EMERSE Networking

Connecting for Prep to Research 9/21/21

Welcome

Session will be recorded

Feel free to use the zoom CHAT feature to submit questions

Program overview:

EMERSE Background, Networking Functional Overview & Demo - David Hanauer, Director of MICHR Informatics Program, Associate Professor - 25 min.

Q&A - 5 min.

EMERSE Network Terms of Use Document - Rebecca Hulea, Compliance Director, Michigan Medicine Compliance - 15 min.

Q&A - 15 min.

EMERSE Background, Networking Functional Overview & Demo

EMERSE – system background

- Search engine for clinical documents in the electronic health record (EHR)
- Originally built for local use only to query for details in clinical notes (diagnoses, symptoms, medications, complications, and much more)
- These notes contain protected health information
- Access to the data is limited based on a variety of factors (e.g., IRB approval for research).
- Audit logs are maintained
- EMERSE is being deployed at multiple institutions nationwide
 - Each site manages their own: system, data, users, security, access privileges, etc.

standard view

(non-network view)

• • •	EMERSE
https://dev-emerse.me	ed.umich.edu/emerse/results/summaries/viewdocument?qk=952&pl=3002&mode=mosaic&source=source4&patientId=100006810&docSortC 🗢 🤤
Ə EMERSE	David Hanauer ^
Patients Demo List	(16)
Filters	
Terms treatmen	t lymphocytes bepatitis cancer "bowel obstruction" lasix seizures "mitral valve prolapse" reflux shingles
Results HIGHLIGH	IT DOCUMENTS FIND PATIENTS SEARCH NETWORK
Overview Summaries	Document 2 of 1 of 16 5 of 7 2 Patient > 2 Document >
Main EHR	sheets, or individual undifferentiated or poorly differentiated malignant epithelial cells surrounded and infiltrated by prominent components of small, mature imphoepited and plasma cells [3]. The more frequent sites of these tumors are the nasopharynx, salivary glands and the larynx [4]. Very few cases of lymphoepithelial carcinoma arising in the maxillary sinus have been reported [5, 6] with none precisely detailing its diagnostic characteristics and management. In this report we present a case of maxillary sinus imphoepithelial carcinoma and review the radiological findings, macroscopic and microscopic features, immunohistology and the treatment options available. Case presentation A 72-year-old Caucasian woman presented to our facility with an isolated right-side epistaxis that had been present for three months without other associated totalaryngological, optithalmological or neurological symptoms. There was no history of traum or chronic sinonasal infections. Ten years prior to the incident, our patient was diagnosed as having breast gancer and was treated with surgery and chemoradiotherapy. The results of a physical examination were normal except for a thread of blood coming out of the right middle meatus. A neck examination revealed no palpable lymph nodes. Routine laboratory investigation results were normal. A paranasal sinus (PNS) computed tomography (CT) scan showed expansion of the right maxillary sinus by a homogenous soft tissue mass orcupying the entire sinus cavity. The sinus walls were intact with no bony erosion. There was an impacted tooth visible inside the mass (Figure 1). Other sinus cavitles were normal. Endoscopic maxillary sinus ganakilectomy, with complete excision of the tumor. Figure 1 Axial computed tomography (CT) scan showed expansion of the right maxillary sinus by a homogenous soft tissue mass occupying the entire sinus cavity. The sinus wallow and palate saving maxillectomy, with complete excision of the tumor. Figure 1 Axial computed tomoligical examination (Figu

EMERSE – network background

- Is there any way to link these sites together in a secure manner to allow limited querying at other sites?
 - Nothing like this has been done for free text
 - Could be useful for developing a study, estimating cohorts (preparatory to research), finding collaborating institutions
- For structured data, there is a precedent for this type of network: i2b2 SHRINE/ACT
 - ACT-Accrual to Clinical Trials-allows users to use structured data (e.g., diagnosis codes) to query other sites for obfuscated counts
 - SHRINE-Shared Health Research Information Network
- EMERSE network is inspired by, but independent from, i2b2/ACT

i2b2 SHRINE/ACT

150 MILLION PATIENTS 50 SITES CONNECTED AND GROWING.



View Patient Count by Sites

i2b2 SHRINE/ACT

Name:

Topic:

Criteria:

Status:

cardiomyopathy

Find patients with (Diagnoses 425 Cardiomyopathy)

In Progress - 26 sites with patients, up to 340,320 total patients ()

Test topic #01

Last Updated: 9/20/21 at 22:57:39 EDT	
Site 1 ^A	Status
	Status
Northwestern	Sent to Site
OHSU - Oregon	Delayed At Site
Ohio State U	Processing at Site
Penn State	9,010 ± 10 patients
Stanford	10 patients or fewer
TuftsMC	9,270 ± 10 patients
U Alabama Birmingham	Delayed At Site
U Arkansas	2,570 ± 10 patients
U Florida	8,575 ± 10 patients
U Illinois-Chicago	Site Error click for details
U Kentucky	7,230 ± 10 patients
U Michigan	10,570 ± 10 patients
U Minnesota	15,135 ± 10 patients
U Pittsburgh	50,575 ± 10 patients

•••	EMERSE	× +				(0
$\leftarrow \ \ \rightarrow \ \ G$		//dev-emerse.med.u	mich.edu/emerse/patient-lists/all-networks?qk=1&pla	=XXX:emerse-qc	• Q \$	🚳 🛃 🗯 J	:
	E					emerse default user	^
Patients	EMERSE 2 (1000	00), Your Site (10000)					
Filters							
Terms	cardiomyopath	עו					
Results	HIGHLIGHT DOG	CUMENTS FIND PA	TIENTS SEARCH NETWORK				
Temporary Pa	atient List Save	d Patient Lists All	Local Patients Network				
	Name	Status	Description	Start Date	End Date	Patient Count	
	EMERSE 2	Available	All patients at the remote site EMERSE 2	01/23/2010	02/19/2020	10000	
	Case Hospital	Offline	All patients at the remote site Case Hospital	01/22/2010	02/18/2020	10000	
	Your Site	Available	All patients at the local site Your Site	01/23/2010	02/19/2020	10000	

● ● ● ⊋ EMERSE	× +
\leftarrow \rightarrow C \triangle $$ https://dev-emerse	e.med.umich.edu/emerse/results/network?qk=1&pl=XXX:emerse-qc-client2,:all 🕶 🔍 🏠 🚳 🤼 🗯 🤳 🗄
Terms cardiomyopathy	
Results HIGHLIGHT DOCUMENTS	Network Search Terms of Use
▶ Network search can not proceed if the An	To use the EMERSE Network you must agree to: 1. Conduct searches only for valid purposes which at this time includes estimating cohort sizes and demonstrating the network capabilities to others 2. Never attempt to identify individuals 3. Never use the counts to gain a "market" advantage over other Participating sites 4. Never share your login credentials with others If you violate these terms you could be prohibited from using the network or face penalties in accordance with the type of violation. DISAGREE AGREE

● ● P EMERSE × +							0
← → C ☆ 🏻 https://dev-emerse.med.umich.edu/emerse/results/network?qk=1&pl=XXX:emerse-qc-client2,:allpt	Q	☆	3	/	*	J	:
<mark>२</mark> EMERSE			em	erse	defaul	t user	
Patients EMERSE 2 (10000), Your Site (10000)							
Filters							
Terms Cardiomyopathy							
Results HIGHLIGHT DOCUMENTS FIND PATIENTS SEARCH NETWORK							
Site Patient Count							
EMERSE 2							
Your Site 1497							

	emerse x +							V
$\leftarrow \ \rightarrow \ G$	☆ https://dev-emerse.med.umich.edu/emerse/results/network?qk=1&pl=XXX:emerse-qc-client2,:allpt	Q	☆	9	/	*	J	•
				em	erse	defau	lt user	^
Patients	EMERSE 2 (10000), Your Site (10000)							
Filters								
Terms	cardiomyopathy							
Results	HIGHLIGHT DOCUMENTS FIND PATIENTS SEARCH NETWORK							
	Site Patient Count							
	EMERSE 2 1495							
	Your Site 1497							

EMERSE – network: practical questions

• What can a user do with the counts?

