

# Enhancing Research with EMERSE Text Processing and Search tool for an AI-Driven Future

University of Michigan Medical-Dental LHS Collaborative

May 13, 2025



Twitter/X: @projectEMERSE  
Web: [project-emerse.org](http://project-emerse.org)  
Email: [hanauer@umich.edu](mailto:hanauer@umich.edu)  
[emerse-team@umich.edu](mailto:emerse-team@umich.edu)

David Hanauer, MD, MS  
Dept of Learning Health Sciences  
University of Michigan

<http://project-emerse.org/presentations.html>

# If you want to take notes or visit links



these slides can be found at:

this link will be  
on most slides

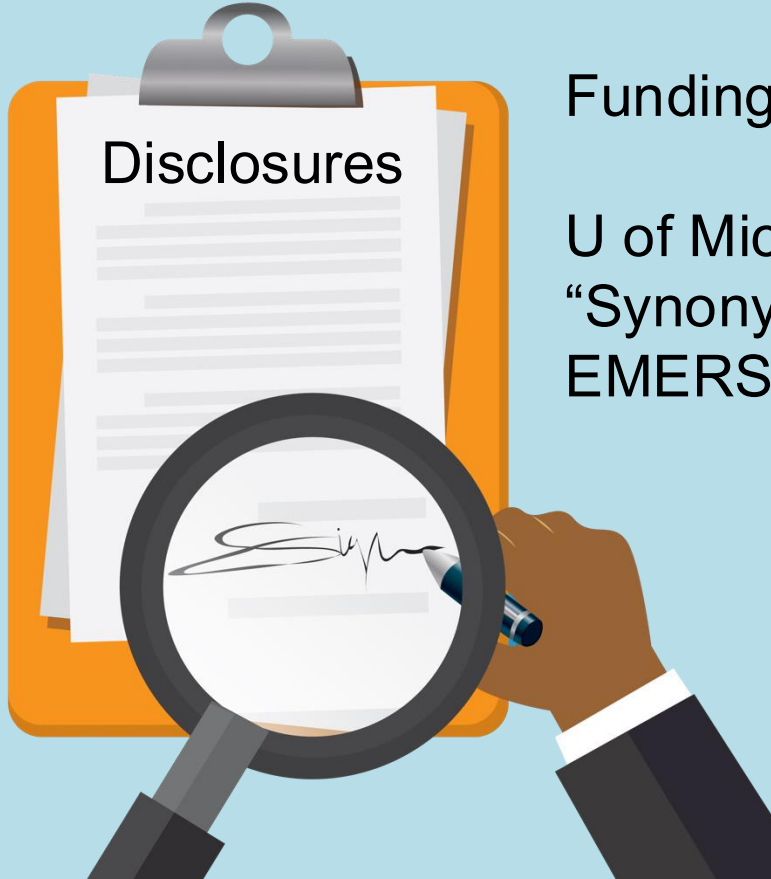


<http://project-emerse.org/presentations.html>

# Disclosures

Funding: NIH (NCI-ITCR, NCATS-CTSA)

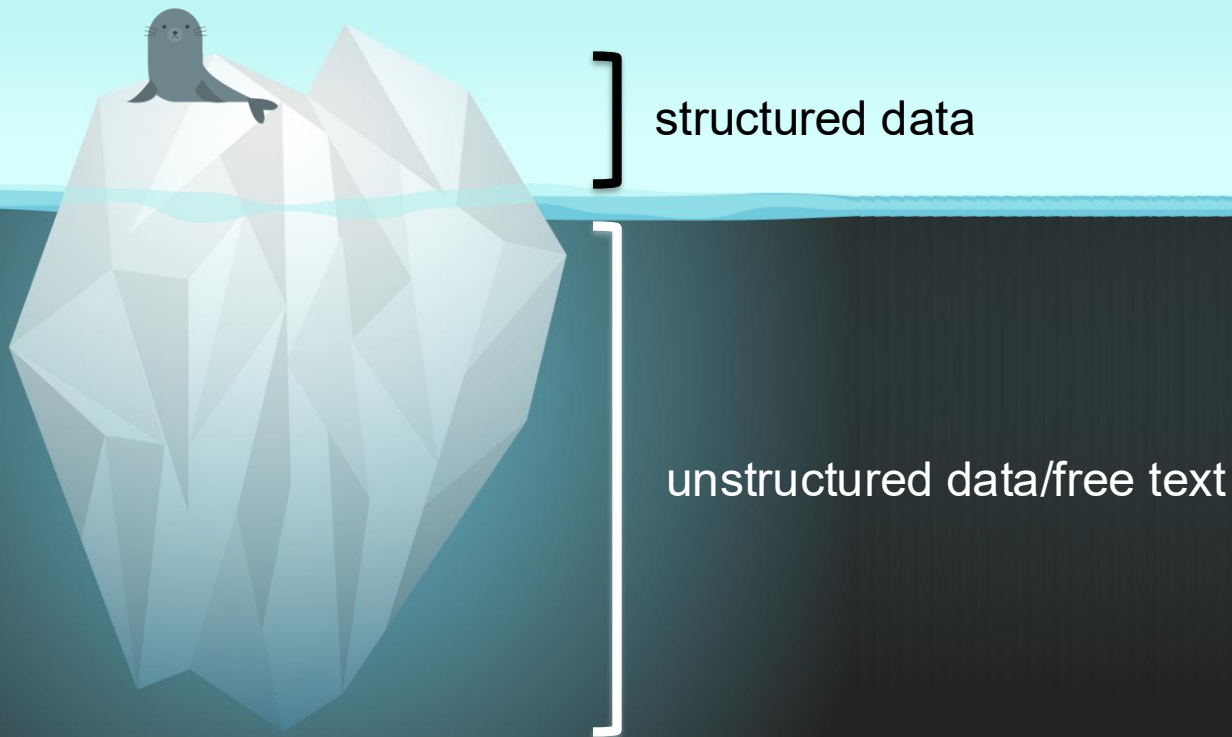
U of Michigan Royalties/Licensing:  
“Synonyms” dataset—optional “plugin” for  
EMERSE.



