

Exploring the Impact of Artificial Intelligence on Medicine

Marc M. Triola, MD

Associate Dean, Educational Informatics

Director, Institute for Innovations in Medical Education

NYU Langone Health



Disclosures

- Employee of NYU Grossman School of Medicine, NYU Langone Health
- Board Member, AAMC
- Institute's research funded by AMA Accelerating Change in Medical Education, the Josiah Macy Jr. Foundation, the Stemmler Fund
- Textbook author, Pearson

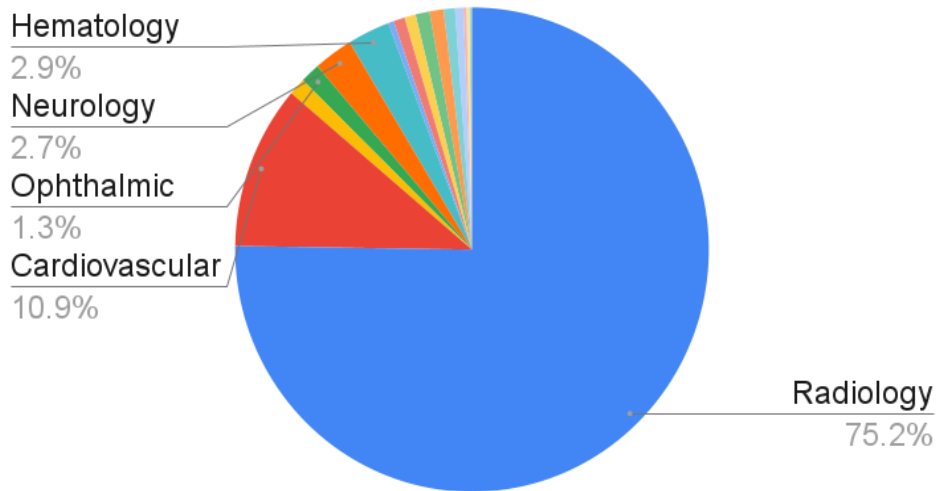
Artificial Intelligence (AI)

Long history of AI innovation in health care and research

FDA has reviewed and authorized almost 750 AI-powered decision support apps and algorithms across many different fields of medicine

30% of radiology practices that responded to a survey indicated that they had adopted AI by 2020

Clinical Area



What is Generative AI?

Generative Artificial Intelligence (AI)

Human language has rules, patterns, and is not random, making it predictable.

Generative AI is a set of algorithms, capable of generating seemingly new, realistic content — such as text, images, video, or music — from examples.

The most powerful generative AI algorithms are built on top of models that are trained on a vast quantity of data to identify underlying patterns for a wide range of tasks.

GPT-4 is the latest release of GPT class of models, a large-scale, multimodal model which can accept image and text inputs and produce text and image outputs.

Create a diagram of the heart's conduction system and a QRS complex to teach a medical student about cardiac electrophysiology

