

From Bytes to Bedside: Exploring the Impact of Artificial Intelligence on Medicine and Education

Marc M. Triola, MD

Associate Dean, Educational Informatics

Director, Institute for Innovations in Medical Education



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

Generative Artificial Intelligence (AI)

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

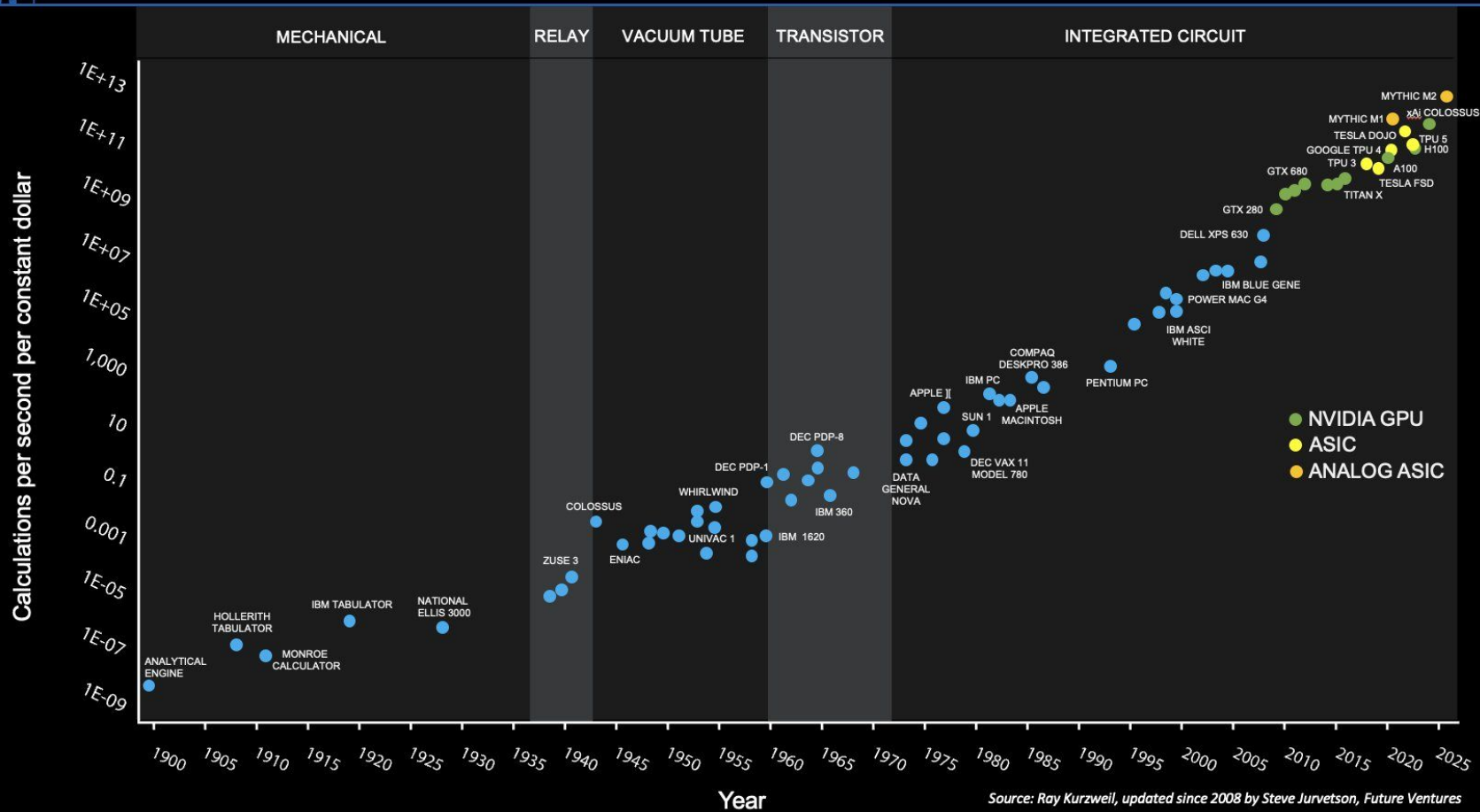
Generative AI is a set of algorithms, capable of generating seemingly new, realistic content — such as text, images, video, or music — from examples.

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.



128 YEARS OF MOORE'S LAW



Miles Driven Per One Accident

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



LLMS are now multimodal

DALL-E 2 (2023)



DreamStudio (2023)



DALL-E 3 (2024)



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

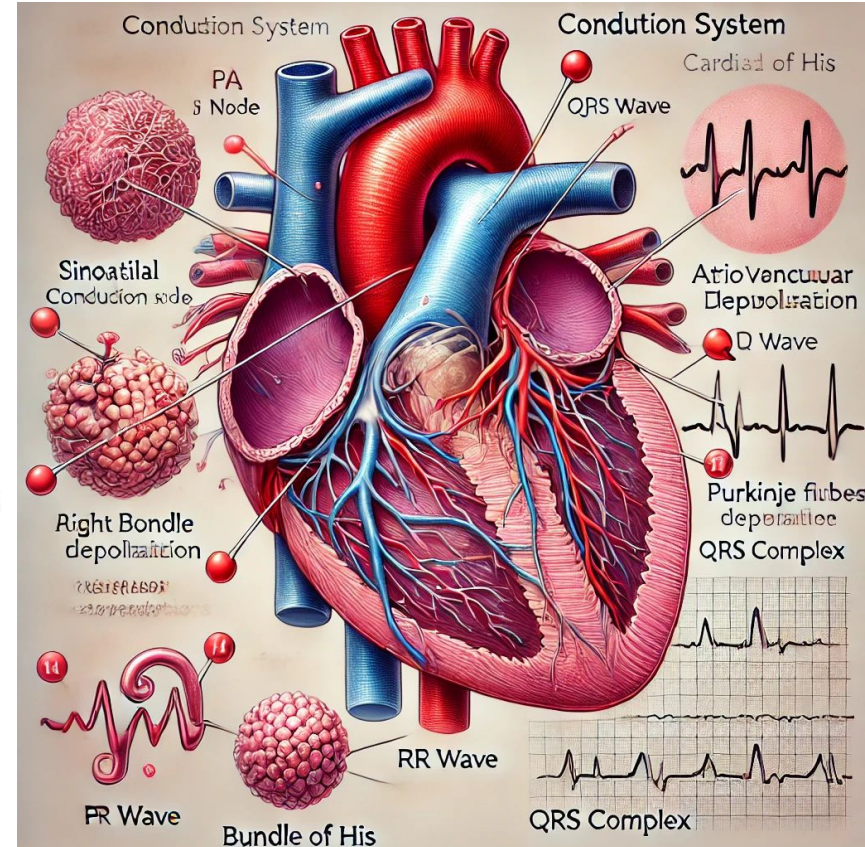
To help explain cardiac electrophysiology to a medical student, we'll diagram both the heart's conduction system and the QRS complex. I'll first walk through the main points of each, and then generate a diagram.

