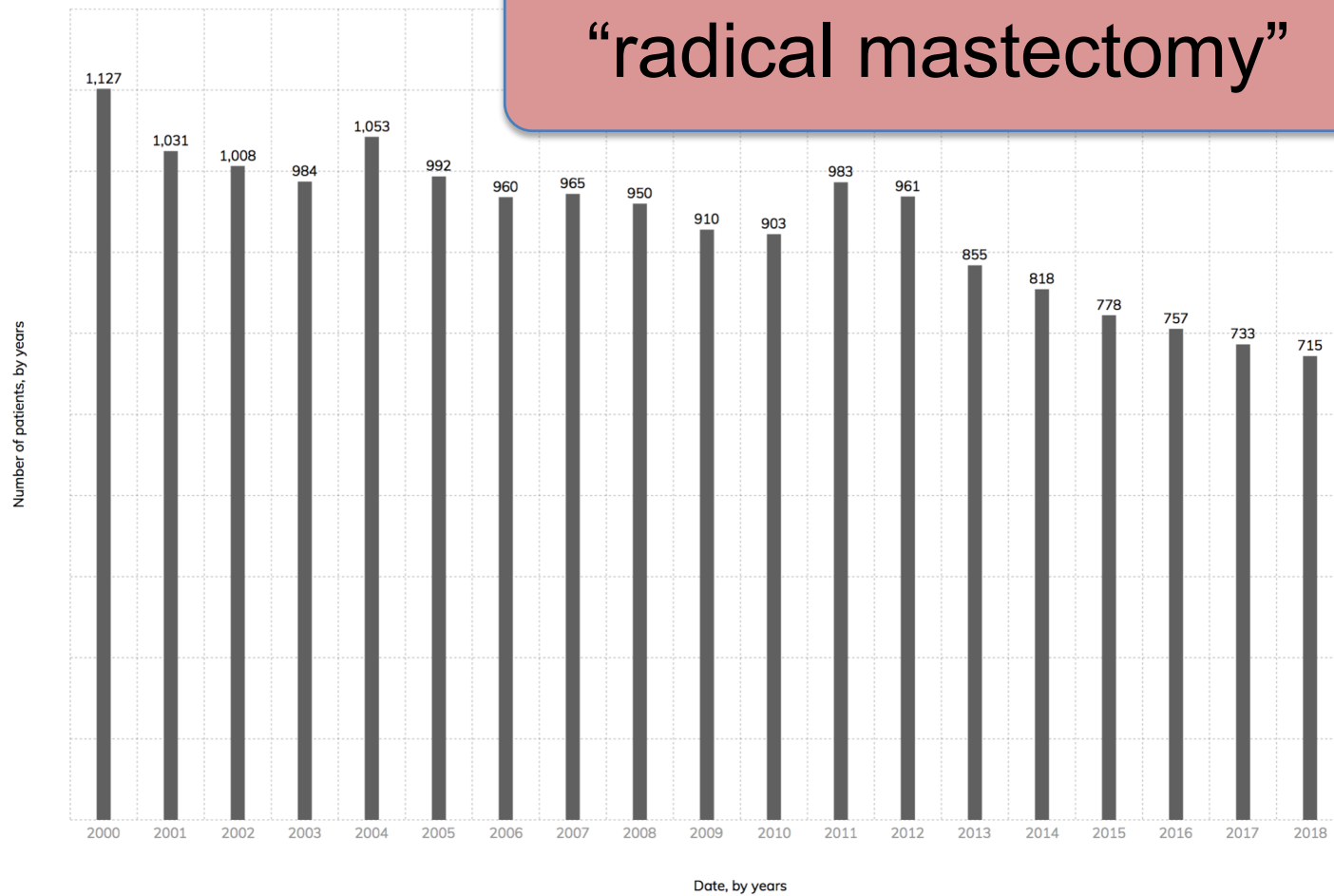
## Patient Trends



Terms

"Radical mastectomy"

