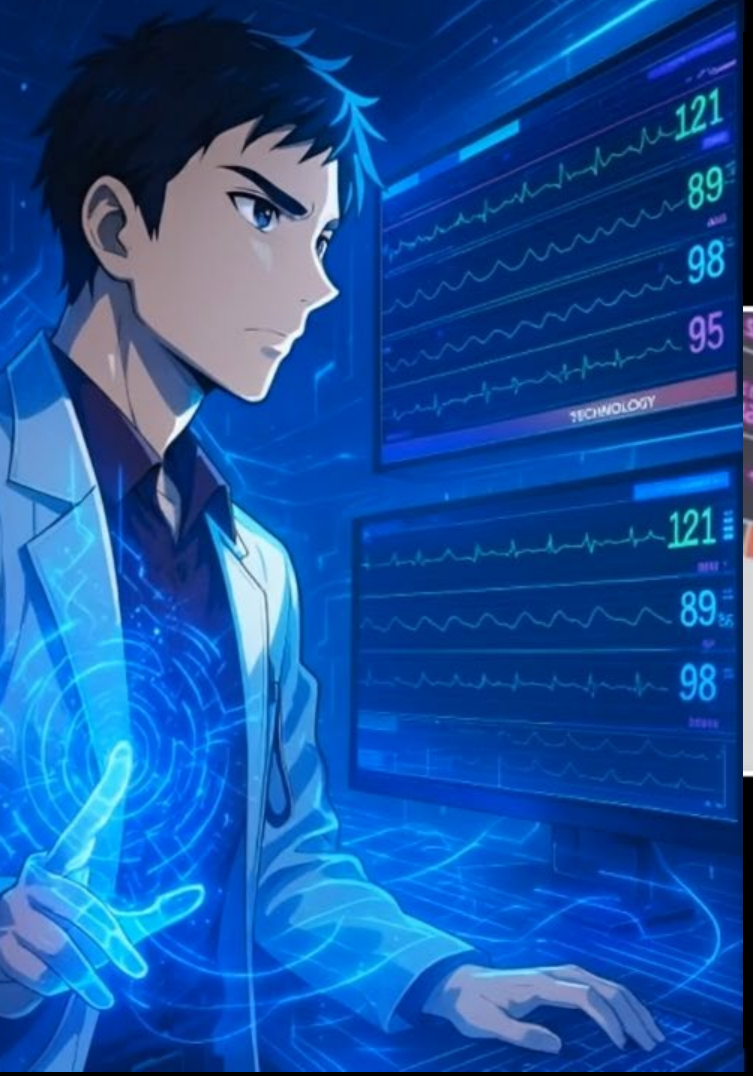
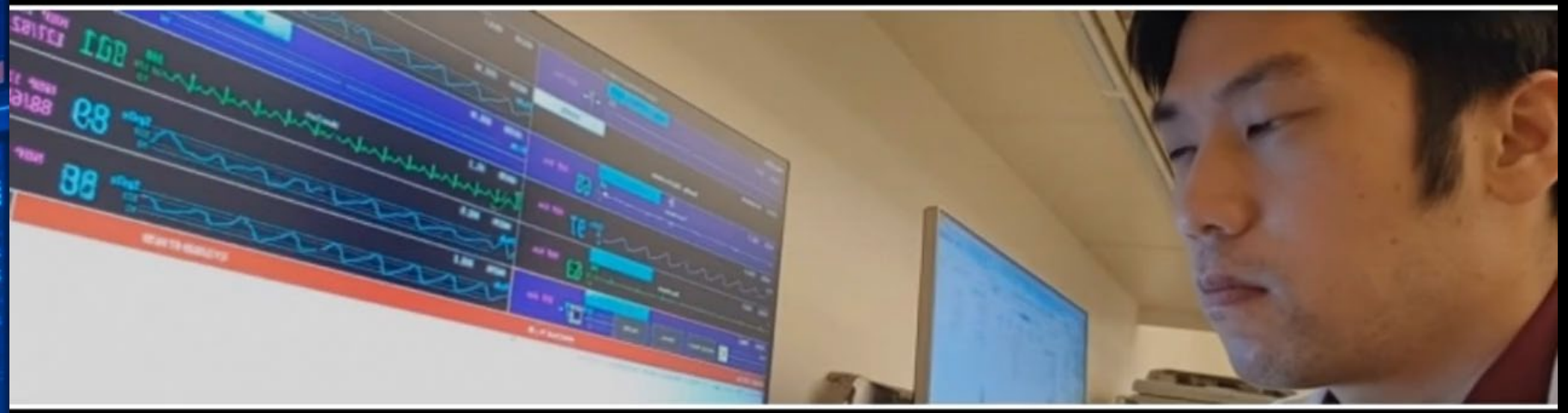


AI in Medicine



Real Magic or Illusion for Serious Illness?