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.

Analyse these blood tests results

WBC (Leukocytes)
REL APPOINTMENT WITH IN

cell count	6.4	$10^9/L$
estimation	146	g/L
Ht	43.9	%
corpuscular volume (MCV)	79.4	fL
basic hemoglobin (MCH)	26.4	pg
ant	259	$10^9/L$
platelet	4.1	$10^9/L$
WBC count	1.5	$10^9/L$
platelet	0.5	$10^9/L$
RBC count	0.3	$10^9/L$
platelet	0.1	$10^9/L$

Answer

Ah, the results of your blood test! Let's dive into this hematological adventure, shall we?

1. Total White Cell Count (WBC):

- **Result:** $6.4 \times 10^9/L$
- **Normal Range:** $3.70 - 9.50 \times 10^9/L$
- **Analysis:** Within the normal range. No need to call the white blood cell cavalry just yet.

2. Hemoglobin Estimation:

- **Result:** 146 g/L
- **Normal Range:** 130.00 - 170.00 g/L
- **Analysis:** Right in the middle of the normal range. Your red blood cells are carrying oxygen just fine.

3. Hematocrit:

- **Result:** 43.9%
- **Normal Range:** 40.00 - 50.00%
- **Analysis:** Perfectly normal. Your blood's viscosity is like a well-blended smoothie.

4. Mean Corpuscular Volume (MCV):

- **Result:** 79.4 fL
- **Normal Range:** 83.00 - 101.00 fL
- **Analysis:** Slightly on the lower side, indicating possible microcytic cells. But not too concerning.

5. Mean Corpuscular Hemoglobin (MCH):

- **Result:** 26.4 pg
- **Normal Range:** 27.00 - 32.00 pg
- **Analysis:** A tad low, which might align with the lower MCV. Keep an eye on your iron levels.

6. Platelet Count:

- **Result:** $259 \times 10^9/L$
- **Normal Range:** $150.00 - 410.00 \times 10^9/L$
- **Analysis:** Normal. Your blood's clotting ability is on point.

“Overall, your blood test results are largely within normal ranges. The slightly lower MCV and MCH might warrant a bit of attention, but nothing too alarming. Keep living life to the fullest, and maybe throw in some spinach for good measure!”

AI as a Diagnostician

Large Language Models Outperform MDs on Diagnosis

*Under Controlled Circumstances

- Several studies evaluated ChatGPT-4's ability to diagnose New England Journal Challenge Cases (CPCs)
- AI did better than human MDs at top diagnosis and differential.
- 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.2300031](https://doi.org/10.1056/AI.2300031)

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

RCT: Large Language Model Influence on Diagnostic Reasoning

POPULATION

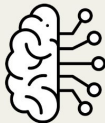
26 Attending physicians
24 Resident physicians



US-trained physicians with training in family medicine, internal medicine, or emergency medicine

INTERVENTION

50 Participants randomized



25 Generative artificial intelligence (AI) chatbot

Participants with access to AI chatbot were allocated 60 min to review up to 6 clinical vignettes



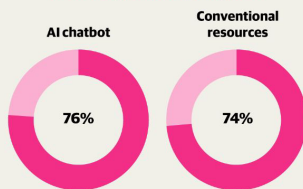
25 Conventional resources

Participants with access to conventional online resources such were allocated 60 min to review up to 6 clinical vignettes

FINDINGS

There was no significant difference in diagnostic performance between physicians in the chatbot group and the conventional resource group

Median diagnostic reasoning score



Adjusted difference between groups:

2 percentage points (95% CI, -4 to 8 percentage points);
 $P = .60$

SETTINGS / LOCATIONS



3 Academic medical centers in the US

PRIMARY OUTCOME

Diagnostic performance, assessed using a standardized rubric including differential diagnosis accuracy, appropriateness of supporting/opposing factors, and next evaluation steps; graded via blinded expert consensus

Goh E, Gallo R, Hom J, et al. Large language model influence on diagnostic reasoning: a randomized clinical trial. *JAMA Netw Open*. 2024;7(10):e2440969. doi:10.1001/jamanetworkopen.2024.40969

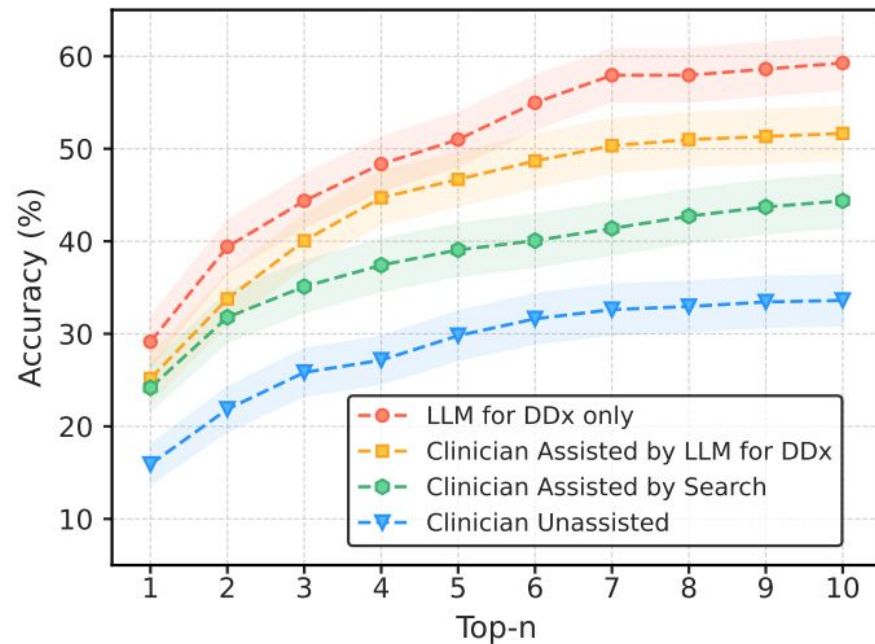
© AMA

- “the LLM alone performed significantly better than both groups of humans” ($P = .60$)
- MD+LLM did not significantly improve clinical reasoning compared with conventional resources
- MDs anchored to their initial diagnoses, even when the chatbot suggested alternatives with better reasoning
- MDs used ChatGPT for targeted questions, failing to exploit its ability to analyze entire case histories