# Unstructured vs Structured Data

EMERSE is for this...	...not this
<i>Unstructured Data (free-text)</i>	<i>Structured Data</i>
Mrs. Jones is a 56 year old female with a history of HTN, hypercholesterolemia, and T2DM who comes to the clinic today with a 3 day h/o dizziness and severe headache on the left side.	WBC: 5.6 Total cholesterol: 182 Weight: 67.4 AST: 30 ALT: 52

# 80% of EHR data are in unstructured free-text



# Most medical centers lack tools for free-text



# The EMERSE solution



- A system “for the people”
- Users search the notes on their own
  - No need to wait in a queue for an analyst or a data scientist
- Easy-to-use for non-technical users
- Unlike with some EHRs, EMERSE can search across all notes and all patients at once
- Continuous refinements for over 19 years

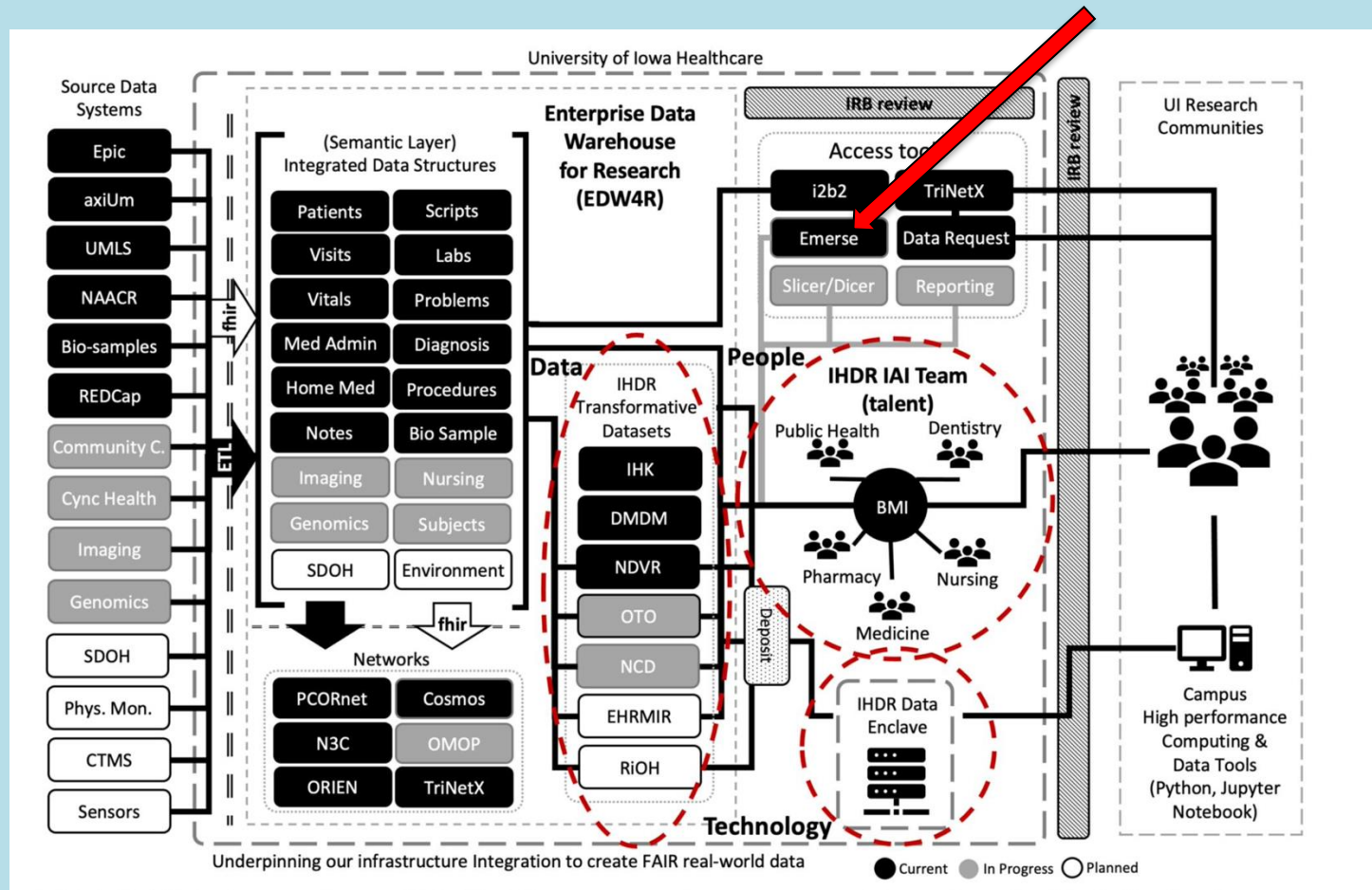


# EMERSE is a useful tool

- It will be one of many you need to build/support your project(s)







# How much does it cost?

Software	\$0.00
Installation Guidance	\$0.00
Access to Documentation	\$0.00
Training (within reason)	\$0.00
Synonyms dataset (optional)	\$0.00*



\* For research use within EMERSE

# We think EMERSE is really good

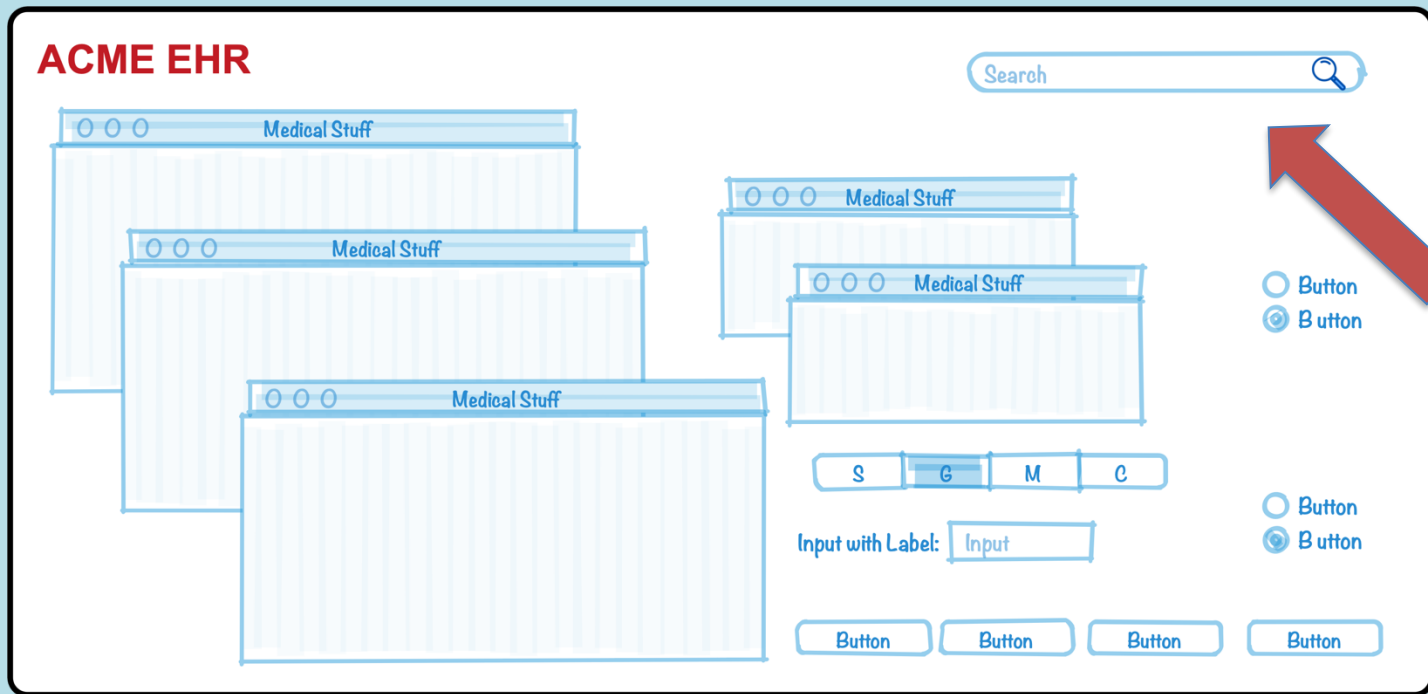
- We're proud of what we created
- We think it's better than other software tools that cost \$\$\$