Veo

Jonathan H. Chen MD, PhD @jonc101x
Director for Medical Education in Artificial Intelligence
Stanford Division of Computational Medicine
+ Division of Hospital Medicine + Clinical Excellence Research Center +
Biomedical Data Science + ARISE Healthcare Network Founding Lead

Deepfake Intro





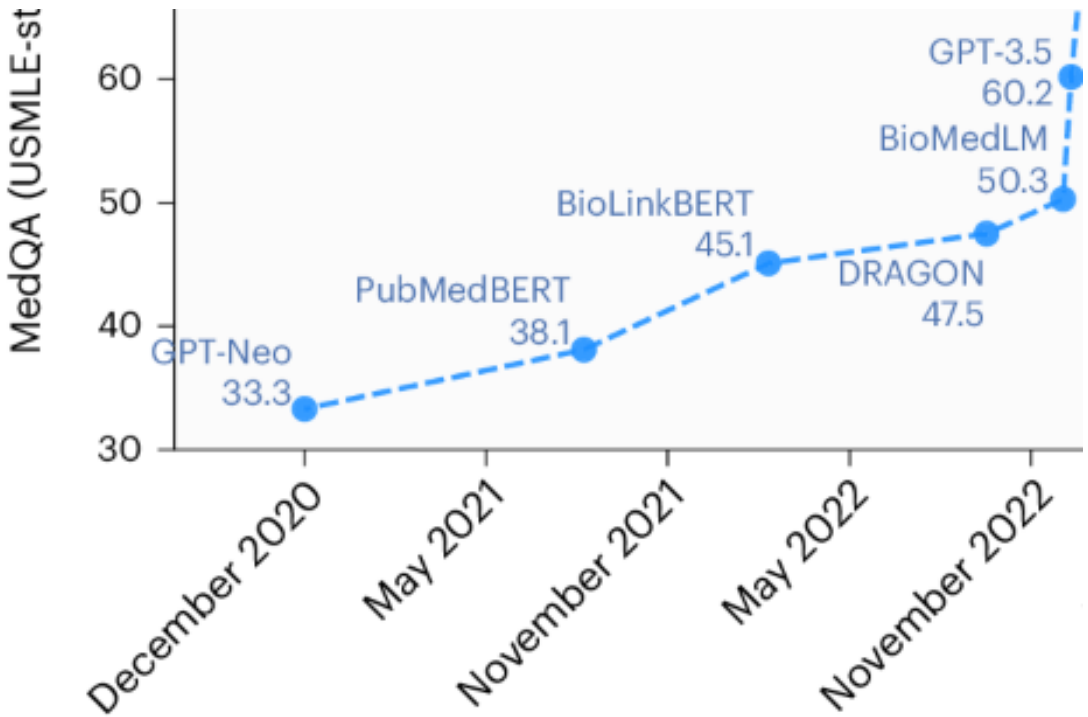
Veo

Acknowledgements

- Students + Collaborators
- ARPA-H PACT (Physician-AI Collaborative Teaming)
- NIH/National Institute of Allergy and Infectious Diseases (1R01AI178121-01)
- NIH-NCATS-Clinical & Translational Science Award (UM1TR004921)
- Stanford Bio-X Interdisciplinary Initiatives Program
- Josiah Macy Jr. Foundation (AI in Medical Education)
- Stanford RAISE Health Seed Grant 2024
- NIH - Center for Undiagnosed Diseases (U01 NS134358)
- GrowTherapy Sponsored Research Award