Towards Accurate Differential Diagnosis with Large Language Models

¹Google Research, ²Google DeepMind

- 302 cases sourced from the NEJM
- Used a Google-developed LLM for DDx
- LLM generated a DDx by itself
- Two internal medicine physicians:
 - Each generated an unassisted DDx
 - One used assistance from search engines and standard medical resources
 - One used an LLM specifically trained to create DDx based on Palm 2
- Both approaches took the same amount of time
- The LLM-assisted physician produced a longer DDx



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 REPRESENTATION

Chief complaint with hypoxemia 1 month

Subacute progressive dyspnea

Chest pain with ST segment elevation

HISTORY AND PHYSICAL

Cardiothoracic

NEXT

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

ADVANCED Differential Diagnosis Draft

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

The following diagnosis are selected 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 electrocardiogram (ECG) is newly revealed to be 30% ST-segment elevation. The following information is provided about the patient's medications, physical exam, 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 30% ST-segment elevation of acute myocardial infarction (AMI), the history of myocardial infarction, and bilateral lower extremity edema. The physical examination findings of bilateral lower extremity edema, an S3 heart sound, and jugular venous distention (JVD) further support the diagnosis. The most recent electrocardiogram (ECG) is 30% ST-segment elevation.

Based on your input

This clinical plan was drafted by a large language model provided with evidence-based clinical guidelines curated by the Glass Health Clinical Team.

The clinical plan draft contains a suggested assessment, diagnostic studies, and treatment options that a clinician may consider but should never replace a clinician's judgment.

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

Generating

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

Generating

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

Home

Diseases

Studies

Pathways

Calculators

5:00

Evidence Review

Close

Pathway assessment

Strong recommendation

Strongly favor this test or intervention.

Moderate recommendation

Favor this test or intervention.

Weak recommendation

Weakly favor this test or intervention.

Do not recommend

Avoid this test or intervention.

Expert recommendation

Expertly favor this test or intervention.

Insufficient evidence

Insufficient evidence to make a recommendation.

Upgraded statement

No formal recommendation made.

Reference

Jean-Philippe Collet, Volker Thiele, Emanuele Di Mario et al. 2020 ESC Guidelines for the management of acute coronary syndromes in patients presenting without prior ST-segment elevation. Eur Heart J. 2021 Apr 1;42(14):1291-1359.

Open reference PDF.

Q Search Clinical Topics...

Feedback

Table of Contents

Home > Diseases > Non-ST-elevation myocardial infarction

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

Non-ST-elevation myocardial infarction

Guideline sources

The following summarized guidelines for the evaluation and management of non-ST-elevation myocardial infarction are prepared by our editorial team based on guidelines from the American Heart Association (AHA/ACC/SCA) (2022), the European Society of Cardiology (ESC) (2021), the European Association for Cardio-Thoracic Surgery (EACTS/ESC) (2019), and the American Heart Association (AHA/ACC) (2014; 2013).

1 2 3 4 5

Calculator

ACTION score for risk of ICU admission

Calculator

Canadian Cardiovascular Society.

Guidelines

1. Screening and diagnosis

2. Classification and risk stratification

3. Prognostic scores

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

Show 6 more >



Digital medicine

A clinical certification pathway for generalist medical AI systems

On demonstrating consistent reliability, the AI would gain conditional autonomy, comparable to a specialist in training, and perform specialty-level tasks with minimal supervision but with strict error monitoring.

After proving proficiency across diverse scenarios, the AI would achieve full certification, equivalent to becoming a consultant or attending physician.

A clinical certification pathway for generalist medical AI systems. Rajpurkar, Pranav et al. The Lancet, Volume 405, Issue 10472, 20

AI as a Patient Communicator

AI Answering Patient Questions

- 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
- “Satisfaction was consistently higher with AI-generated responses than with clinicians overall and by specialty. “

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 Titze ¹; Paul Alfred Grützner ¹; Holger Freischmidt ¹

Patient-Facing Chronic Disease Management at Scale

Screen 1: Woebo Health Chatbot

1:34

Topics Tools Journal

Can you give me a little description of what led to you feeling this way?

I've been thinking about the future and honestly it seems very depressing...

So you're struggling with feeling low, is that true?

Yes

Sadness can make the smallest tasks seem like an unclimbable mountain but I want you to know: I'm here to help and we can get through this together

Yeah

Can I just check: is this something you would like my help with or did you just want to let me know? Both are OK

Work on it

I've got 2 great tools that can help you with this

Which do you want to do?

Thought challenger Self-Care

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

Understands the user state

Shows empathy

Gives autonomy to the user to choose their own path

Offers actionable tools to work on the problem

<https://woebothealth.com/>

Screen 2: Hello Heart Blood Pressure App

9:41

Blood pressure: 118/78

Your latest reading is in the following BP category:

NORMAL HYPERTENSION

Relax, you're doing great!

Here's what it means

- Your blood pressure is in the normal range. Great job!
- The most important thing you can do is track your blood pressure daily.
- You can also try [relaxing activities](#) to reduce your stress and help keep your blood pressure under control.

Please note: Hello Heart does not diagnose, treat, cure or prevent any disease, or act as a substitute for medical care. Contact your doctor if you have questions. Sources: AHA, CDC, and Mayo Clinic.

Got it, thanks!

Screen 3: Hello Heart Medication Management App

9:41

About 4 hours ago

Hey Melton, we've got a great update! It looks like **your medication is still having an effect.**

On Oct 7, 2020, you added a medication called [Lisinopril](#). When you started taking it, your blood pressure went down and has stayed steady since.

Here's a look at your average monthly blood pressure:

Lisinopril added on Oct 7

141 sys 116 sys

Sep 7 - Oct 6 Oct 7 - Today

Remember, Lisinopril works as long as you take it. **Keep up the good work!**

[Learn how Lisinopril works](#)