# Why do we think it's better?

## *We've been laser-focused on it*

*EHRs have thousands of features*



*The EMERSE team has dedicated 19+ years on just this one feature (search) because it's so important*

# Why do we think it's better?

## *We've studied it*

- “new users of the EMERSE system are able to complete basic but critical workflow tasks in the system with a high rate of success...are **highly satisfied with the interface**, and have highly positive perceptions of its expected utility in their work.”
- “survey results indicate **very high ratings of usability and satisfaction** with the EMERSE system by new users after only one session of use.”

# Why do we think it's better?

## *Our users tell us*

Thank you for...this important tool which is **proving extremely valuable** in enhancing patient safety and quality of care delivered at Michigan Medicine.

..continued gratitude for the availability of this powerful research tool. As always, **thank you for this innovation!**

EMERSE is **working out great for casefinding**...have found quite a bit of cases that we would have missed otherwise. **Thank you and everyone on your team!**

To say **“it is the most useful tool I use in my job”** would be an understatement, so **thank you** for creating and sharing it with others!!

EMERSE is **an absolute gem** @umichmedicine. The functionality is very friendly and it saved hours of time during the data collection process. **Thank you to the team** that created this powerful research tool!



# Why do we think it's better?

## *Researchers mention it in their publications*

- “Reviewers used the EMERSE search tool to **ensure thorough review** of the available documentation...” [PMID 36119396]
- “the **tool avoids the pitfalls of diagnostic inaccuracy** seen with tools querying on ICD and billing codes...” [PMID 36114099]
- “[EMERSE] provides software features to **comprehensively scan all clinical documents**...for keywords and phrases to ensure that even rarely mentioned events are detected.” [PMID 36550198]
- “...information [was] captured via EMERSE...in order to **obtain the most accurate and complete information** per patient.” [PMID 36752520]

# Why do we think it's better?


## *Top-tier medical centers use it*

Installed	Installing
U of Michigan	Weill Cornell Medical Center
Harvard U – Dana Farber Cancer Center	U of Virginia
Columbia U Cancer Center	Moffitt Cancer Center, Tampa, FL
U of North Carolina – Chapel Hill	U of California – Irvine
U of California – San Francisco	Children's Health Orange County, Calif.
U of Kentucky	
U of Cincinnati	
Case Western Reserve U	
U of Minnesota – Minneapolis	
Utrecht University, Netherlands	
U of Iowa	



# What can EMERSE be used for?

Lots of things!



EMERSE for Cardiac Surgery Research

David Hanauer

Linda Farhat, Clinical Research Coordinator  
Department of Cardiac Surgery, Michigan Medicine

01:28

Cardiac Surgery Research

"It allows us to basically target exactly what we need to know."

--Linda Farhat, Clinical Research Coordinator



EMERSE for Quality Analytics

David Hanauer

Andrew Heiler, Clinical Quality Coordinator

01:18

Quality Analytics

"I can't imagine having to do what I do without having EMERSE."

--Andrew Heiler, Quality Analytics Coordinator



EMERSE for Infection Control

David Hanauer

Jayna Berger Heiler, Infection Prevention Coach

01:33

Infection Prevention

"The value of EMERSE is immeasurable."

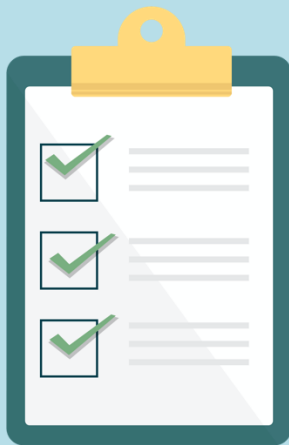
--Jayna Berger Heiler, Infection Prevention Coach

Watch our videos: [https://project-emerse.org/use\\_cases.html](https://project-emerse.org/use_cases.html)

# Find cohorts

EMERSE allows you to find cohorts based on things mentioned in the notes

- diseases
- drugs
- symptoms
- anything\*



\*if it is mentioned

# Find cohorts

It's perfect for finding rare things...



...like rare cancers such as  
cutaneous leiomyosarcoma

See this talk for more details:

<https://vimeo.com/677482835>

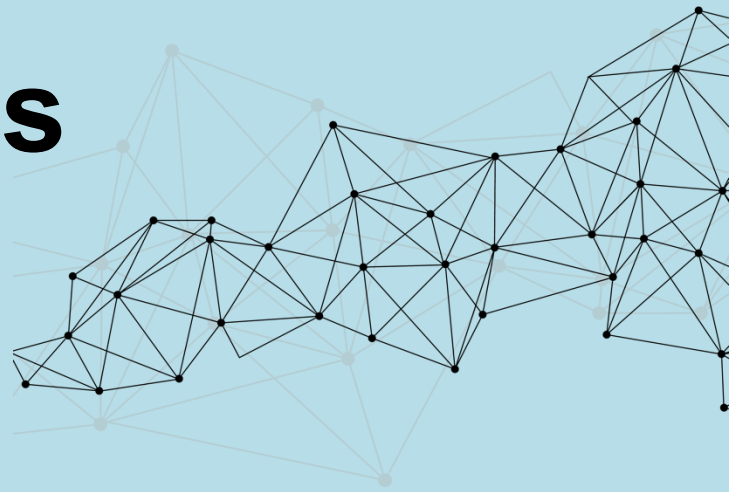
“Using EMERSE to Improve Research  
Involving Rare Cancers”



# Highlight documents for chart review

Thoracocentesis confirmed the recurrence of mantle cell lymphoma. Disease restaging work-up revealed multicompartiment lymphadenopathy in the neck, mediastinal, retrocrural, retroperitoneal and pelvic regions. Bone marrow was also involved. The patient was treated with a total of six cycles of rituximab, cyclophosphamide, vincristine, doxorubicin and dexamethasone (R-HyperCVAD) completed in January 2007. That treatment led to complete remission that lasted until October 2008, when the disease was found to have recurred in the left pleural space and retroperitoneum without bone marrow involvement.