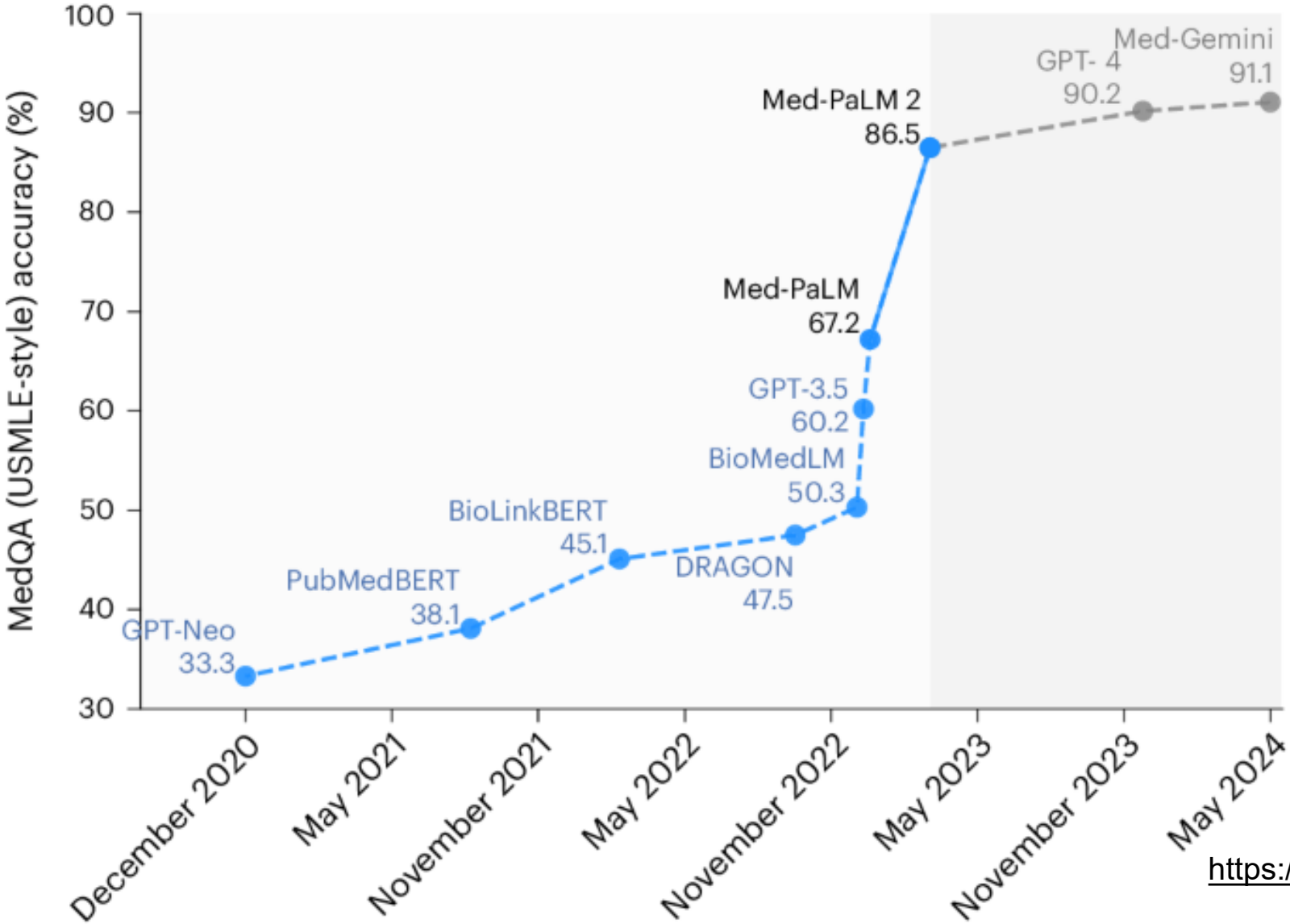
Content does not necessarily represent the views of any agency or institution
Financial disclosures: Reaction Explorer LLC – Co-founder. Expert witness / Consulting / Speaking fees from Elite Experts, Pfizer, insitro, Gen Re, AASCIF, R1RCM

Advancing Medical Q&A Accuracy



https://www.nature.com/articles/s41591-024-03423-7

Advancing Medical Q&A Accuracy



https://www.nature.com/articles/s41591-024-03423-7

Chatbot vs. Medical Student Clinical Reasoning

PROMPT 1:

A 72 y/o man with a history of hypertension and smoking presented with shortness of breath for two weeks.

The patient was in his usual state of health until about two weeks ago when he noted the gradual onset of shortness of breath. It was initially present only with moderate exertion such as climbing multiple flights of stairs, but now occurs while he walks across the room. He does not feel short of breath at rest, and reports no orthopnea, paroxysmal nocturnal dyspnea, lower extremity swelling, cough, chest pain, or palpitations. He reports that he thinks his weight has remained unchanged during this period, but that he doesn't routinely weigh himself.

Of note, six months ago the patient was hospitalized for four days with COVID, but he thought he had fully recovered from that. The patient is now concerned he may have "long COVID".