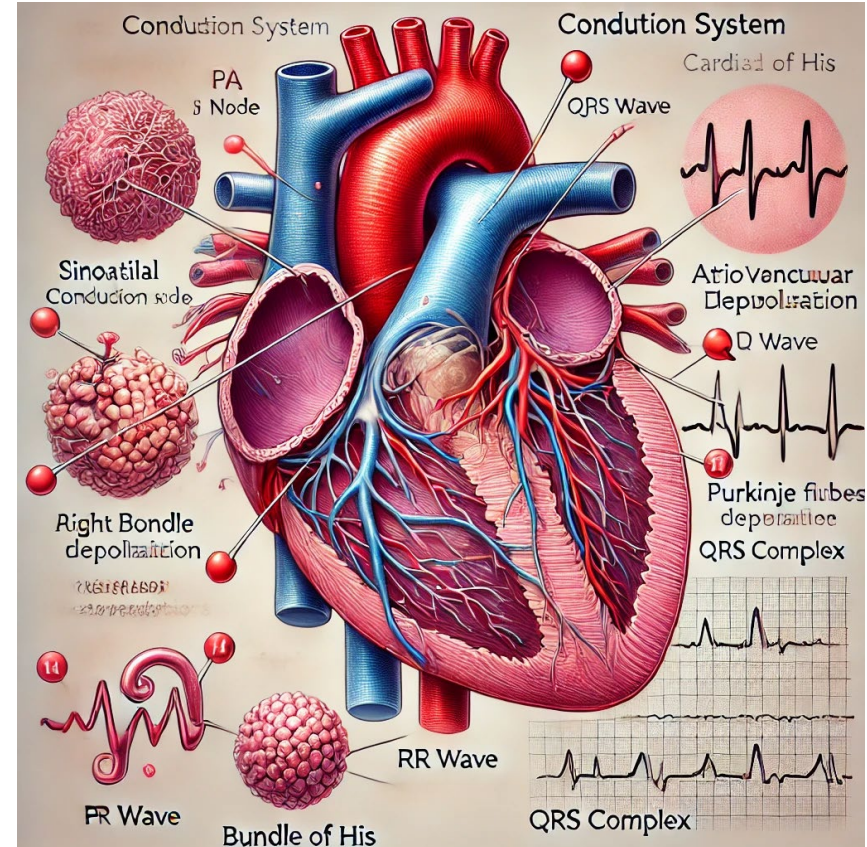
To help explain cardiac electrophysiology to a medical student, we'll diagram both the heart's conduction system and the QRS complex. I'll first walk through the main points of each, and then generate a diagram.

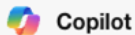
1. Heart's Conduction System

- **Sinoatrial (SA) Node:** The natural pacemaker of the heart, located in the right atrium. It generates electrical impulses that spread through the atria.
- **Atrioventricular (AV) Node:** Located between the atria and ventricles. It slows down the electrical signal before sending it to the ventricles.
- **Bundle of His:** Conducts the electrical signal from the AV node to the ventricles.
- **Right and Left Bundle Branches:** Conduct the signal to the right and left ventricles.
- **Purkinje Fibers:** Spread the electrical signal throughout the ventricles, leading to contraction.

2. QRS Complex

- Represents **ventricular depolarization**, which triggers ventricular contraction.
- **P wave:** Represents atrial depolarization.
- **Q wave:** Small downward deflection after the P wave.
- **R wave:** The major upward deflection representing the bulk of ventricular depolarization.
- **S wave:** Small downward deflection following the R wave.





Copilot

The image you provided shows a **crash cart tray**, which contains essential medications and supplies used during emergency situations, such as cardiac arrests, code blue scenarios, intubation, anaphylaxis, and arrhythmias. Let's take a closer look at some of the medications commonly found in a crash cart:

1. Epinephrine (Epi):

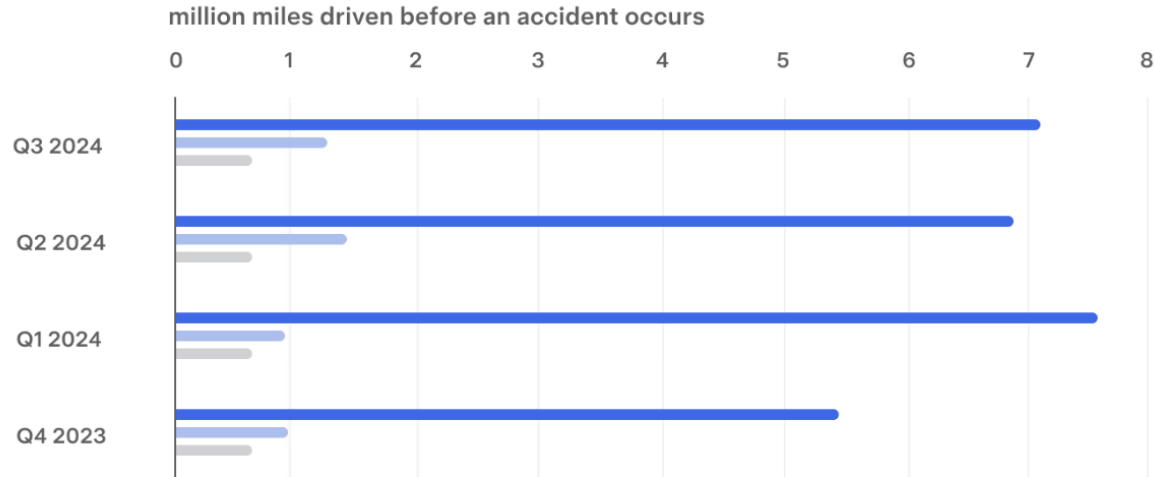
- **Purpose:** Epinephrine is the primary medication used during a code blue. It serves several critical functions:
 - Restarting the heart.
 - Improving heart muscle contractility.
 - Increasing cardiac output.
- **Indications:** Used in situations like Pulseless Electrical Activity (PEA), asystole, ventricular fibrillation (V-fib), pulseless ventricular tachycardia (V-tach), anaphylaxis, and status asthmaticus.
- **Side Effects:** Hypertension, nervousness, restlessness, tachycardia, and chest pain.
- **Dosing:** Two different concentrations exist:
 - **Epinephrine 1:10,000:** Commonly used for cardiac arrest and asystole.
 - **Epinephrine 1:1000:** Used for anaphylaxis (intramuscular administration only due to high concentration).

2. Amiodarone:

- **Purpose:** Used to treat ventricular fibrillation (VF) and ventricular tachycardia (VT) during cardiac arrest.
- **Indications:** Typically administered after epinephrine and vasopressin have failed to convert VF/VT.

Miles Driven Per One Accident

- Tesla vehicles using Autopilot technology
- Tesla vehicles not using Autopilot technology
- United States average



AI as a Diagnostician

Large Language Models Outperform MDs on Diagnosis

***Under Controlled Circumstances**

- AI is more accurate than human MDs at top diagnosis and differentials
- Even though ChatGPT was not trained for diagnosis, it has similar performance to purpose-built diagnostic AI systems
- AI diagnostic accuracy has improved by >15% in the past few months alone

Research Letter

June 15, 2023

Accuracy of a Generative Artificial Intelligence Model in a Complex Diagnostic Challenge

Zahir Kanjee, MD, MPH¹; Byron Crowe, MD¹; Adam Rodman, MD, MPH¹



NEJM AI 2023; 1 (1)
[DOI: 10.1056/AI.p2300031](https://doi.org/10.1056/AI.p2300031)

PERSPECTIVE

Use of GPT-4 to Diagnose Complex Clinical Cases

Alexander V. Eriksen , M.D.,^{1,2} Sören Möller , M.Sc., Ph.D.,^{3,4} and Jesper Ryg , M.D., Ph.D.^{1,2}

Received: July 10, 2023; Revised: September 15, 2023; Accepted: September 29, 2023; Published: November 9, 2023

Clinical Reasoning of a Generative Artificial Intelligence Model Compared With Physicians

Stephanie Cabral, MD¹; Daniel Restrepo, MD²; Zahir Kanjee, MD, MPH¹; [et al](#)

Superhuman performance of a large language model on the reasoning tasks of a physician

Peter G. Brodeur, Thomas A. Buckley, +15 authors [Adam Rodman](#) • Published 14 December 2024 • Medicine, Computer Science

Large Language Models Outperform MDs on Diagnosis

***Under Controlled Circumstances**

- Physicians using AI diagnose better than those using conventional resources
- AI by itself is better than physicians using AI
- MDs anchored to their initial diagnoses, even when the AI suggested alternatives with better reasoning
- MDs used ChatGPT for targeted questions, failing to exploit its ability to analyze entire case histories

nature medicine

Article | Published: 05 February 2025

GPT-4 assistance for improvement of physician performance on patient care tasks: a randomized controlled trial