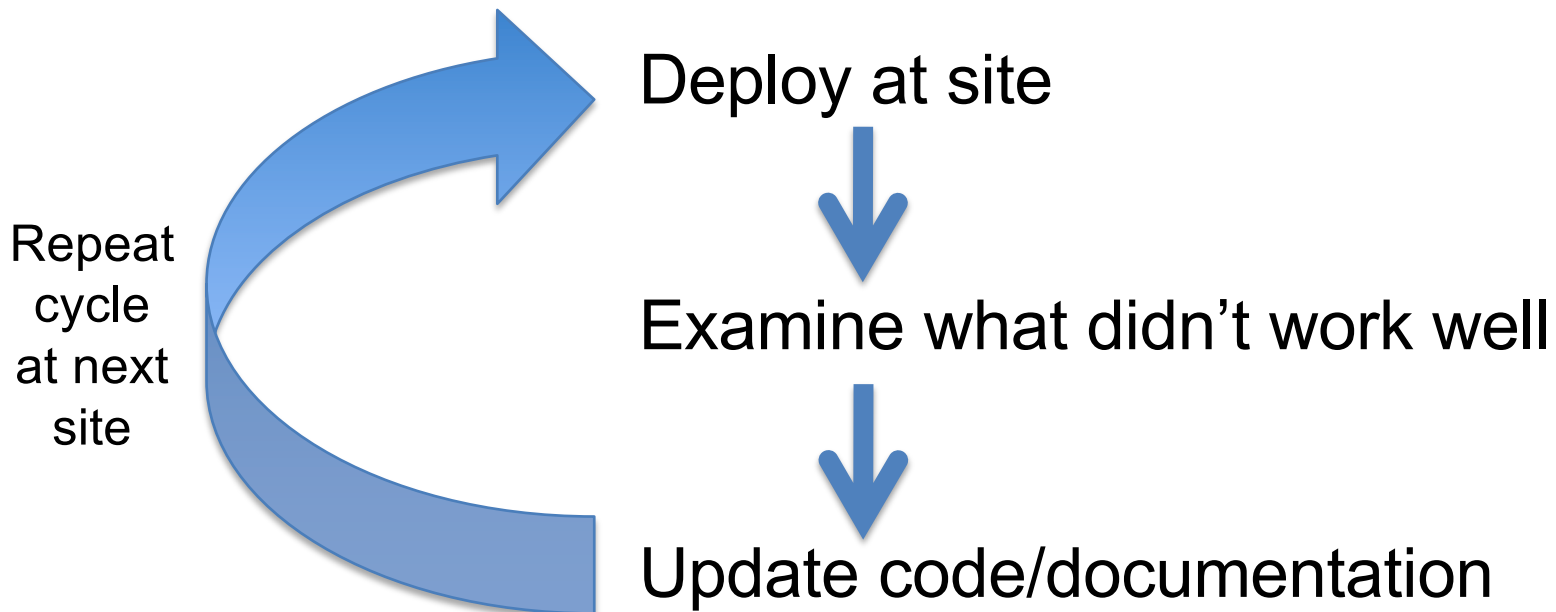
## Patient Trends



# PROJECT PLANS

- Securely network sites for cross-institution searches
- Continue work on detailed documentation
- Improve handling of synonyms
- Incorporate de-identified documents
- Feature and usability enhancements
- User evaluation
- Dissemination

# PROJECT PLANS



# LEARNING ABOUT VARIABILITY

- Variable database needs
- Variable local support structures
  - Self-service vs pay-per-use
- Variable levels of experience with the technologies
- Variable understandings of local infrastructures
- Variable security requirements



# COLLABORATION WITH UTAH

- Understand how sites obtain documents to index
- Develop a reproducible process for extracting documents in bulk from an EHR
- Explore the current use of/potential for formatted notes
  - Formatting can contain useful information

Formatted

HPI

██████████ is a 42 y.o. male who is here today for a comprehensive visual examination and to update his glasses. He has no visual complaints.

Assessment/Plan

Refractive error.

Ocular health is clear OU.

Eyeglass Final Rx

Eyeglass Final Rx

	Sphere	Cylinder	Axis
Right	-1.50	+1.00	165
Left	-2.50	+0.50	28

The patient is to return in 1-2 years.

Unformatted

HPI

XXXXXXXXXXXXXXXXX is a 42 y.o. male who is here today for a comprehensive visual examination and to update his glasses. He has no visual complaints.

Assessment/Plan

Refractive error.

Ocular health is clear OU.

Eyeglass Final Rx

|

Eyeglass Final Rx

Sphere

Cylinder

Axis

Right

-1.50

+1.00

165

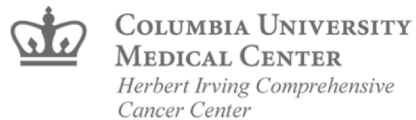
Left

-2.50

+0.50

28

# COLLABORATORS



# ACKNOWLEDGEMENTS

The logo for the ITCR (Institutional Training and Career Resource) is displayed within a light gray rectangular box. It consists of the letters "ITCR" in a bold, black, sans-serif font.

**ITCR**

The logo for the Rogel Cancer Center at Michigan Medicine is shown within a dark blue rectangular box. It features a large, bold, yellow "M" at the top. Below the "M", the words "ROGEL CANCER CENTER" are written in a white, sans-serif font, and "MICHIGAN MEDICINE" is written in a smaller, white, sans-serif font below that.

**M**  
**ROGEL CANCER CENTER**  
MICHIGAN MEDICINE

The logo for MICHHR (Michigan Health Information and Research) is displayed within a dark blue rectangular box. It features a large, bold, yellow "M" on the left, followed by a vertical line and the word "MICHHR" in a white, sans-serif font.

**M** | **MICHHR**

The logo for Health Information Technology & Services is shown within a dark blue rectangular box. It features a small, bold, yellow "M" on the left, with the words "MICHIGAN MEDICINE" in a small, white, sans-serif font below it. To the right of the "M" and "MICHIGAN MEDICINE" text, the words "Health Information Technology & Services" are written in a white, sans-serif font.

**M**  
MICHIGAN  
MEDICINE  
Health Information  
Technology & Services

ITCR (NCI): U24CA204863  
U of Michigan Rogel Cancer Center (NCI): P30CA046592  
MICHHR (CTSA): UL1TR000433 and UL1TR002240

# WEBSITE

<http://project-emerse.org>