He has a history of hypertension and osteoarthritis of the knees.

His outpatient medications include losartan, amlodipine, and occasional ibuprofen.

His immunizations are up to date, including influenza and COVID.

He has no known allergies to medications, food, or environmental factors

His father died at 65 from emphysema. His mother died at 80 from heart disease. The patient also has a 68 year old brother with coronary artery disease and diabetes.

The patient lives with his wife and his brother-in-law in a private home in San Jose. His brother-in-law has Down's syndrome, and they have been caring for him since his wife's parents passed away about five years ago. The patient is an active smoker, currently 1 pack/day. He drinks "a few" beers or mixed alcoholic drinks per week. He denies any history of illicit drug use. He is a retired plumber and suspects he's had occupational exposure to various chemicals over the years, but doesn't know any details. He has not been sexually active for several years. He describes his diet as "pretty bad, I guess. A lot of meat and salt, and probably too many sweets." His last travel out of the Bay Area was a trip to visit his children in Southern California shortly before the COVID pandemic. He has no pets or animal exposures.

Taking all the information you have into account, what are the three most likely diagnosis for this patient's primary complaint of shortness of breath?

Example Tasks

- Summarize the case
- Describe a diagnostic schema and illness script
- Write a prioritized problem list
- Explain key factors that favor possible diagnoses

GRADING RUBRIC 1 (4 points total):

Inclusion of any of the following diagnoses are justifiable to include in a list of the top 4 possibilities, and should receive 1 point each. Additional diagnoses not listed may be justifiable and should receive credit at grader's discretion. Students may list variations of diagnoses as separate items (e.g. multiple forms of ILD, multiple forms of pneumonia, etc…); graders should use their discretion to determine if the variations are sufficiently clinically distinct to receive full credit.

- Pneumonia
- COVID
- Post-acute COVID syndrome (a.k.a. "Long COVID")
- Interstitial lung disease
- Pulmonary embolism
- Pulmonary hypertension
- Heart failure
- CAD / unstable angina
- Valvular heart disease (e.g. aortic stenosis)
- COPD
- Lung cancer
- Lymphoma
- Other cancer metastatic to lung
- Anemia (with long list of possible etiologies)

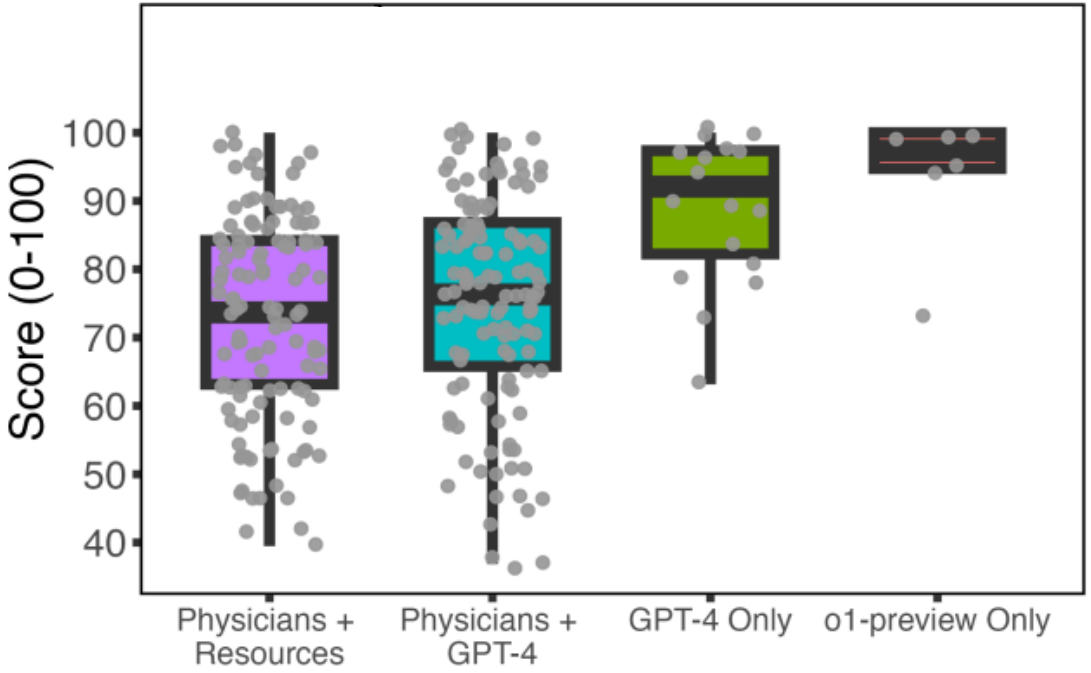
Eric Strong, MD[1]; Alicia DiGiammarino, MS[2]; Yingjie Weng, MHS[3]; Andre Kumar, MD, MEd[1]; Poonam Hosamani, MD[1]; Jason Hom, MD[1]; Jonathan H. Chen, MD, PhD[1,4,5]