- Hope is that it can encourage research collaborations among sites
- This still needs to be explored
- Similar issues to use of i2b2/ACT

• Why should a site bother?

- No additional software needs to be installed or maintained
- Everything is already built into EMERSE
- The network component just needs to be turned on and configured
- Sites have to feel "comfortable" with this setup, but we have made great efforts to enable this in a secure and trustworthy manner

EMERSE – network: two phases

• Phase 1: Obfuscated counts

- Technical work is complete
- Goal now is to get sites to agree to participate

• Phase 2: Shared Search Term "Bundles"

- Sharing of saved search terms across sites
- Users can already share Bundles internally to all users
- This can promote re-use of search knowledge
- Have already had a request for this feature from a multi-site study involving EMERSE sites
- Technical work still required, but will leverage existing network components

EMERSE – network architecture



EMERSE – network security

Built with security in mind

- Developed with input from the Michigan Medicine Information Assurance (IA; aka IT security) team
- Has undergone external (from the EMERSE team) code reviews, penetration testing, vulnerability scans, architectural reviews
- Has received the approval of IA
- Requires ongoing reviews

No "data" are transmitted

- No data/text/no protected health information ever leaves a local EMERSE instance.
- No data/text/no protected health information ever travels across the network
- Only obfuscated counts are sent out by each EMERSE instance this is not even considered to be clinical data (no data use agreement is needed)

Obfuscation

- Obfuscation is determined locally by each site
- Each site has control over the approach for obfuscation

Distribution value	Description	Additional Parameters	Parameter description
normal	The normal distribution	XXX.obfuscate.count.distribution.normal.s=2	The standard deviation
binomial	The binomial distribution	XXX.obfuscate.count.distribution.binomial.n=6 XXX.obfuscate.count.distribution.binomial.p=0.5	Number of "trials", a choice between 0 or 1. Probability of a "1"
uniform	The uniform distribution	XXX.obfuscate.count.distribution.uniform.scale=6	This is the uniform distribution on [0, S], where S is this parameter.
disabled	A uniform distribution on [0,0]	(none)	This distribution effectively disables the random variation of results.

Thresholds

- Thresholds are controlled by each site
- Thresholds are important to prevent a user from narrowing a query down to extremely small samples sizes, which could potentially identify and individual.

🗦 EMERSE		emerse default user ^
Patients	EMERSE 2 (10000), Your Site (10000)	
Filters		
Terms	nephrectomy myocarditis	
Results	HIGHLIGHT DOCUMENTS FIND PATIENTS SEARCH NETWORK	
	Site	Patient Count
	EMERSE 2	<15
	Your Site	<20

Encryption

• Connections between sites are encrypted

Authentication

- Each site authenticates to the broker by a username and password, over a secure TLS connection
- This helps ensure attackers don't try to directly connect to the message broker, or that one site doesn't attempt to impersonate another

Query limits

- Sites can set a limit for how many queries a user can submit in a given time
- This can help prevent a user from running the same query over-and-over to try to "guess" the real count

Randomized response times

- In general, smaller counts can be returned faster, especially 0 counts.
- Random amounts of delays are added to each response time
- This helps prevent a malicious person from trying to estimate the count based on the response time

Detailed audit logs

• All sites maintain their own audit logs and track details about both incoming and outgoing queries