<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

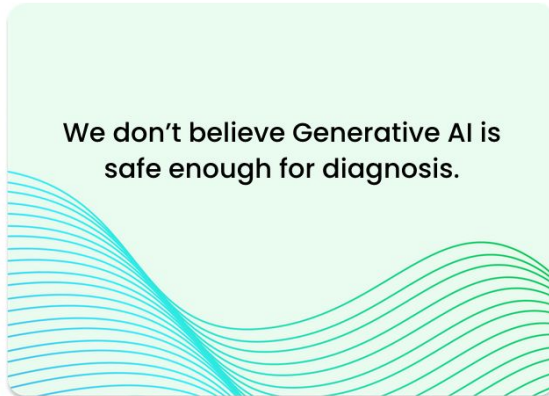


Nancy

Pre-Op Colonoscopy

Rating by Nurses: 88%

Style: Direct

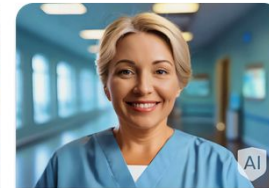


Jasmine

Pre-Op Colonoscopy

Rating by Nurses: 85%

Style: Engaging



Nina

HRT Follow Up

Rating by Nurses: 85%

Style: Engaging

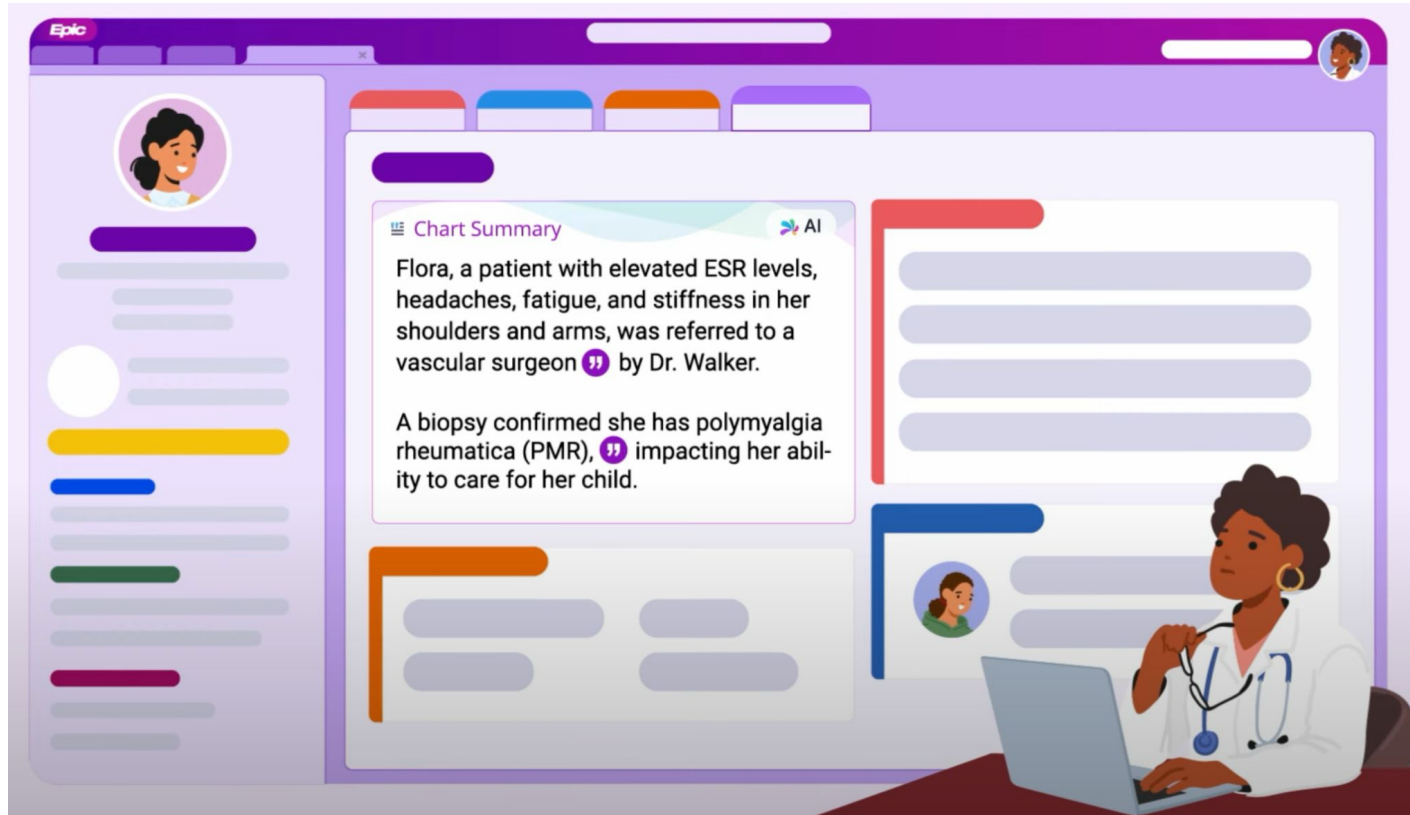
AI as a Scribe

Adapting Discharge Instructions to Patient Need

- Used Chat-GPT4 to generate numerous pediatric ER discharge summaries in English and Spanish, and at different reading levels
- **English-language GPT-generated discharge instructions contained a significantly higher proportion of must-include discharge instructions than those in Spanish**
- “Future research in prompt engineering and GPT-4 performance, both generally and in multiple languages, is needed to reduce potential for health disparities by language and reading level.”

Grade level	English	Spanish
Fifth grade	<p>Dear Parents or Guardians, Your child has had a minor head injury. Good news is, it's not a concussion, and doctors believe it is OK for your child to go home. Still, it's important to keep an eye on them and help them get better. Here's what you should do:</p> <ol style="list-style-type: none"> Rest and relaxation: Your child needs to rest a lot. They should avoid rough play, running, bike riding, or anything that could hurt their head again. Even activities that need a lot of thinking, like schoolwork, should be reduced for a few days. Pain relief: If your child feels some pain, you can give them an over-the-counter medicine like Tylenol (or its generic form, acetaminophen). Don't give them aspirin—it's not safe for kids. Remember to always follow the instructions on the bottle... 	<p>Instrucciones para el Alta: Lesión Menor en la Cabeza de su Hijo Estimado(a) padre/madre o tutor(a):</p> <ol style="list-style-type: none"> Vigile los Síntomas: Su hijo ha sufrido una lesión menor en la cabeza, pero no ha mostrado signos de conmoción cerebral. Aun así, vigile a su hijo por si aparecen nuevos síntomas, como dolor de cabeza fuerte, vómitos, mareos, dificultad para caminar o hablar, o cambios en su comportamiento. Dolor: Para el dolor, puede darle a su hijo medicamentos sin receta, como paracetamol (Tylenol). Siga las instrucciones de envase para la dosis adecuada. No le de aspirina a su hijo...
Eighth grade	<p>Dear Parent/Guardian, Your child has had a minor head injury. We're glad to tell you that there are no signs of a concussion, but we do need you to monitor them at home for any changes.</p> <p>Rest and activity</p> <ol style="list-style-type: none"> Ensure your child gets plenty of rest in the first 24-48 hours. This 	<p>Instrucciones de Alta para un Lesión Menor de Cabeza Sin Conmoción en Pediatría Estimados padres, Su hijo ha tenido una lesión menor en la cabeza pero no muestra signos de conmoción cerebral. A continuación, encontrará las instrucciones a seguir en casa para ayudarlo a recuperarse.</p> <p>Cuidado en casa</p>