Author Affiliations | Article Information

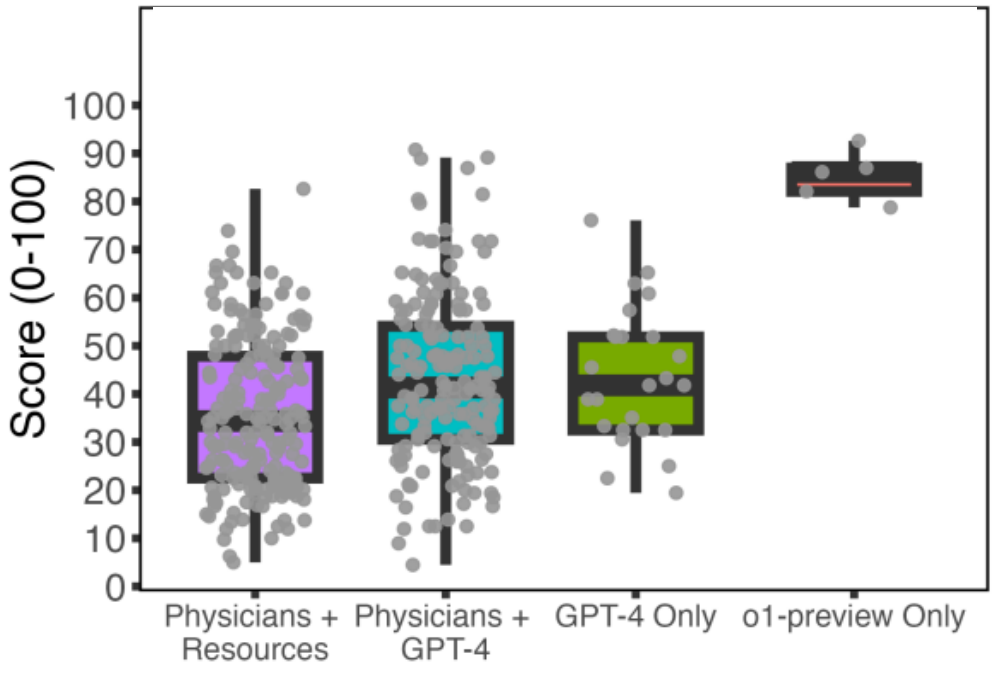
JAMA Intern Med. Published online July 17, 2023. doi:10.1001/jamainternmed.2023.2909

Superhuman Clinical Reasoning?