Towards Accurate Differential Diagnosis with Large Language Models

Daniel McDuff · Mike Schaekermann · Tao Tu · Anil Palepu · Amy Wang · Jake Garrison · Karan Singhal · Yash Sharma · Shekoofeh Azizi · Kavitaulkarni · Le Hou · Yong Cheng · Yun Liu · Sara Mahdavi · Sushant Prakash · Anupam Pathak · Christopher Semturs · Shwetak Patel · Dale Webster · Ewa Dominowska · Juro Gottweis · Joelle Barral · Kat Chou · Greg Corrado · Yossi Matias · Jake Sunshine · Alan Karthikesalingam · Vivek Natarajan · Arxiv (2023)

Original Investigation | Health Informatics

October 28, 2024

Large Language Model Influence on Diagnostic Reasoning **A Randomized Clinical Trial**

AI-Powered Clinical Decision Support

We empower clinicians with our AI platform for developing differential diagnoses and drafting clinical plans.

[TRY GLASS](#)

GLASS AI NOTEBOOK

Enter Patient Summary

Provide age, sex, relevant past medical history, medications, presenting symptoms, associated symptoms, descriptions of relevant studies (including labs and imaging), the illness course and any additional information you might include when consulting another physician about your patient.

Examples

CLINICAL PROBLEM IDENTIFICATION

Chief pain with hypoxemia 1 month Subacute progressive dyspnea Confusion and disorientation Chest pain with ST segment elevation

RECENT AND PRESENT

Confusion

71-year-old male with a history of MI presents with subacute progressive dyspnea on exertion and is found to have bilateral lower extremity edema, an S3 heart sound, and JVD on physical exam. ST-segment elevation is 30%.

Draft Differential Diagnosis **Draft Clinical Plan**

DIFFERENTIAL DIAGNOSIS DRAFT

This differential diagnosis was drafted by a large language model fine-tuned on complex algorithms used by the Glass Health Clinical Team.

The following diagnoses are entered that a clinician may want to consider and investigate, but should never replace a clinician's judgment.

COMPREHENSIVE REVIEW OF THE CASE The patient is a 71-year-old male with a history of myocardial infarction. He presents with subacute progressive dyspnea on exertion and is found to have bilateral lower extremity edema, an S3 heart sound, and jugular venous distention (JVD) on physical exam. The elevation factor (EF) is newly revealed to be 30%. The information is provided about the patient's medications, physical, mental history, laboratory data, course of illness, or imaging data and other studies.

ACUTE DECOMPENSATED HEART FAILURE

The 71-year-old patient presents with subacute progressive dyspnea on exertion, a confirmed elevation of jugular venous distention (JVD), the history of myocardial infarction, and bilateral lower extremity edema. Physical examination findings of bilateral lower extremity edema, an S3 heart sound, and jugular venous distention (JVD) further support the diagnosis. The newly revealed elevation factor (EF) of 30% adds to the diagnosis.

Expand your Differential Diagnosis

Glass helps you draft a differential diagnosis for complex patients. We analyze your patient summary and suggest diagnoses you may consider and investigate.

DIFFERENTIAL DIAGNOSIS **CLINICAL PLAN** **Consenting**

1. Pulmonary Embolism

The patient presented with sudden onset dyspnea, chest pain, tachycardia, and hypoxemia which could be consistent with a pulmonary embolism. The presence of risk factors such as recent surgery and history of smoking could further support this diagnosis.

2. Pneumonia

The patient presented with sudden onset dyspnea, chest pain, and hypoxemia which could be consistent with pneumonia. The presence of fever, cough, sputum production, leukocytosis, elevated inflammatory markers, and lobar or interstitial infiltrates on chest imaging could further suggest this diagnosis.

3. Acute Exacerbation of Chronic Obstructive Pulmonary Disease (COPD)

The patient presented with sudden onset dyspnea.

Draft Clinical Plans

Glass drafts an assessment and plan based on your patient summary in seconds, suggesting evidence-based diagnostic and treatment steps for you to consider.