# Synonyms



- Used for query expansions
- User-controlled
- Multiple datasets can be included
- EMERSE Synonyms
  - acronyms, abbreviations, professional/consumer terms, misspellings, trade/generic drug names, species, chemo regimens, phrase variations, malapropisms, idioms, neologisms, organizations, companies, & more
  - 2.8 million unique entries

# Synonyms for bite guard

Click individual terms to highlight or de-highlight.

WRAPAROUND VIEW



Synonyms (42)

SORT A TO Z

HIGHLIGHT ALL

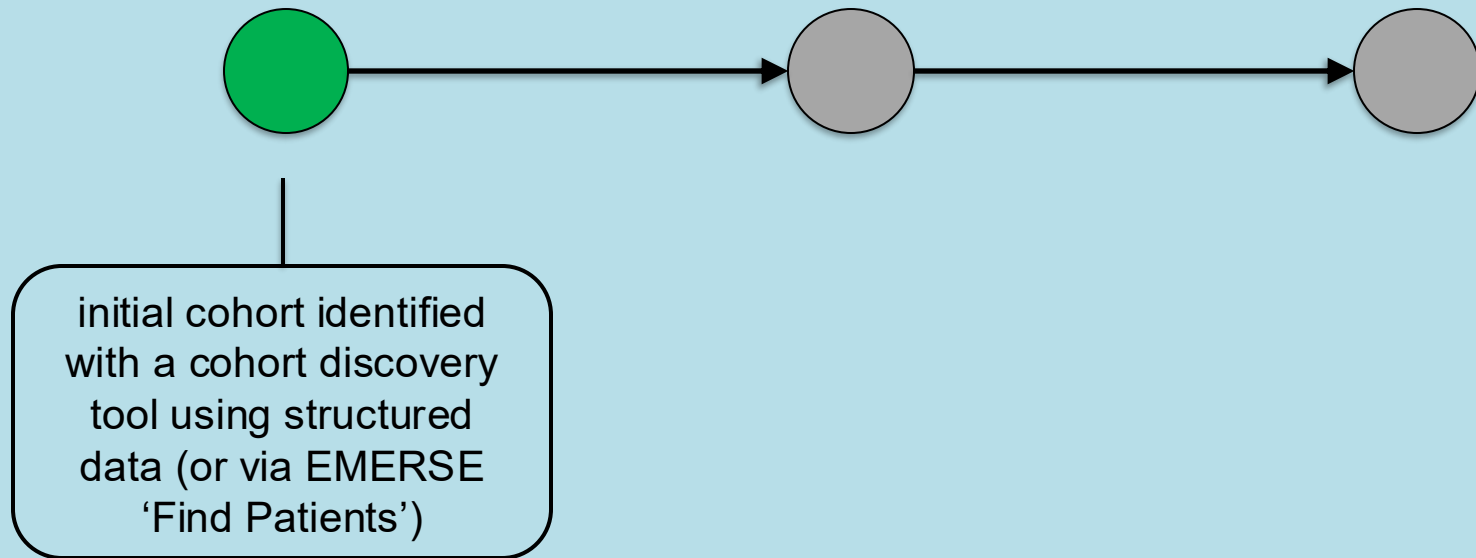
HIGHLIGHT NONE

- bite guards
- bite splint
- bite splints
- dental bite guard
- dental bite guards
- dental guard
- dental guards
- dental night guard
- dental night guard appliance
- dental night guard appliances
- dental night guards
- dental occlusal guard
- dental occlusal guards
- night guard
- night guard appliance
- night guard appliances
- night guards
- nightguard
- nightguards
- occlusal appliance
- occlusal appliance therapy
- occlusal appliances
- occlusal dental guard
- occlusal dental guards
- occlusal dental splint
- occlusal dental splints
- occlusal device
- occlusal device night guard
- occlusal device night guards
- occlusal devices
- occlusal guard
- occlusal guards
- occlusal mouth guard
- occlusal mouth guards
- occlusal mouthguard
- occlusal mouthguards
- occlusal orthotic device
- occlusal orthotic devices
- occlusal splint
- occlusal splints
- stabilization bite splint
- stabilization bite splints

ADD HIGHLIGHTED TERMS (42)