Diagnostic Reasoning Cases

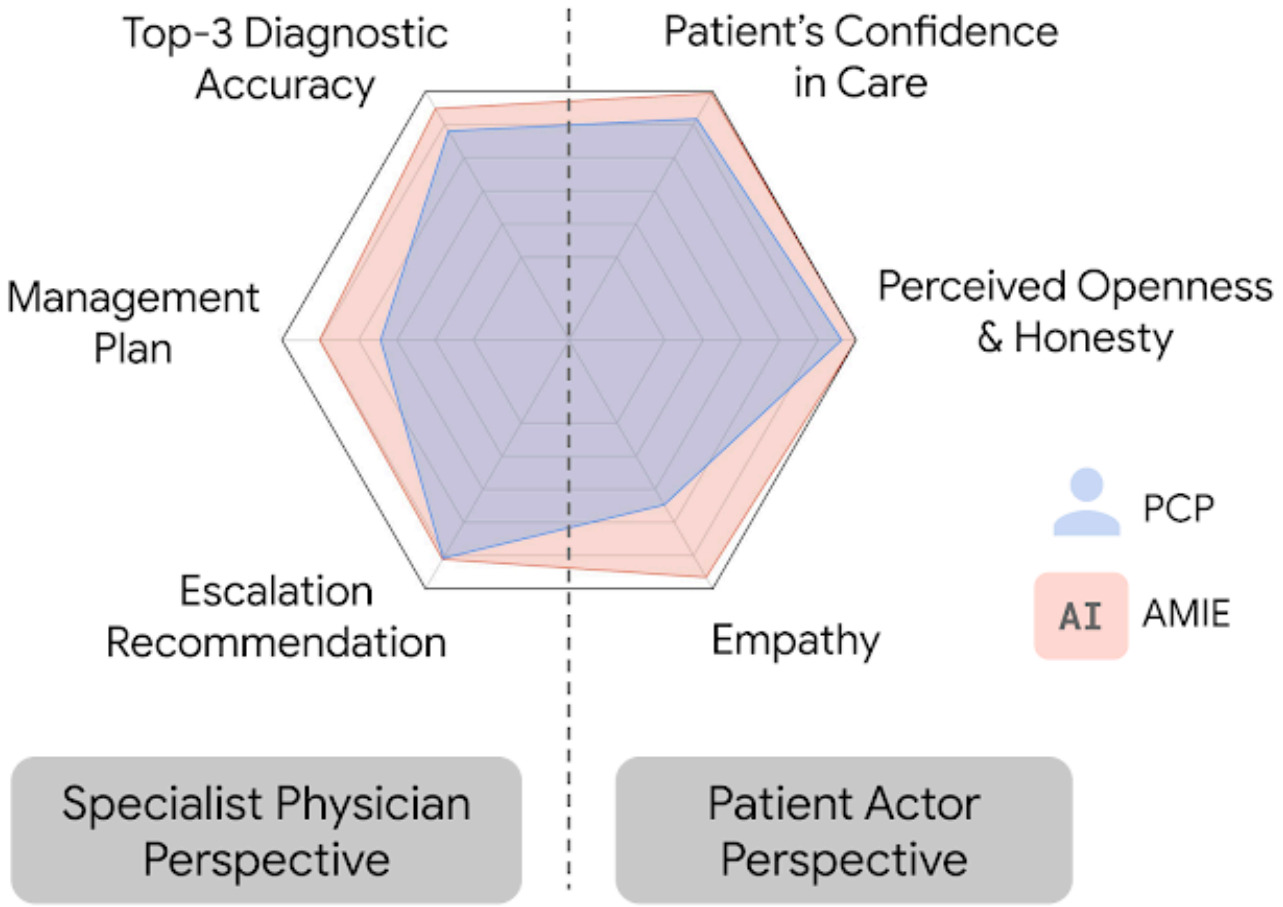


Management Reasoning Cases



Large Language Model Influence on Diagnostic Reasoning: A Randomized Clinical Trial, *JAMA Network Open* (2024)
GPT-4 assistance for improvement of physician performance on patient care tasks: A randomized controlled trial, *Nature Medicine* (2025)
Performance of a large language model on the reasoning tasks of a physician, *Science* (2026)

Bot vs. Human Empathy?



https://research.google/blog/amie-a-research-ai-system-for-diagnostic-medical-reasoning-and-conversations/

https://www.nytimes.com/2023/06/12/health/doctors-chatgpt-artificial-intelligence.html

FIERCE Healthcare

Providers | Health Tech | Finance | Payers | Regulatory | Special Reports | Podcasts

AI AND MACHINE LEARNING

Chatbots outperformed doctors in answering patient questions with accuracy and empathy: JAMA study

By **Annie Burky** • May 1, 2023 01:57pm

JAMA Internal Medicine | Natural Language Processing | Artificial Intelligence | Google



ChatGPT's responses to patient questions outperform doctors in both empathy and quality, according to a new study. (Getty)