DIFFERENTIAL DIAGNOSIS **CLINICAL PLAN** **Consenting**

Daytime Hypoxemia in Obstructive Sleep Apnea

A 50-year-old male with a history of obstructive sleep apnea, diabetes mellitus, and BMI 35 presents with daytime hypoxemia, fatigue, and headache. The patient's obesity and history of obstructive sleep apnea suggest that his hypoxemia may be related to hypoventilation or inadequate treatment of his sleep apnea. Additionally, the presence of diabetes mellitus raises concern for potential microvascular complications affecting oxygen delivery. The differential diagnosis.

Rapid clinical guidance

Pathway provides access to concise summaries of clinical guidelines, trials and more, saving you hours every week.

[Start for Free](#)

pathway

Q Search Clinical Topics... Feedback

Home > Diseases > Non-ST-elevation myocardial infarction

Table of Contents

Overview

Guidelines

Screening and diagnosis

Classification and risk stratification

Diagnostic investigations

Respiratory support

Medical management

Inpatient care

Nonpharmacologic interventions

Therapeutic procedures

Perioperative care

Surgical interventions

Specific circumstances

Patient education

Preventative measures

Follow-up and surveillance

Quality improvement

Clinical findings

Studies

References

Guidelines

1. Screening and diagnosis

2. Classification and risk stratification

3. Prognostic scores

4. As per ESC 2021 guidelines, consider using established risk scores for prognosis estimation.

Show 6 more >

Calculator

ACTION score for risk of ICU admission

Calculator

Canadian Cardiovascular Society

AI as a Patient Communicator

AI Answering Patient Questions

- Patients prefer AI-generated text
- AI responses perceived as being more empathic, friendlier
- Human brevity viewed negatively by patients
- AI responses rated as higher quality, more accurate, safer
- AI can meet reading level and health literacy needs of patients more effectively and consistently

JAMA Internal Medicine | [Original Investigation](#)

Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum

John W. Ayers, PhD, MA; Adam Poliak, PhD; Mark Dredze, PhD; Eric C. Leas, PhD, MPH; Zechariah Zhu, BS; Jessica B. Kelley, MSN; Dennis J. Faix, MD; Aaron M. Goodman, MD; Christopher A. Longhurst, MD, MS; Michael Hogarth, MD; Davey M. Smith, MD, MAS

JAMA Oncology | [Brief Report](#)

Physician and Artificial Intelligence Chatbot Responses to Cancer Questions From Social Media

David Chen, BMS; Rod Parsa, MSc; Andrew Hope, MD; Breffni Hannon, MBChB; Ernie Mak, MD; Lawson Eng, MD; Fei-Fei Liu, MD; Nazanin Fallah-Rad, MD; Ann M. Heesters, PhD; Srinivas Raman, MD, MAS



Putting ChatGPT's Medical Advice to the (Turing) Test: Survey Study

Oded Nov ¹; Nina Singh ²; Devin Mann ^{2,3}

JAMA
Network | **Open**



[Research Letter](#) | [Health Informatics](#)

Perspectives on Artificial Intelligence–Generated Responses to Patient Messages

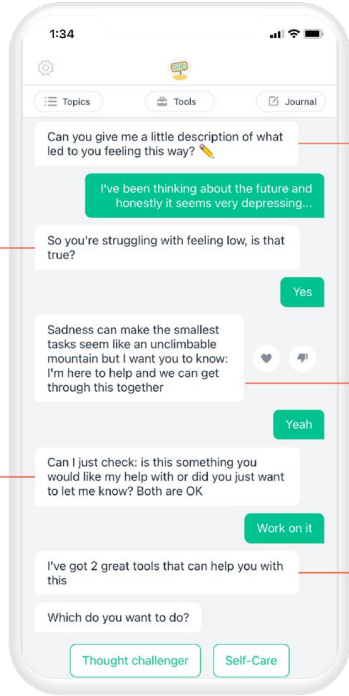
Jiyeon Kim, PhD, MPH; Michael L. Chen, BA; Shawheen J. Rezaei, MPH; April S. Liang, MD; Susan M. Seav, MD; Sonia Onyeka, MD; Julie J. Lee, MD, MPH; Shivam C. Vedak, MD, MBA; David Mui, MD, MBA; Rayhan A. Lal, MD; Michael A. Pfeffer, MD; Christopher Sharp, MD; Natalie M. Pageler, MD, MEd; Steven M. Asch, MD, MPH; Eleni Linos, MD, DrPH