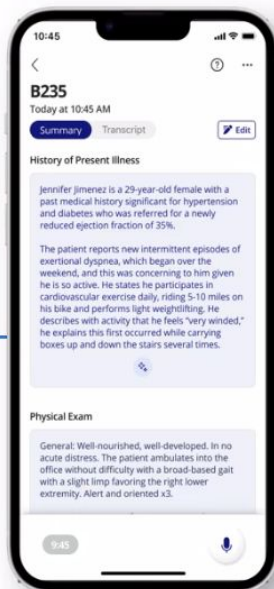
AI-Generated Hospital Course



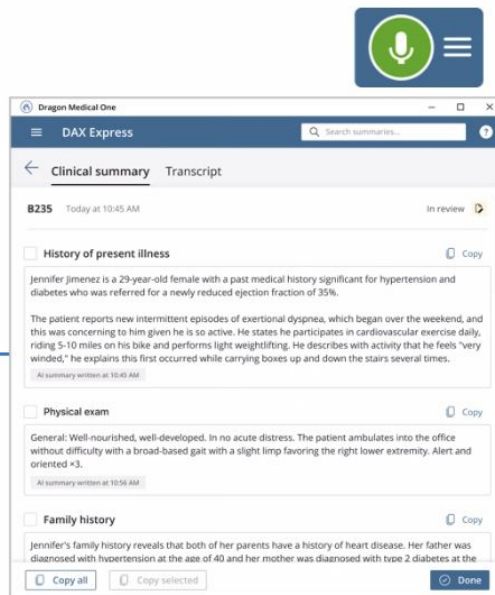
Dragon Ambient Experience (DAX) CoPilot



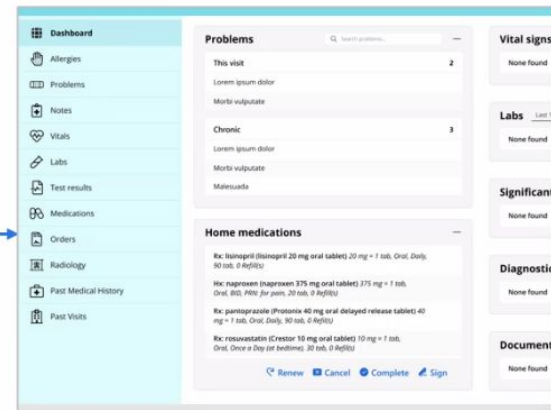
Ambient conversation recorded via unified mobile app for ambient recordings and smartphone mic



Draft clinical summary available in mobile app within seconds after encounter



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



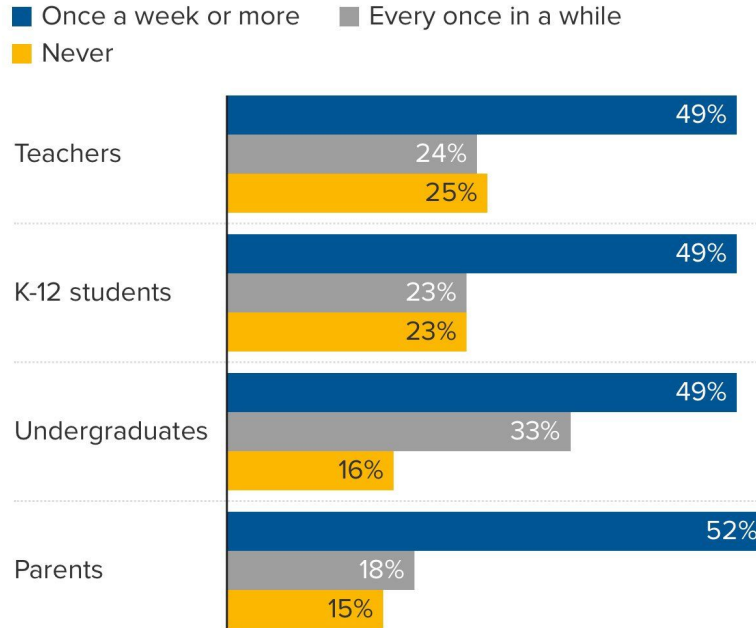
Once transferred, clinician signs off on final note in EHR

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

Generative AI in Medical Education

About half of teachers, students and parents are using AI chatbots at least once a week for work or school



Source: Impact Research
Online survey of 4,007 teachers, parents and students in the U.S. conducted May 7-15, 2024

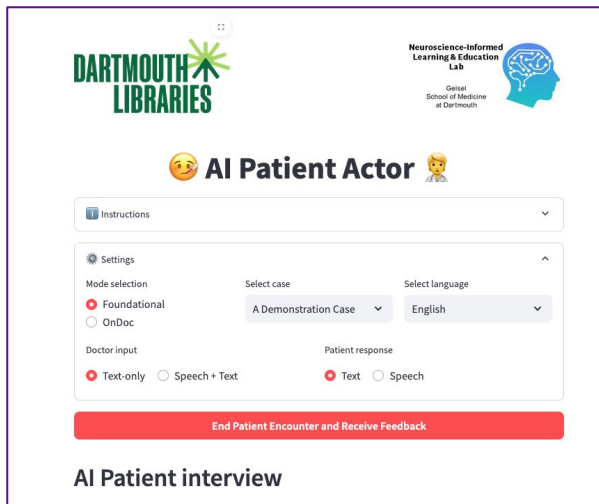
AI % Correct of USMLE-Style Questions

Model	Organization	Release Year	Accuracy (%)
BioLinkBERT	-	2021	45.1
DRAGON	-	2022	47.5
BioMedLM	Stanford	2022	50.3
GPT-3.5 base	OpenAI	2022	60.2
Med-PaLM	Google	2022	67.2
GPT-4 base	OpenAI	2023	86.1
Med-PaLM 2	Google	2023	86.5
GPT-4 (Medprompt)	OpenAI	2023	90.2
Med-Gemini	Google	2024	91.1
o1-preview	OpenAI	2024	96.0

Medical Students and Residents are using ChatGPT to...