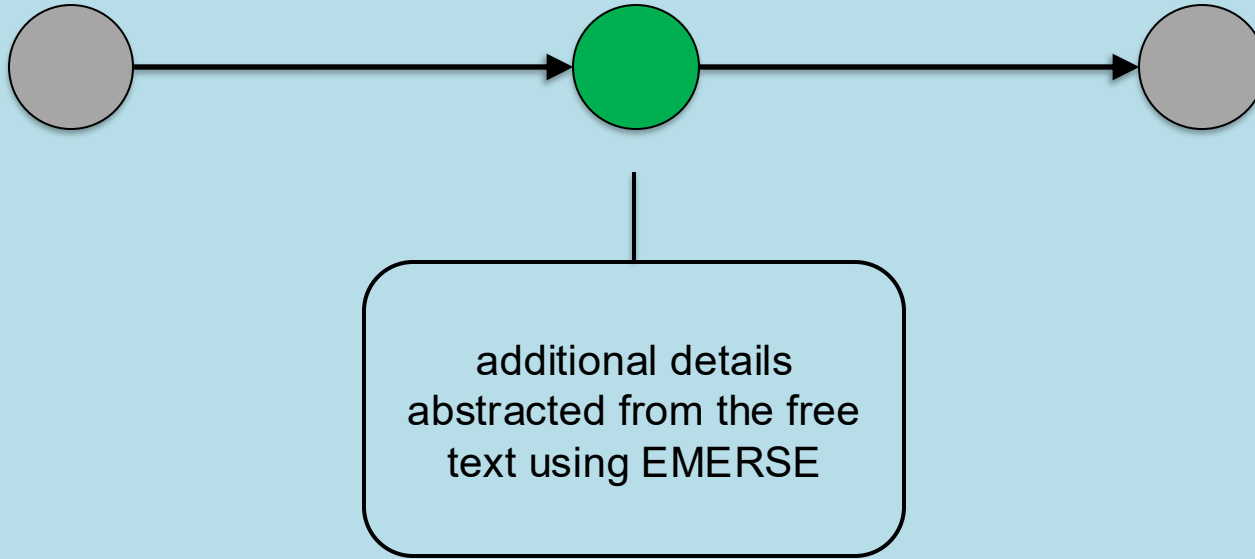
CANCEL

# Typical workflow



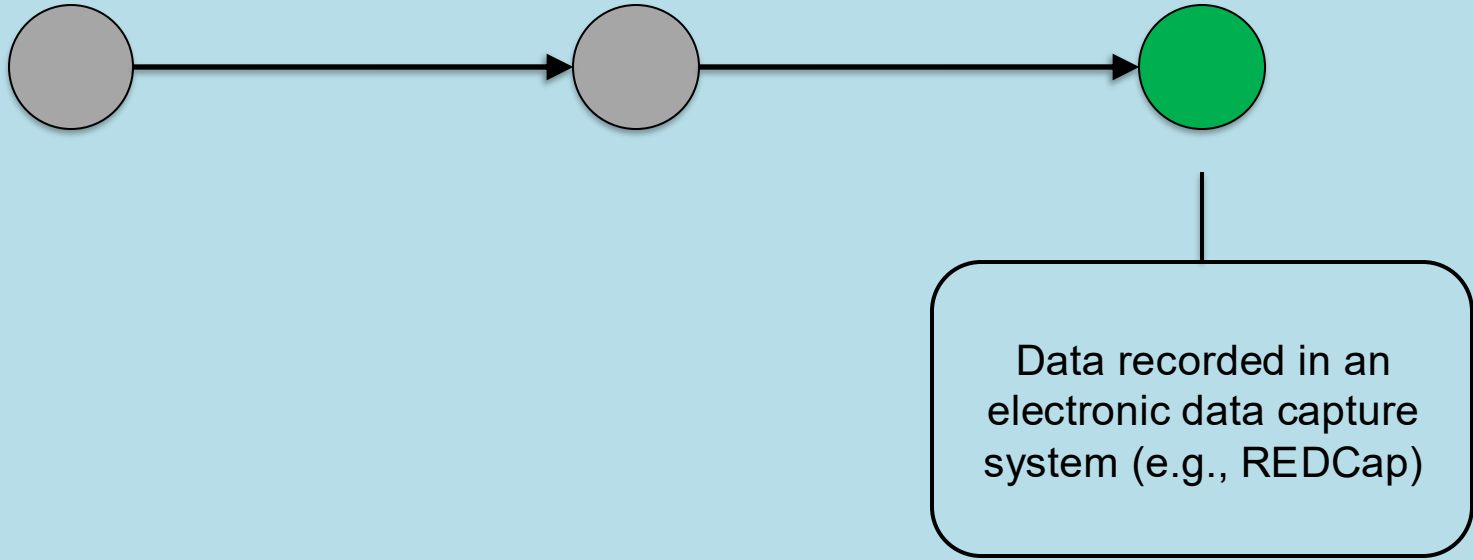
Cohort discovery tools:  
i2b2/ENACT, Leaf, etc.

# Typical workflow





# Typical workflow



# Publications using EMERSE

701

papers and abstracts



Full list at:

<http://project-emerse.org/publications.html>

# Recently released

Incorporation of natural language processing (NLP) features

- negation
- uncertainty
- subject (patient vs other)
- named entity recognition/mapping to ontologies

If you don't like our built-in NLP, add your own



Anatomy (40) Attribute (0) Device (2) Disorder (70) Drug (33) Event (10) Finding (300) Lab (8) Procedure (83) Subject (16) Title (0)

Note: Sections that overlap will be underlined/highlighted in black  
.....

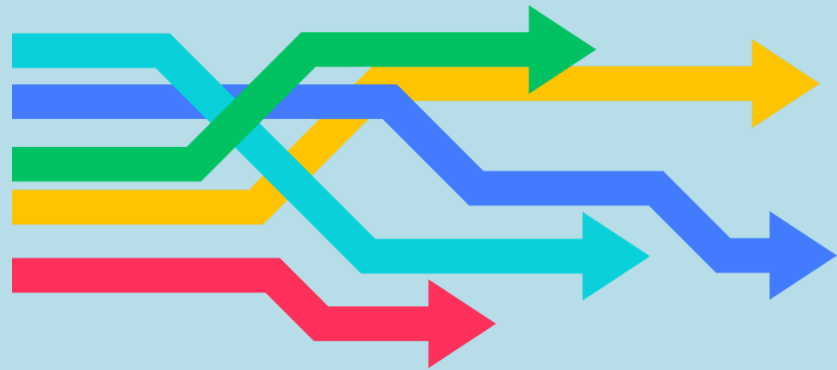
Annotations

Negation (17) Uncertainty (4) Non-patient subject (5) History of (1) Note: Sections that overlap will be underlined in black

> Summary

Case Report from the Journal of Medical Case Reports Title: Familial glucocorticoid deficiency presenting with hyperpigmentation, gigantism, and motor development delay: a case report | Journal of Medical Case Reports | Full Text DOI: 10.1186/s13256-019-2206-5 Abstract Background Familial glucocorticoid deficiency is a rare autosomal recessive disorder characterized by isolated glucocorticoid deficiency. Most patients are diagnosed following episodes of hypoglycemia or convulsion. We report the case of an infant with familial glucocorticoid deficiency who presented with hyperpigmentation, gigantism, and motor developmental delay without documented hypoglycemia, convulsion, or circulatory collapse. Case presentation A 10-month-old Sri Lankan Sinhalese baby boy born to consanguineous parents presented with generalized hyperpigmentation and overgrowth since birth. He had marginal gross motor developmental delay. His weight, length, and head circumference were above normal range for his age. Investigations revealed low serum cortisol and high adrenocorticotrophic hormone levels with no cortisol response following adrenocorticotropin stimulation. Serum electrolytes and aldosterone levels were normal. A diagnosis of familial glucocorticoid deficiency was made based on isolated glucocorticoid deficiency, hyperpigmentation, and tall stature. Conclusions This case report highlights that glucocorticoid deficiency can present without documented hypoglycemia and circulatory collapse and a high degree of suspicion is needed in diagnosis. Open Peer Review reports Background Familial glucocorticoid deficiency (FGD) is a rare autosomal recessive disorder characterized by isolated glucocorticoid deficiency with normal mineralocorticoid activity [1]. Two main types are described: type 1, which is caused by mutations in the gene encoding for adrenocorticotrophic hormone (ACTH) receptor (melanocortin-2 receptor, MC2R); and type 2, due to a mutation in melanocortin-2 receptor associated protein (MRAP) [2, 3]. Most of the reported cases of patients with FGD were diagnosed following episodes of hypoglycemia, recurrent convulsions, or circulatory collapse. Here, we report the case of an infant with FGD who presented with hyperpigmentation, gigantism, and isolated gross motor developmental delay. Case presentation A 10-month-old Sri Lankan Sinhalese baby boy presented with generalized hyperpigmentation and overgrowth since birth (Fig. 1). He was born at term following an uncomplicated antenatal period and had a normal perinatal period without episodes of hypoglycemia or circulatory collapse. He was the only child of a pair of consanguineous parents. His growth parameters at birth were within normal limits; weight 3.3 kg (at median for age), length 50 cm (at median for age), and head circumference 35 cm (between median and +1 SD). Since birth his growth chart demonstrated accelerated growth in weight, length, and head circumference (Fig. 2). His parents observed hyperpigmentation at birth which worsened gradually over time. His developmental history revealed marginal gross motor developmental delay; at 10 months of age, he was able to sit without support, however, he was unable to come to seated position on his own, stand with support, or crawl. His vision, fine motor, speech, and social development milestones were age appropriate. There were no previous hospital admissions, recurrent infections, seizures, episodes of shock, or documented hypoglycemia. Fig. 1 Photograph of the child demonstrating marked hyperpigmentation Full size image Fig. 2 Growth charts demonstrating accelerated growth during first 10 months and slowing down of growth rate after commencement of treatment Full size image On examination, his length was 84 cm (above +3 SD), weight was 12.6 kg (at +3 SD), and head circumference was 48 cm (at +2 SD). He had generalized hyperpigmentation involving oral mucosa, palms, and soles. The rest of the physical examination, including cardiovascular system, blood pressure, abdomen, and genitalia was clinically normal. Basic hematological and biochemical investigations that included full blood count, C-reactive protein, renal and liver function tests, and serum electrolytes were normal. His random blood glucose was 76 mg/dl. Subsequent investigations revealed very low serum cortisol (< 22 nmol/l; normal 55–304 nmol/l) and very high ACTH (> 1734 pg/ml; normal 10–50 pg/ml) suggesting glucocorticoid deficiency. His serum