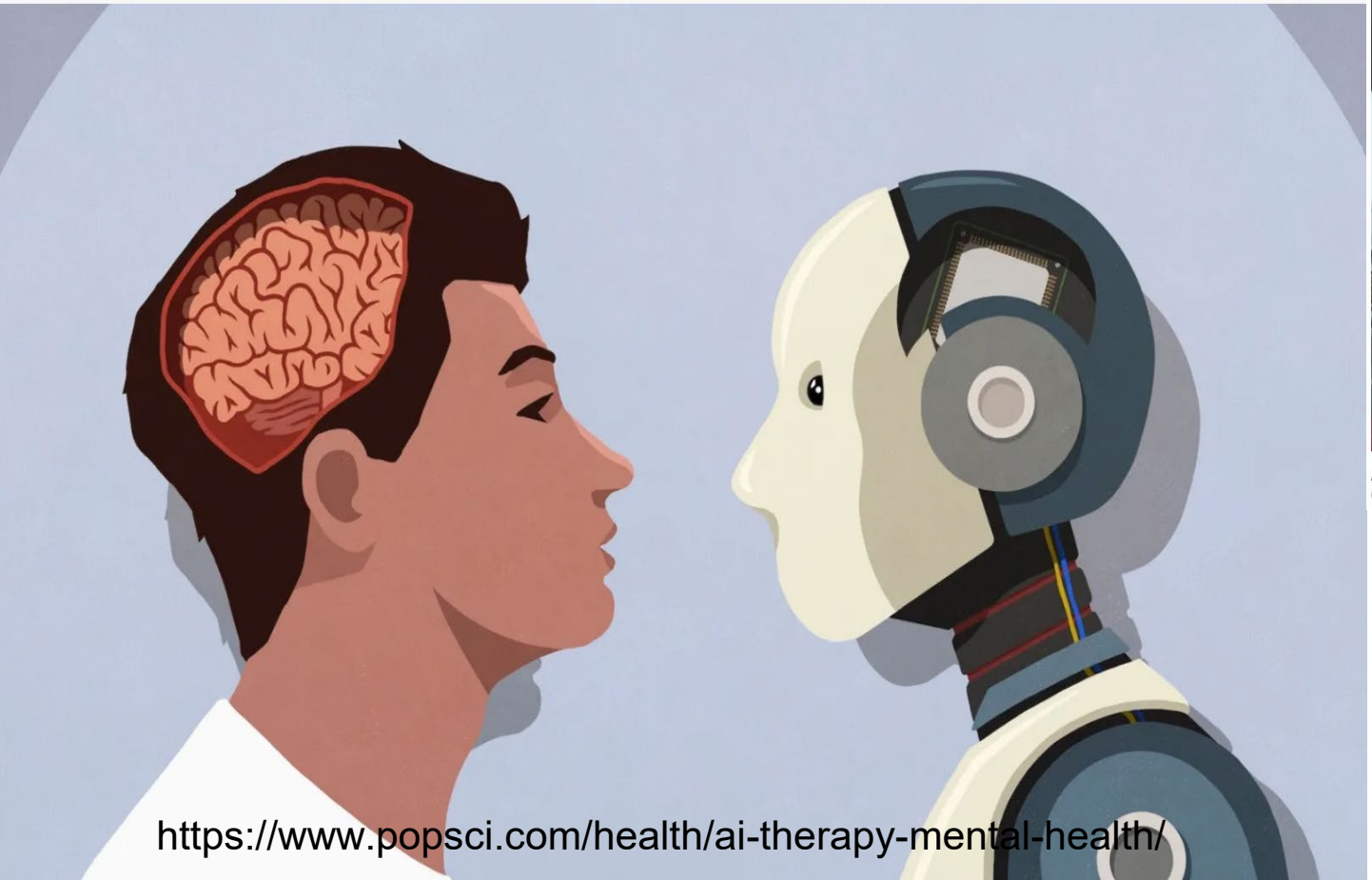
https://jamanetwork.com/journals/jamainternalmedicine/article-abstract/2804309

POPULAR SCIENCE

Illinois' ban on AI therapy won't stop people from asking chatbots for help

Americans are turning to AI for mental health support; sometimes with dangerous consequences.

MACK DEGEURIN / AUG 6, 2025 4:06 PM EDT



https://www.popsci.com/health/ai-therapy-mental-health/

AP

WORLD U.S. POLITICS SPORTS ENTERTAINMENT BUSINESS SCIENCE FACT CHECK ODDITIES MORE

2,400 Kaiser mental health professionals strike in Northern California over AI concerns



About 2,400 Kaiser Permanente mental health workers are striking in Northern California over fears the company will replace therapists with artificial intelligence. Kaiser says AI will not replace human assessments. (AP Video/Terry Chea)

Half of people would could benefit from therapy

Don't have access



AI, TECHNOLOGY & INNOVATION | ISSUE 3 / 2023

Who's training whom?

A physician's surprising encounter with ChatGPT

By Jonathan H. Chen
Illustration by Harry Campbell
November 10, 2023

https://ai.nejm.org/doi/pdf/10.1056/AIp2400006

https://stanmed.stanford.edu/surprising-chatgpt-revelation/

High-Stakes Conversations

“Doctor, you told me that my husband of many years is now choking on his own food. That's why he has pneumonia now. I know he's been having worsening dementia over the years. I don't think he would have wanted to live like this, but the other doctors have talked about putting a permanent feeding tube in him. If we don't do that, we'd be leaving him to starve to death. What choice is there?”