- Generate differential diagnoses and plans for PBL cases
- Be a virtual simulated patient by giving the AI OSCE cases
- Create USMLE questions and flashcards
- Draft clinical write ups, summarize the literature
- Inform clinical reasoning on challenging cases and CPCs

AI-Created OSCEs



AI Patient Actor

Instructions

Settings

Mode selection

Foundational

OnDoc

Select case

A Demonstration Case

Select language

English

Doctor input

Text-only

Speech + Text

Patient response

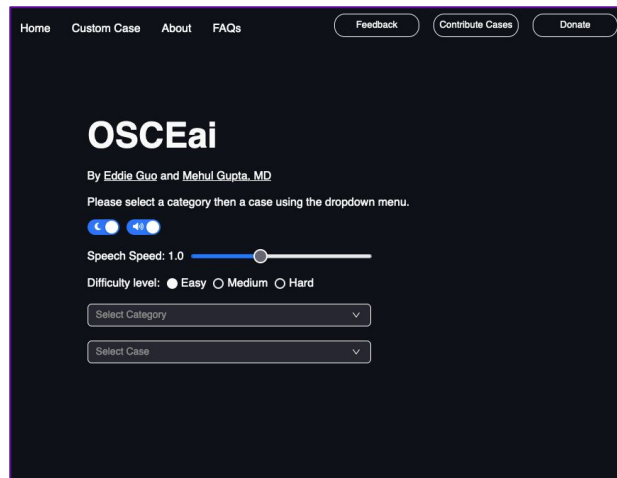
Text

Speech

End Patient Encounter and Receive Feedback

AI Patient interview

<https://ai.dartmouth.edu/patient-actor>



Home Custom Case About FAQs Feedback Contribute Cases Donate

OSCEai

By Eddie Guo and Mehul Gupta, MD

Please select a category then a case using the dropdown menu.

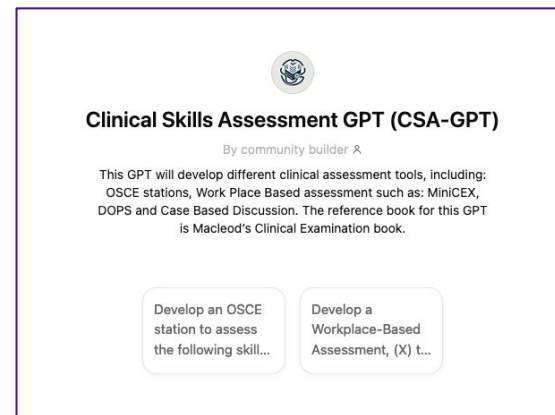
Speech Speed: 1.0

Difficulty level: Easy Medium Hard

Select Category

Select Case

<https://oscegpt.com>



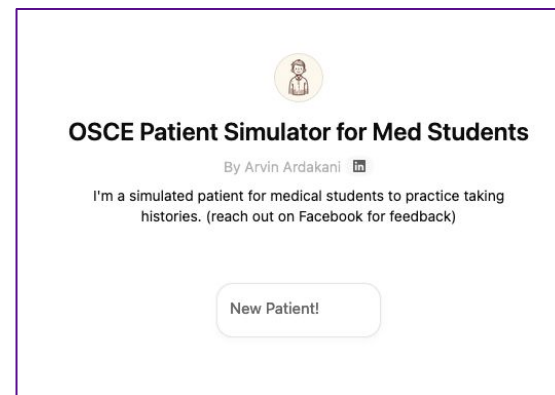
Clinical Skills Assessment GPT (CSA-GPT)

By community builder

This GPT will develop different clinical assessment tools, including: OSCE stations, Work Place Based assessment such as: MiniCEX, DOPS and Case Based Discussion. The reference book for this GPT is Macleod's Clinical Examination book.

Develop an OSCE station to assess the following skill...

Develop a Workplace-Based Assessment, (X) t...



OSCE Patient Simulator for Med Students

By Arvin Ardakani



I'm a simulated patient for medical students to practice taking histories. (reach out on Facebook for feedback)

New Patient!

<https://chatgpt.com>

AI-Graded OSCEs

Rubrics to Prompts: Assessing Medical Student Post-Encounter Notes with AI

Authors: [Andrew R. Jamieson, Ph.D.](#) , [Michael J. Holcomb, M.S.](#) , [Thomas O. Dalton, M.D.](#) , [Krystle K. Campbell, D.H.A.](#) , [Sol Vedovato, M.S.](#) , [Ameer Hamza Shakur, Ph.D.](#) , [Shinyoung Kang, B.S.](#) , [David Hein, M.S.](#) , [Jack Lawson, B.S.](#) , [Gaudenz Danuser, Ph.D.](#) , and [Daniel J. Scott, M.D.](#)  [Author Info & Affiliations](#)

Published November 25, 2024 | NEJM AI 2024;1(12) | DOI: 10.1056/AIcs2400631 | [VOL. 1 NO. 12](#)

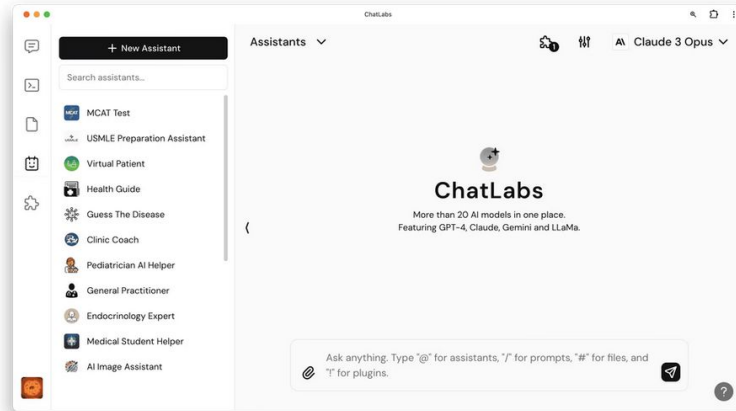
- AI grading system for medical student post-encounter OSCE notes
- Reduced human effort by an 91% and dramatically reduced turnaround time
- Zero-shot architecture with minimal prompt engineering, requires no prior training data
- 89.7% agreement with human expert graders
- Local small open-weight models (such as Llama, Mistral, etc.) can be fine-tuned to achieve similar performance

AI Tutors and Virtual Coaches


Best AI for Medical Students


Your AI study helper with 30+ AI Models


Get the best AI for any task by using
OpenAI, Claude, Gemini, Groq, Mistral, and more all in one place





<https://writingmate.ai/medical>

**File upload**
Unlock deeper insights! Discuss and decode documents in real-time with HPE-Bot, including your handwritten notes.

**Personalized follow up**
Not sure what next? Get AI-recommended follow up questions, for deeper understanding.

**Dive Deeper**
Engage actively and dive deeper into topics with our specialized feature.

**Challenge Your Knowledge**
Test your understanding with tailored vignette MCQs, matched to your conversation.

**Clinical Reasoning**
Improve your critical/clinical reasoning with different challenging open ended questions; choose the orange icon for more challenging scenarios.

<https://hpe-bot.com/>

AI Tutors and Virtual Coaches

- Randomized controlled experiment in a large undergraduate physics course (N = 194) at Harvard
- “We have found that when students interact with our AI tutor, at home, on their own, they learn more than twice as much as when they engage with the same content during an actively taught science course, while spending less time on task.