# Future Work



Data extraction from templated notes

Integration with ChatGPT or similar large language models (LLMs)

# Interested in EMERSE?



Contact us to schedule a time with your team for:

- Discussions about usage strategies
- Training
- Live demonstrations (abstractors, IT teams, etc)

# Community Meeting



TOMORROW: May 14, 2025, 1-2 PM ET  
Open to everyone, find registration link at  
<http://project-emerse.org/>

EMERSE-team@umich.edu



Lisa Ferguson  
David Hanauer  
Kellen McClain  
Guan Wang



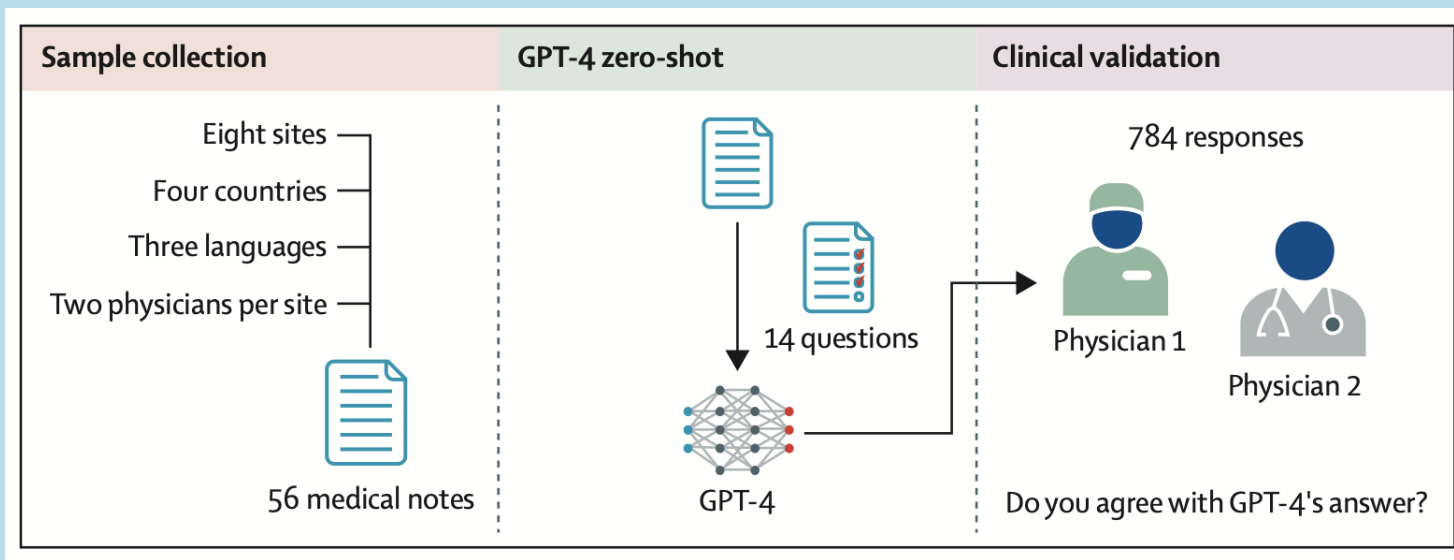
# Large language models (LLMs)

Logical next step for integration

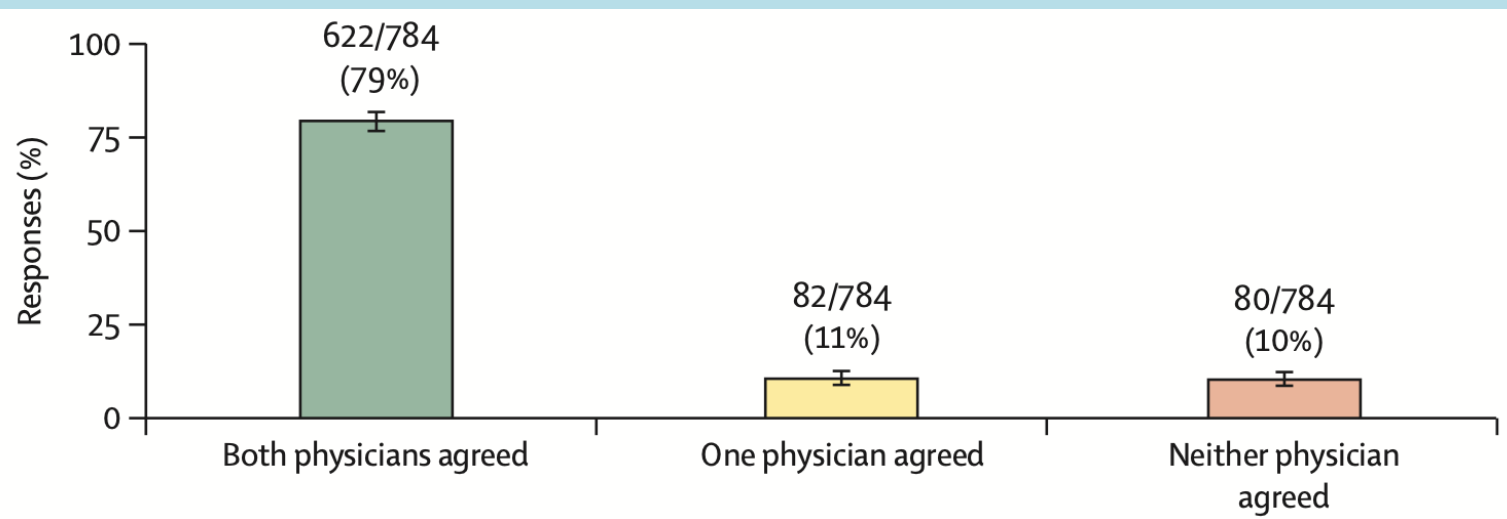
Very exciting, lots of potential




# The potential of Generative Pre-trained Transformer 4 (GPT-4) to analyse medical notes in three different languages: a retrospective model-evaluation study