emerse default user

EMERSE

Users Roles Synonyms Lo	og System Ale	erts Reports Netw	orks			
Networks Network Info	Date/Time	Site/Node	User ID	Туре	Result	Search Terms
Network Log	9/14/21, 9:49 AM	XXX:emerse-qc-client1	emerse	External - Incoming	485	acetaminophen
Firewall	9/14/21, 9:48 AM	XXX:emerse-qc-client1	emerse	External - Incoming	<15	"jane doe"
	9/14/21, 9:47 AM	XXX:emerse-qc-client1	emerse	External - Incoming	<15	nephrectomy,myocarditis
	9/14/21, 9:45 AM	XXX:emerse-qc-client1	emerse	External - Incoming	357	myocarditis
	9/14/21, 9:41 AM	XXX:emerse-qc-client1	emerse	External - Incoming	2163	ab
	8/16/21, 11:52 AM	XXX:emerse-qc-client1	emerse	External - Incoming	773	abc
	8/16/21, 11:47 AM	XXX:emerse-qc-client1	emerse	External - Incoming	770	abc
	8/16/21, 11:45 AM	XXX:emerse-qc-client1	emerse	External - Incoming	773	abc
	8/15/21, 10:04 PM	XXX:emerse-qc-client1	emerse	External - Incoming	2160	ab
	7/16/21, 11:52 AM	XXX:emerse-qc-client1	emerse	External - Incoming	207	apples
	7/16/21, 11:50 AM	XXX:emerse-qc-client1	emerse	External - Incoming	9796	test
	7/16/21, 11:04 AM	XXX:emerse-qc-client1	emerse	External - Incoming	9798	test
	7/16/21, 10:44 AM	XXX:emerse-qc-client1	emerse	External - Incoming	9796	test
	7/15/21, 4:13 PM	XXX:emerse-qc-client1	emerse	External - Incoming	9794	test
	7/15/21, 10:27 AM	XXX:emerse-qc-client1	emerse	Local - Outgoing	9795	test
	7/13/21, 3:18 PM	XXX:emerse-qc-client1	emerse	External - Incoming	771	abc

Built-in firewall

- Each site has its own firewall which allows them to:
 - Block a specific site on the network (all users from that site)
 - Block a specific user from a specific site on the network (including local/internal users as well as external users)

्रि	EMERSE								emerse default user 🕠
User	s Roles	Synonyms	Log System	Alerts Reports	letworks				
Ne	tworks Jetwork Info					INPUT			
I	letwork Log		Network	Node	User	Action	Comment		
,	irewall		ALL	ALL	ALL	ACCEPT	Default INPUT Rule	Ø	
			Local Network	EMERSE	emerse	REJECT		Ø	\otimes
									ADD NEW RULE
						OUTPUT			
			Network	Node	User	Action	Comment		
			ALL	ALL	ALL	ACCEPT	Default OUTPUT Rule	Ø	
									ADD NEW RULE
			UPDATE FIREW	ALL					

Outbound connections only

- All sites only need to make an outbound connection to the query coordinator
- No inbound connections are needed
- No network security firewall exceptions need to be made
- The EMERSE server can be maintained behind the secure firewalls of each institution
 - The server simply has to be able to make an outbound connection, but most server configurations should be capable of doing that while keeping the system secure

Terms of Use

- For every EMERSE session, the network Terms of use will be shown to the user and they will have to agree to the Terms
- The Terms are brief and written in easily understandable language

● ● ● ⊋ EMERSE	× +
\leftarrow \rightarrow C \triangle $$ https://dev-emerse	e.med.umich.edu/emerse/results/network?qk=1&pl=XXX:emerse-qc-client2,:all 🕶 🔍 🏠 🚳 🤼 🗯 🤳 🗄
Terms cardiomyopathy	
Results HIGHLIGHT DOCUMENTS	Network Search Terms of Use
▶ Network search can not proceed if the An	To use the EMERSE Network you must agree to: 1. Conduct searches only for valid purposes which at this time includes estimating cohort sizes and demonstrating the network capabilities to others 2. Never attempt to identify individuals 3. Never use the counts to gain a "market" advantage over other Participating sites 4. Never share your login credentials with others If you violate these terms you could be prohibited from using the network or face penalties in accordance with the type of violation. DISAGREE AGREE

Network Terms of Use Document

Approach to Understanding – Terms of Use

- Network Agreement describes terms of use for network participants
- Network sites are responsible for its users and setting up security / privacy features
- Network sites may withdraw participation at anytime
- Network Terms of Use requires acknowledgement of understanding (institutional signature on the document)

EMERSE Research Informatics Network Terms of Use

Background

EMERSE is the Electronic Medical Record Search Engine, a software tool for searching free text clinical documents derived from electronic health record (EHR) systems. EMERSE is intended to be installed, operated and queried locally, but has an optional networking component that provides query capability which allows for the transfer of non-protected health information in the form of obfuscated counts between EMERSE systems installed at different Participating Sites. Technical documentation detailing the network and its configuration is available on the main EMERSE website (project-emers.erg).