AI Tutoring Outperforms Active Learning

Gregory Kestin*, Kelly Miller*, Anna Kiales, Timothy Milbourne, and 1 more

nature portfolio

The screenshot displays the TeachGPT.us web application. On the left, a sidebar lists activities: Bernoulli equation (Parts 1-3), Bucket Demo, and Venturi effect (Parts 1-2). The main content area is titled 'Activity 4: Venturi effect || Part 1'. It includes a diagram of a Venturi tube with cross-sectional areas A_1 and A_2 and fluid velocities v_1 and v_2 . The text instructs the user to use the Bernoulli equation to determine pressure changes. Below this, an AI chat interface shows a user asking for the Bernoulli equation, the AI providing the formula $P_1 + \frac{1}{2}\rho v_1^2 + \rho gh_1 = P_2 + \frac{1}{2}\rho v_2^2 + \rho gh_2$, and the user asking for a hint. The AI responds that since heights are equal, the equation simplifies. The user then says 'I am confused about wh'. The interface also includes a 'Load Assignment' button, a 'Log Out' button, and a user login status at the bottom: 'Logged in as happy_student_33@harvard.edu'.

Activity 4: Venturi effect || Part 1

Use the Bernoulli equation to figure out if the pressure of the fluid increases or decreases when it enters a smaller cross-sectional area A of the tube shown below. Use $h_1 = h_2$ to simplify the equation.

Could you please remind me what the expression for the Bernoulli equation is?

Sure! The expression for the Bernoulli equation is:

$$P_1 + \frac{1}{2}\rho v_1^2 + \rho gh_1 = P_2 + \frac{1}{2}\rho v_2^2 + \rho gh_2$$

Here, P stands for pressure, ρ is the fluid density, v represents the velocity of the fluid, g is the acceleration due to gravity, and h refers to the height from a reference level. This equation applies along a streamline of steady, incompressible flow with negligible viscosity.

Thanks! I'd love a hint about what to do next.

Since we're starting with the Bernoulli equation and the heights are the same ($h_1 = h_2$), you can simplify

I am confused about wh

© 2024 TeachGPT.us

AI-Authored Medical Knowledge Assessments

- Great potential - unlimited student-specific questions and assessments
- Almost there:
 - Decent at question stems, issues with choices and distractors
 - “approximately 15% of the questions generated from the detailed prompt required some correction, primarily due to inaccuracies in content or faulty methodology.”
 - Still require human review and editing
 - Human-authored questions are more discriminating of student performance

ChatGPT prompts for generating multiple-choice questions in medical education and evidence on their validity: a literature review

Yavuz Selim Kiyak^{1,*}, Emre Emekli²

Innovation Report

Large Language Models in Medical Education: Comparing ChatGPT- to Human-Generated Exam Questions

Matthias Carl Laupichler, MSc, Johanna Flora Rother, MSc, Ilona C. Grunwald Kadow, PhD, Seifollah Ahmadi, PhD, and Tobias Raupach, MD, MME

Example Use at NYU: Curriculum Mapping

Lecture Description

In this lecture, you will learn about the mechanisms of liver injury and how that manifests in a variety of hepatic diseases. In addition to learning about the ways liver injury manifest, you will learn about the liver's amazing regenerative response and fibrosis that occurs with chronic injury. You will hear about the diseases that result in acute and chronic types of hepatocellular and cholestatic injury and learn the means to distinguish them. Perhaps most importantly, you will learn about drug-induced liver disease, which is quite common and potentially deadly. In all, this lecture will provide you an amazing framework for characterization of all manner of liver disease. Goal: Be able to describe the major histology of common types of hepatitis, identify the salient histological features of the etiology-specific hepatitis, and associate the histology with the clinical outcome.



ChatGPT Prompt

"Map this lecture to USMLE Content Area keywords and to MeSH terms:"



Result

USMLE Content Area Keywords:

- Liver injury mechanisms
- Hepatic diseases
- Liver regeneration
- Fibrosis
- Acute hepatocellular injury
- Chronic hepatocellular injury...

MeSH terms:

- Liver injuries
- Liver diseases
- Liver regeneration
- Liver cirrhosis
- Hepatitis...

Example Use at NYU: Student Goals

Suggested ways to accomplish this goal:

1. Join NYU Grossman School of Medicine's academic clubs, interest groups or student chapters of professional associations related to psychiatry. This will provide opportunities to expand your knowledge, participate in organized discussions, and attend seminars or conferences.
2. Set up a weekly schedule to review journal articles related to psychiatry. Utilize resources such as the American Journal of Physical Medicine & Rehabilitation and Archives of Physical Medicine and Rehabilitation. Staying up-to-date on recent research findings will enhance your understanding of the field.
3. Read and analyze the Cuccurullo textbook, "Physical Medicine and Rehabilitation Board Review." Cover one chapter or section each week, taking notes and discussing key points with classmates or mentors to ensure a thorough understanding of the material.
4. Participate in clinical rotations or observe psychiatrists at Family Health Centers affiliated with NYU Langone Hospital – Brooklyn or other NYU-sponsored clinics. This hands-on experience will give you insight into the practice of psychiatry while expanding your knowledge.
5. Attend psychiatry-related webinars, conferences, and workshops offered by NYU Grossman School of Medicine, other medical schools, or professional associations. These opportunities will provide updates on the latest advances in the field, as well as the chance to network with professionals and fellow students with similar interests.

ChatGPT may produce inaccurate information about people, places, or facts

Save these suggestions as a note

Generate different suggestions

Rate this suggestion:  

Example Use at NYU: EHR Triggered Educational Nudges

NYU Grossman School of Medicine

Institute for Innovations in Medical Education

For: [Redacted]

Below are suggested personalized learning resources based on recent notes you wrote in Epic. These recommendations were automatically generated by an AI education system developed at NYU Grossman.

[View More >](#)

Recent Cases

42 YEAR-OLD FEMALE WITH A DIAGNOSIS OF ACUTE PANCREATITIS, UNSPECIFIED COMPLICATION STATUS, UNSPECIFIED PANCREATITIS TYPE

Amboss: *Acute pancreatitis*

[Open Amboss ->](#) [Related Q-Bank Questions ->](#)

Suggested Review Article
Acute Pancreatitis: A Review. JAMA. Jan 2021. [Read now](#)

- "Several scoring systems, such as BISAP and APACHE II, have good predictive capabilities for disease severity and mortality in acute pancreatitis."
- "Early and aggressive fluid resuscitation and early enteral nutrition are associated with lower rates of mortality and infectious complications."
- "Prompt diagnosis and stratification of severity influence proper management in acute pancreatitis."