# The potential of Generative Pre-trained Transformer 4 (GPT-4) to analyse medical notes in three different languages: a retrospective model-evaluation study



# A comparative analysis of privacy-preserving large language models for automated echocardiography report analysis

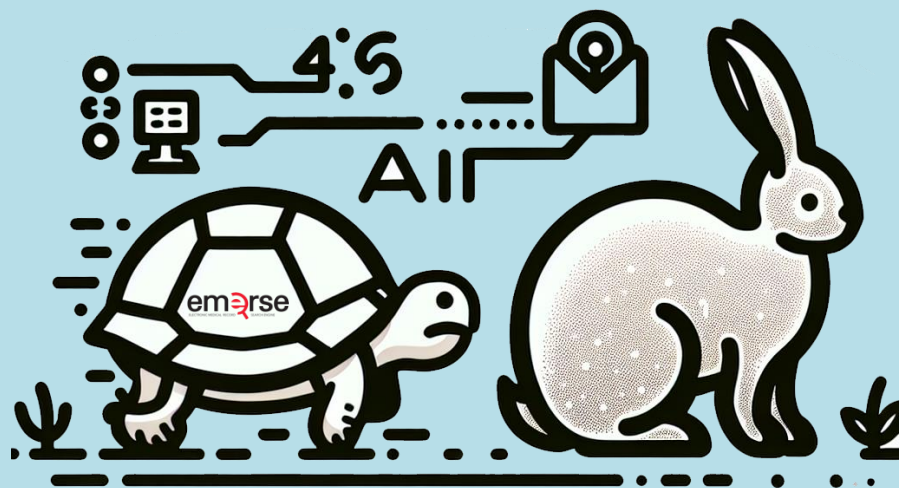
Elham Mahmoudi, MD, MPH<sup>1,2</sup>, Sanaz Vahdati, MD<sup>1</sup>, Chieh-Ju Chao, MD<sup>1,2</sup>,  
Bardia Khosravi, MD, MPH, MHPE<sup>1</sup>, Ajay Misra<sup>1</sup>, Francisco Lopez-Jimenez , MD, MSc<sup>2</sup>,  
Bradley J. Erickson, MD, PhD<sup>\*,1</sup>

**Conclusions:** Our study demonstrates the near-perfect performance of open-source LLMs for automated echocardiography report interpretation with the purpose of registry formation and disease surveillance. While larger models achieved exceptional accuracy through prompt optimization, practical implementation requires balancing performance with computational efficiency.

# Large language models (LLMs)

Some concerns too

We are moving ahead cautiously



# GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language Models

Iman Mirzadeh<sup>†</sup>  
Oncel Tuzel

Keivan Alizadeh  
Samy Bengio

Hooman Shahrokhi\*  
Mehrdad Farajtabar<sup>†</sup>

Apple

benchmark. Furthermore, we investigate the fragility of mathematical reasoning in these models and demonstrate that their performance significantly deteriorates as the number of clauses in a question increases. We hypothesize that this decline is due to the fact that current LLMs are not capable of genuine logical reasoning; instead, they attempt to replicate the reasoning steps observed in their training data. When we add a single clause that appears relevant to the question, we observe significant performance drops (up to 65%) across all state-of-the-art models, even though the added clause does not contribute to the reasoning chain needed to reach the final answer. Overall, our work provides a more nuanced understanding of LLMs' capabilities and limitations in mathematical reasoning.

## GSM-NoOp

Oliver picks 44 kiwis on Friday. Then he picks 58 kiwis on Saturday. On Sunday, he picks double the number of kiwis he did on Friday, but five of them were a bit smaller than average. How many kiwis does Oliver have?

**o1-mini:** Oliver picks kiwis over three days as follows:

- Friday: 44 kiwis
- Saturday: 58 kiwis
- Sunday: Double the number he picked on Friday, which is  $2 \times 44 = 88$  kiwis

However, on Sunday, 5 of these kiwis were smaller than average. We need to subtract them from the Sunday total:  $88$  (Sunday's kiwis)  $- 5$  (smaller kiwis)  $= 83$  kiwis. Now, summing up the kiwis from all three days:  $44$  (Friday)  $+ 58$  (Saturday)  $+ 83$  (Sunday)  $= 185$  kiwis. Oliver has a total of 185 kiwis.

**Llama3-8B:** Let's break it down step by step:

Oliver picks 44 kiwis on Friday. He picks 58 kiwis on Saturday. On Sunday, he picks double the number of kiwis he did on Friday, which is  $2 \times 44 = 88$  kiwis.

Five of the kiwis picked on Sunday are a bit smaller than average, so we subtract 5 from the total number of kiwis picked on Sunday:  $88 - 5 = 83$  kiwis. Now, let's add up the total number of kiwis Oliver has:

$44$  (Friday)  $+ 58$  (Saturday)  $+ 83$  (Sunday)  $= 185$  kiwis

So, Oliver has 185 kiwis in total.

# Counterfactual bias

“...the notion of counterfactual bias is the idea that large language models will often assume that a factual premise in a question by the user is true, even if it's flat-out wrong.”

Source:

**AI can't handle the truth when it comes to the law**

Marketplace Tech

<https://www.marketplace.org/shows/marketplace-tech/ai-cant-handle-the-truth-when-it-comes-to-the-law/>



origin and common usage of "to get into school for dental, eat more lentils"

All Short videos Forums Images Videos News Shopping : More

### ◆ AI Overview

The saying "to get into school for dental, eat more lentils" is a humorous and informal way of suggesting that a good diet, particularly one rich in protein and nutrients like lentils, can help with academic performance or potentially lead to better chances of getting into a desirable dental program. It's a playful, not literal, way of saying that healthy eating can support overall well-being, which might indirectly improve academic success. [🔗](#)

### Origin and Common Usage:

#### **Playful Connection to Health and Success:**

The saying uses the idea that lentils, a healthy and nutritious food, can somehow help with getting into dental school. It's a lighthearted way of suggesting that a healthy lifestyle contributes to achieving academic goals.

#### **Not Literally True:**

While lentils are a good source of protein, fiber, and other nutrients, they don't directly translate into acceptance into dental school. The saying is meant to be a joke or a playful exaggeration.