Chatbot: I understand this is a really hard time for you. Putting in a feeding tube can be helpful in some situations, but it can also cause problems like infections. With your husband's dementia, he might not even feel hunger the way we do. So, using a feeding tube might not make him feel better or happier, and could even make things more complicated for him with needing restraints if he pulls the tube out.

High-Stakes Conversations

“I can't give up on him. He would never give up on me. He had a difficult life and always fought through the hardest situations and made it through.”

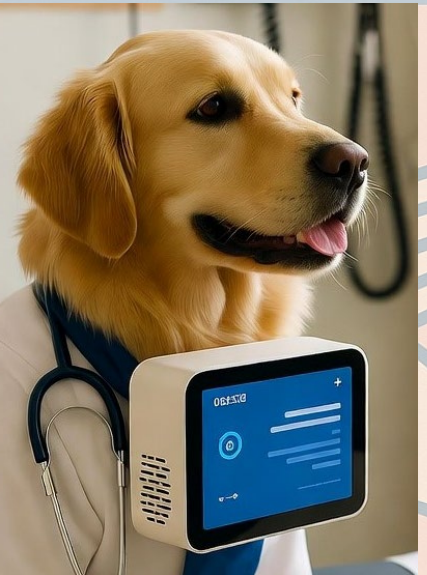
Chatbot: Your feelings are totally understandable. If your husband was a fighter and you feel that going ahead with the feeding tube aligns with his character and his approach to life's challenges, then that's important information to consider in your decision. Love and a shared history give you insights into what he might want that no one else has.

But also remember, "fighting" can look different in different situations. Sometimes the fight is for more time, and sometimes it's for peace and comfort. Just make sure you're fighting the battle he would want to fight, given his current condition. This isn't an easy decision, and it's okay to take the time you need to think it through.

What Makes a Clinician?

Knowledge?

Empathy?

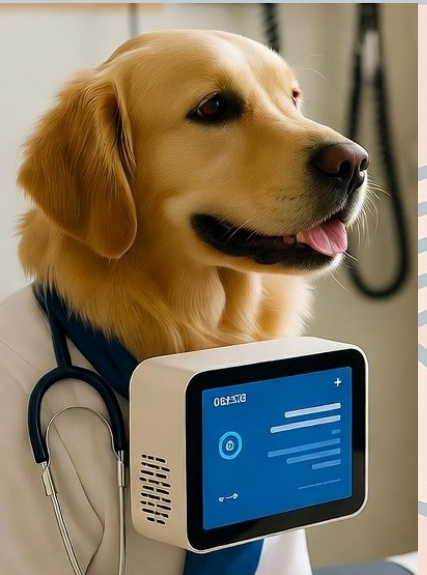


Liability?

What Makes a Clinician?

Competence
Knowledge → Judgment

Communication
Empathy → Influence



Why are humans still in the loop with advancing AI capabilities?
BMJ Digital Health and AI (2026)

Character
Liability → Responsibility

Acknowledgements

- Students + Collaborators
- ARPA-H PACT (Physician-AI Collaborative Teaming)
- NIH/National Institute of Allergy and Infectious Diseases (1R01AI178121-01)
- NIH-NCATS-Clinical & Translational Science Award (UM1TR004921)
- Stanford Bio-X Interdisciplinary Initiatives Program
- Josiah Macy Jr. Foundation (AI in Medical Education)
- Stanford RAISE Health Seed Grant 2024
- NIH - Center for Undiagnosed Diseases (U01 NS134358)
- GrowTherapy Sponsored Research Award



Content does not necessarily represent the views of any agency or institution

Financial disclosures: Reaction Explorer LLC – Co-founder. Expert witness / Consulting / Speaking fees from Elite Experts, Pfizer, insitro, Gen Re, AASCIF, R1RCM

Self-Directed

- ChatGPT, Gemini, etc.
  - Role Play / Generate Images
- OpenEvidence / DoxGPT
  - Ask Questions, Lookup Articles
- NotebookLM
  - Generate Summary/Podcast
- Claude Cowork
  - Automate Tasks
- Approved Tools:
  - https://uit.stanford.edu/ai/

Discussion Forum

- Updated Competencies?
  - Skills not replaceable by AI?
  - Skills to deprioritize?
- AI for Clinical Use Policy?
  - Who is responsible for use and missed opportunities?
- AI Direct to Patients?
  - Trust vs. Access?
- Core Curriculum:
  - https://stanfordmed.github.io/aimeded-website/learn