"Doctor ChatGPT, Can You Help Me?" The Patient's Perspective: Cross-Sectional Study

Jonas Armbruster ¹; Florian Bussmann ¹; Catharina Rothhaas ¹; Nadine Titz ¹; Paul Alfred Grützner ¹; Holger Freischmidt ¹

Patient-Facing Chronic Disease Management at Scale



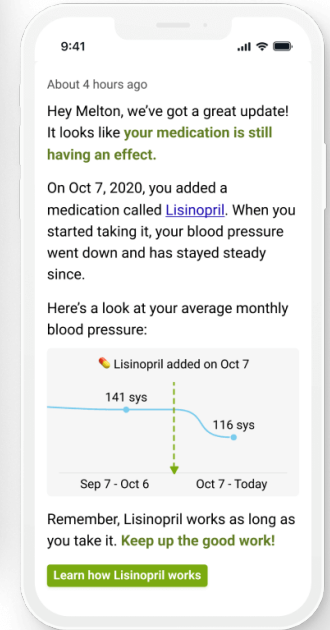
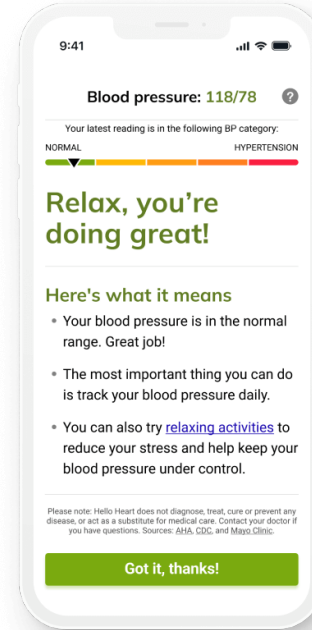
Understands the user state

Gives autonomy to the user to choose their own path

Uses NLP to allow the user to share their problems in their own words.

Shows empathy

Offers actionable tools to work on the problem



<https://woebothealth.com/>

<https://www.helloheart.com>



Hippocratic AI
— Do No Harm —

Foundation Model

Safety

Research

Company



Keisha

CHF Discharge

Rating by Nurses: 82%

Style: Direct



As seen on:



Diane

CKD Chronic Care

Rating by Nurses: 80%

Style: Engaging



Wyatt

CKD Chronic Care

Rating by Nurses: 86%

Style: Engaging

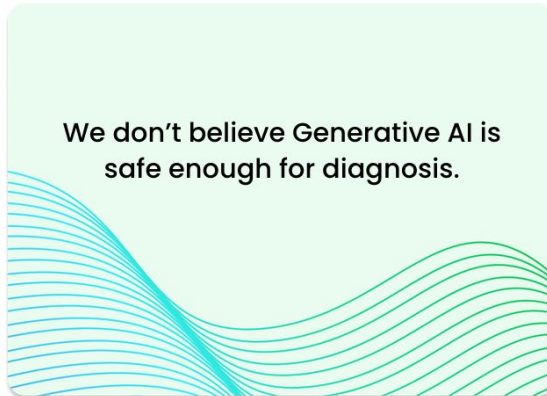


Nancy

Pre-Op Colonoscopy

Rating by Nurses: 88%

Style: Direct

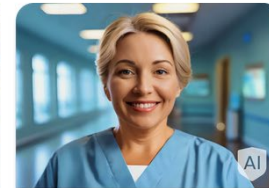


Jasmine

Pre-Op Colonoscopy

Rating by Nurses: 85%

Style: Engaging



Nina

HRT Follow Up

Rating by Nurses: 85%

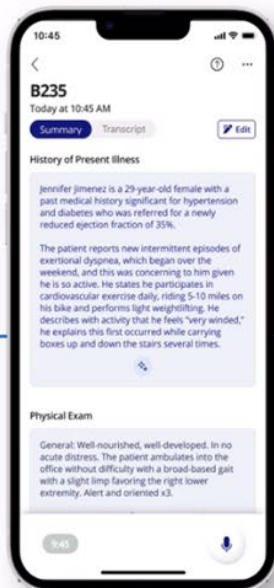
Style: Engaging

AI as a Scribe

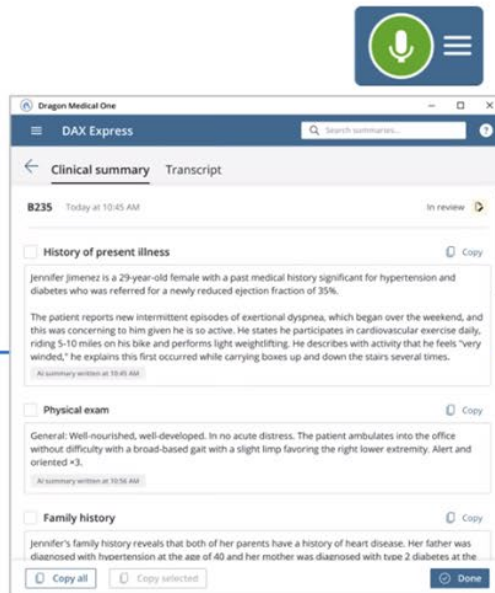
Dragon Ambient Experience (DAX) CoPilot



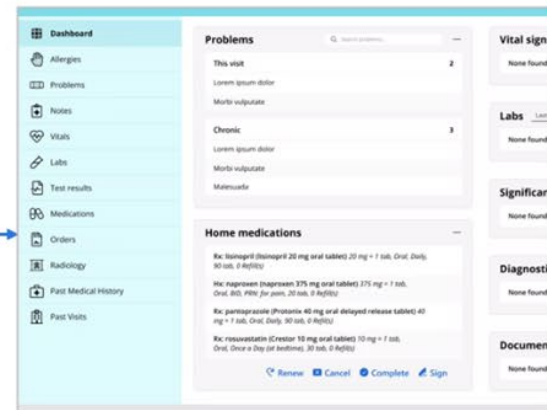
Ambient conversation recorded via unified mobile app for ambient recordings and smartphone mic



Draft clinical summary available in mobile app within seconds after encounter



Transfer to Dragon Medical One for clinician review, edit, and transfer to EHR



Once transferred, clinician signs off on final note in EHR

Considerations

- **Hallucinations:** fabricated facts, values
- **Lack of evidence:** most studies have been on structured 'perfect' cases
- **Accuracy:** How do we define the best AI?
- **Non-generalizable models:** What works for one health system/specialty may not perform as well in others
- It's a first draft for now, but progress is happening very quickly