#### **Common Usage in Informal Settings:**

It's typically used in conversations between friends, family, or in online forums where students may be joking about the challenges of getting into dental school or discussing healthy eating tips.

# Counterfactual bias

What does this have to do with clinical notes?

An EMERSE user searched for:

“deleted cesarean”

What is that?

deleted cesarean

All

Images

Videos

News

Shopping

Short videos

Forums

: More

### ✦ AI Overview

A deleted C-section, or cesarean hysterectomy, refers to **a procedure where the uterus is removed during a cesarean section**. This is becoming more common due to the rise in primary cesarean delivery rates. It differs from a postpartum hysterectomy, which is performed after vaginal or cesarean delivery, but not at the same time as the C-section. [🔗](#)

RESOLVED: Supervision of other normal pregnancy	ICD-10-CM: Z34.80 ICD-9-CM: V22.1
-------------------------------------------------	--------------------------------------

<del>DELETED: Cesarean section wound complications</del>	<del>ICD-10-CM: O90.9</del> <del>ICD-9-CM: 674.34</del>
----------------------------------------------------------	------------------------------------------------------------

RESOLVED: Normal labor	ICD-10-CM: O80, Z37.9 ICD-9-CM: 650
------------------------	-------------------------------------------

# Hallucinations

OpenAI o3 and o4-mini System Card

OpenAI

April 16, 2025



## 3.3 Hallucinations

We evaluate hallucinations in OpenAI o3 and o4-mini against the following evaluations that aim to elicit hallucinations from the models:

Table 4: Hallucination evaluations

Dataset	Metric	o3	o4-mini	o1
SimpleQA	accuracy (higher is better)	0.49	0.20	0.47
	hallucination rate (lower is better)	0.51	0.79	0.44
PersonQA	accuracy (higher is better)	0.59	0.36	0.47
	hallucination rate (lower is better)	0.33	0.48	0.16

# ECRI (Emergency Care Research Institute) Top 10 technology hazards for 2025:

- 1.
- 2.
- 3.
- 4.
5. Fire risk from supplemental oxygen
- 6.
- 7.
8. Poorly managed infusion lines
9. Harmful medical adhesive products
- 10.

<https://home.ecri.org/blogs/ecri-news/artificial-intelligence-tops-2025-health-technology-hazards-list>

# ECRI (Emergency Care Research Institute) Top 10 technology hazards for 2025:

- 1. Risks with AI-enabled health technologies**
2. Unmet technology support needs for home care patients
3. Vulnerable technology vendors and cybersecurity threats
4. Substandard or fraudulent medical devices and supplies
5. Fire risk from supplemental oxygen
6. Dangerously low default alarm limits on anesthesia units
7. Mishandled temporary holds on medication orders
8. Poorly managed infusion lines
9. Harmful medical adhesive products
10. Incomplete investigations of infusion system incidents

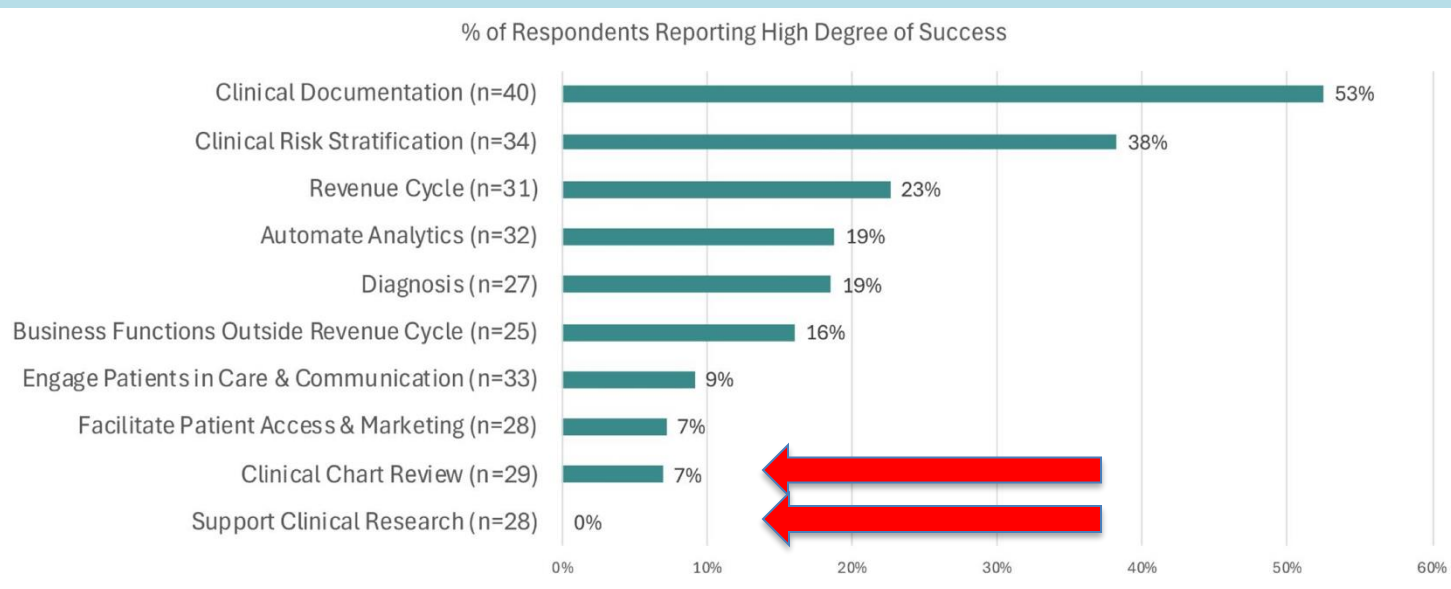
<https://home.ecri.org/blogs/ecri-news/artificial-intelligence-tops-2025-health-technology-hazards-list>



<http://project-emerse.org/presentations.html>

# Adoption of artificial intelligence in healthcare: survey of health system priorities, successes, and challenges

Eric G. Poon , MD, MPH<sup>1,2,3,\*</sup>, Christy Harris Lemak, PhD<sup>4,5</sup>, Juan C. Rojas, MD, MS<sup>6</sup>, Janet Guptill, MPH<sup>5</sup>, David Classen, MD, MS<sup>7</sup>





# This sums it up well

“...while AI technologies like ChatGPT have demonstrated promising potential in the field of dentistry, their integration into clinical practice should be approached with significant caution.”

**Transforming dental diagnostics with artificial intelligence: advanced integration of ChatGPT and large language models for patient care**  
<https://www.frontiersin.org/journals/dental-medicine/articles/10.3389/fdmed.2024.1456208/full>

# Thank you

Questions? Comments?

David Hanauer

hanauer@umich.edu