Suggested Randomized Controlled Trial Article
Aggressive or Moderate Fluid Resuscitation in Acute Pancreatitis. N Engl J Med. Sep 2022. [Read now](#)

- Implications:** The study suggests that aggressive fluid resuscitation in acute pancreatitis may increase the risk of fluid overload without improving clinical outcomes.
- Safety Concerns:** The higher incidence of fluid overload in the aggressive resuscitation group raises safety concerns, which warrants further investigation into safe fluid management strategies in acute pancreatitis.
- Future Directions:** Research should focus on identifying effective, safe fluid management strategies in acute pancreatitis, potentially including individualized fluid resuscitation based on patient response and risk factors.

Institute for Innovations in Medical Education
EducationIT@nyulangone.org | med.nyu.edu

DxMentor uses Epic Data to tailor suggested educational nudges for our medical students and residents

- Attribution determined by authorship of H&P or Progress notes
- Learners emailed each morning with AI-generated nudges for their new cases
 - Amboss
 - Podcasts
 - Diagnostic Frameworks
 - PubMed Articles
- Learners have access to all prior cases and recommendations (patient log).

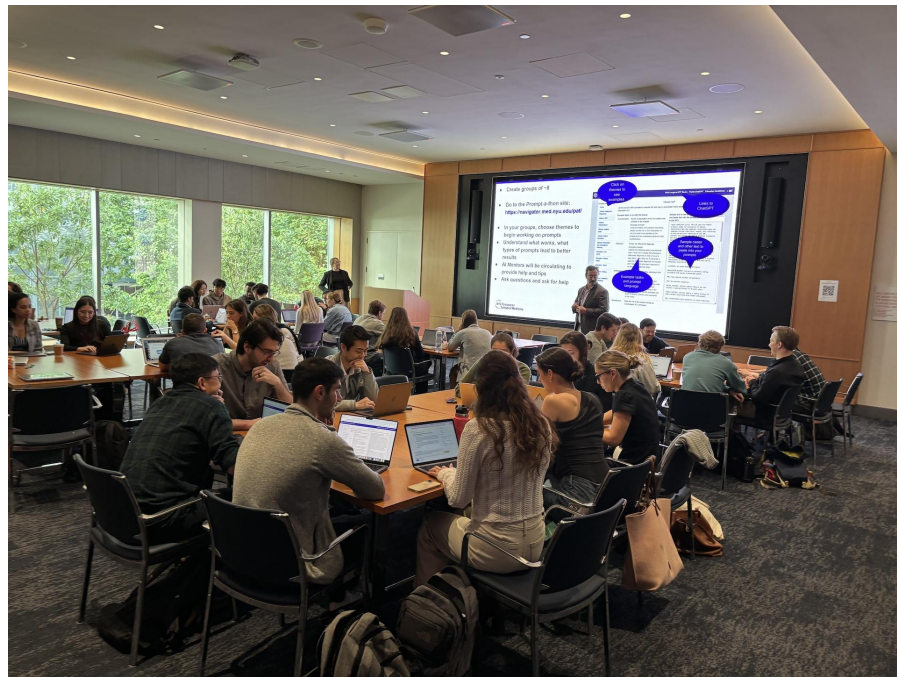
AI innovation at NYU Langone

- Formed GenAI Advisory Workgroup
- Established 1st Academic Medical Center HIPAA/PHI GPT environment
- Elicited 100+ ideas for initial use cases; Launched 10+ pilots in clinical, research and education
- Hosted world's 1st healthcare “Prompt-a-thon”
- Built NYUTron our own GenAI model



NYU Grossman School of Medicine + AI

- **AI and Generative AI now a required part of our core curriculum**
- **Issued guidelines on safe and responsible use of AI for our students and residents**
- **Host prompt-a-thons with hands-on AI skills training**
 - <https://navigator.med.nyu.edu/pat/>
- **AI electives**
- **Launched a new Masters program in Health Care AI for students, residents, and faculty**



What's Next?

Unique Challenges of GenAI: Power

MICROSOFT / TECH / SCIENCE

Microsoft wants Three Mile Island to fuel its AI power needs



Photo by Andrew Caballero-Reynolds / AFP via Getty Images

/ Microsoft has signed a 20-year deal to exclusively access 835 megawatts of energy from a nuclear plant.

By **Tom Warren**, a senior editor and author of *Notepad*, who has been covering all things Microsoft, PC, and tech for over 20 years.
Sep 20, 2024, 8:23 AM EDT

[Link](#) [Facebook](#) [Twitter](#) | 65 Comments (65 New)

[HOME](#) > [NEWS](#) > [THE ENERGY & SUSTAINABILITY CHANNEL](#)

AWS hiring for data center 'principal nuclear engineer' to evaluate SMRs and nuclear fuel strategy roadmaps

After the blockbuster deal to build campus next to a traditional nuclear power plant

- AI is estimated to consume 4% of global electricity by 2026, equivalent to the entire country of Japan
- Tech companies are spending > \$20 billion to explore nuclear power for AI data centers

Turbulent, Existentially Difficult, but Ultimately Better for Patients

- Humans judgement and decision-making is influenced by myriad cognitive biases, limited by working memory. The data available on patients is vast and far exceeds the capacity of even the best physicians to use it to make precision medicine decisions. It's hubris to think we can practice in the same way.
- AI is less susceptible to common factors that lead doctors to make diagnostic errors: fatigue, lack of time and cognitive bandwidth when treating many patients, gaps of knowledge and reliance on mental shortcuts.
- AI is not perfect, but can be improved quickly and widely. Fixing biases or insufficient underlying data is “easy” in AI and difficult to impossible in humans.
- What it means to be a physician will change, for some dramatically.

Medical Education and AI

- Your students and trainees are and will use these tools
- AI tools are ubiquitous - so much so that people do not realize they are using them
- AI-generated text is indistinguishable and not detectable
- Create a policy for responsible use
- Provide a safe AI space and encourage experimentation
- Be transparent about how you are using AI and with which learner data
- Be growth-minded - AI can solve many age-old problems and unlocks incredible new possibilities.

Emerging Questions For Medical Education

- How do we train students for a future that will include AI agents and even autonomous AI?
- How do we tackle the risk of 'never-skilling' or 'de-skilling' when AI does the work?
- Changing scope of practice
- What happens when the AI is consistently better than humans at certain tasks?
- Etiquette - when is it appropriate to use these tools?

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/>
- High yield X/Twitter users to follow:
 - <https://twitter.com/emollick>
 - <https://twitter.com/EricTopol>
 - https://twitter.com/nejm_ai
 - <https://twitter.com/AdamRodmanMD>
 - <https://twitter.com/ylecun>