Unique Challenges of GenAI in Medicine

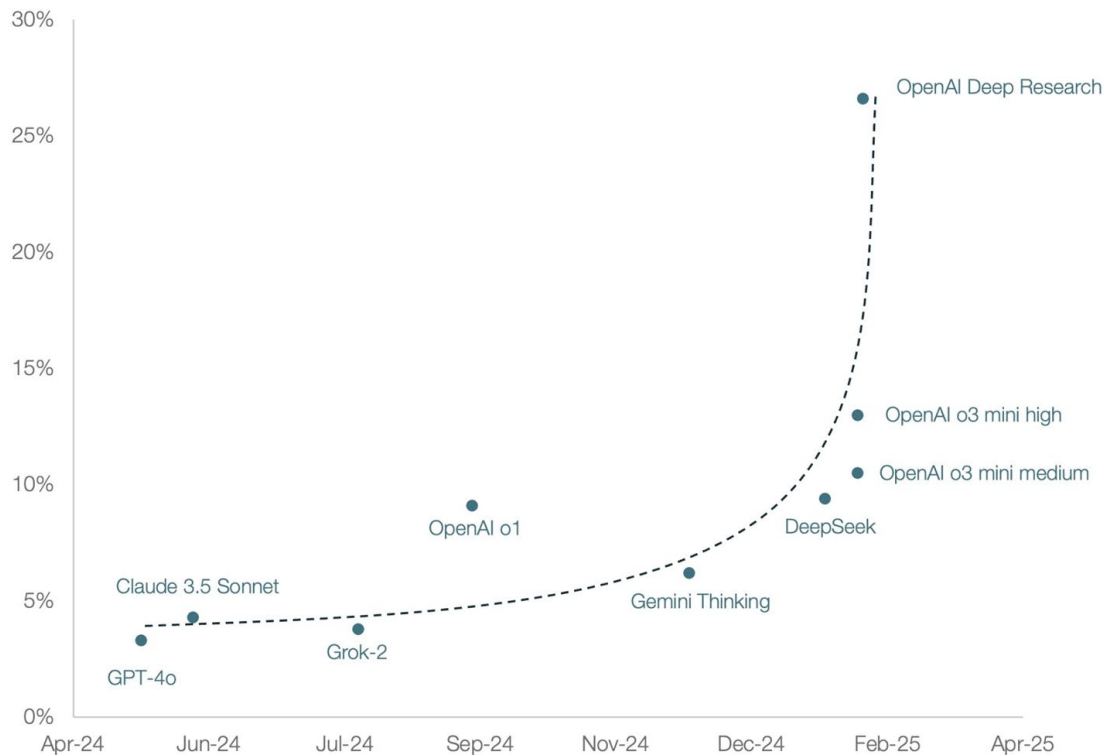
The need for a human in the loop. The increasing need to verify whether generated content is based on fact or inference elevates the need for a new level of quality control.

Explainability. A lack of transparency into the origins of generated content and traceability of root data could make it difficult to update models and scan them for potential risks.

Privacy. Generative AI's use of clinical images and medical records could increase the risk of data leak and challenge patient privacy.

Evidence of effectiveness: Lack of evaluative frameworks; Few studies or RCTs to date of these technologies

AI Scores on *Humanity's Last Exam*



**3,000 challenging questions
across over a hundred subjects**

Source: Tomas Pueyo for Uncharted Territories, with data from Dan Hendrycks, of Humanity's Last Exam

<https://agi.safe.ai/>

Healthcare is an optimal place for AI transformation

- Demand for healthcare far outstrips supply.
- Patients want care that is both more cutting-edge AND more human.
- AI will be a new member of the health care team - part scribe, part co-pilot. This will redefine some roles.
- AI can tackle the data overload facing physicians enabling higher quality, safer, efficient, and more precise care.

Turbulent, Existentially Difficult, but Ultimately Better for Patients

- Risk of 'never-skilling' or 'de-skilling' when AI does the work
- Changing scope of practice
- Automation bias
- What happens when the AI is consistently better than humans at certain tasks?

AI as a Licensed Health Care Provider

H.R.238 - Healthy Technology Act of 2025

- “The Healthy Technology Act of 2025 would amend the Federal Food, Drug, and Cosmetic Act to allow AI and machine learning to qualify as practitioners eligible to prescribe drugs if authorized by the state involved and approved, cleared, or authorized by the US Food and Drug Administration (FDA) for other purposes.”
- Similar bills in the US House of Representatives in 2021 and 2023. Both were referred to the House Committee on Energy and Commerce and to its Subcommittee on Health, where no further action was taken.

How to Stay Current on AI

- My email: marc.triola@nyulangone.org
- AI Breakfast Newsletter: <https://aibreakfast.beehiiv.com/>
- Ground Truths from Dr. Eric Topol: <https://erictopol.substack.com/>
- NEJM AI: <https://ai.nejm.org/>
- JAMA+ AI: <https://jamanetwork.com/channels/ai>

Links, tutorials, examples available at:
ai.iime.cloud