Definitions

Participating Site. A Participating Site is an institution/organization, such as an academic medical center, that has signed the Terms of Use document and agrees to adhere to all of the guidelines outlined within that document. Each Participating Site will have oversight of its Users. Participating Sites are able to withdraw from the network at any time but will then no longer be able to connect to the network and must inform the other sites of their inkention to withdraw.

User: A User is an individual that has an EMERSE account at one of the Participating Sites.

Network: The group of Participating Sites that have agreed to the principles outlined in this document and have signed the EMERSE Research Informatics Network (ERIN) Terms of Use document.

Security Features

EMERSE has multiple security features built into the system that are configurable by each site. These include the ability to:

- 1. Set the threshold under which counts are not returned; the recommended setting is less than 10
- Apply an obfuscation algorithm to the results which will provide a range of results when a query is run rather than a precise count

Data Privacy Features

- Obfuscated Counts only no PHI
- Limits the number of queries:
 - By site
 - By individuals
- Monitor / track search logs for users across all sites
- Randomizes returned search time
- Block users and sites
- Users agree to Privacy Terms of use

•	Network Search Terms of Use To use the EMERSE Network you must agree to:
	To use the EMERSE Network you must agree to:
	Conduct searches only for valid purposes which at this time includes estimating cohort sizes and demonstrating the network capabilities to others Never attempt to identify individuals Never use the counts to gain a "market" advantage over other Participating sizes
	Never share your login credentials with others
	If you violate these terms you could be prohibited from using the network or face penalties in accordance with the type of violation.
	DISAGREE AGREE
	Participation Rules
	All Participating Sites must agree to:
	1. Allow the other Participating Sites in the Network to run queries against their
	Iocal EMERSE instance to retrieve aggregate, obfuscated patient counts. 2 Provide education to their own Users about the properties of EMERSE
	 Keep their own local data secure based on their site's local regulatory, privacy
	and compliance requirements.
	 Provide oversight of their Users, including: Only apprinting accounts to Users who have undergoes the appropriate
	 Only provisioning accounts to Users who have undergone the appropriate

3. Limit the number of times a guery can be run by a site

4. Capture the search logs for each user across all participating sites

Participation Rules

- Allow Network sites to run queries
- Educate Users
- Provide oversight of users
- Monitor access / terminate when no longer needed by a user
- Investigate concerns of other Network sites when raised

3. Limit the number of times a query can be run by a site 4. Capture the search logs for each user across all participating sites 5. Randomizing the time it takes for a query to be returned 6. Block a participating site or a specific user from a participating site from querying the local data for counts. In addition to this Terms of Use Document, each user who utilizes the network will be required to agree to their own Terms of Use each time they use the network. This is built into the EMERSE system and is as follows: Network Search Terms of Use To use the EMERSE Network you must agree to: 1. Conduct searches only for valid purposes which at this time includes estimating cohort sizes and demonstrating the network capabilities to others 2. Never attempt to identify individuals 3. Never use the counts to gain a "market" advantage over other Participating sites 4. Never share your login credentials with others If you violate these terms you could be prohibited from using the network or face penalties in accordance with the type of violation. DISAGREE AGREE Participation Rules All Participating Sites must agree to: 1. Allow the other Participating Sites in the Network to run queries against their local EMERSE instance to retrieve aggregate, obfuscated patient counts. 2. Provide education to their own Users about the proper use of EMERSE. 3. Keep their own local data secure based on their site's local regulatory, privacy and compliance requirements. 4. Provide oversight of their Users, including: a. Only provisioning accounts to Users who have undergone the appropriate training for using EMERSE, as determined by each participating site.

Finalizing the Network Terms of Use

Ready to implement networking phase	Acknowledge use terms are in place and will abide by use terms by Signing Agreement
Not ready to implement networking phase	Don't sign until all use terms are in place and agree to be bound by use terms.

Thank you for attending!

• Hold the date

- o **10/19/21**
- 12-1:30 p.m. ET
- Making Sense of Unstructured Data