Department of Neurological Surgery
Department of Radiology
NYU Institute for Neuroscience
NYU Center for Data Science
Global AI Frontier Lab

Off Label Intelligence: Books, Language Models, and AI As A Medical Device

Eric K. Oermann, M.D.
Associate Professor of Neurosurgery, Radiology, and Data Science
Associate Program Director, Neurosurgery Residency Program
Director of the OLAB, Health AI Research Lab at NYU Langone Health
Investigator, Neuroscience Institute, NYU
Principal Investigator, Global AI Frontier Lab
Chair of Data Science, Congress of Neurological Surgeons
Associate Editor for Data Science, Neurosurgery

Congress of Neurological Surgeons



Disclosure Statement of Financial Interest

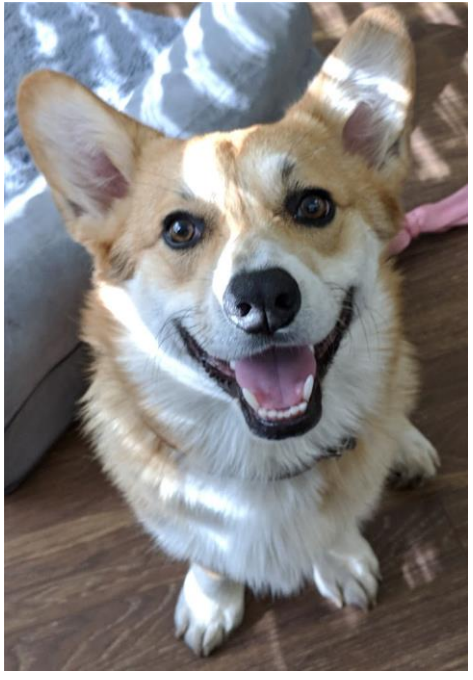
Within the past 12 months, I or my spouse/partner have had a financial interest/arrangement or affiliation with the organization(s) listed below.

Grant/research support: NIH, Keck Foundation, Ministry of Science and ICT (Republic of Korea)

Consultant: Google Inc., Sofinnova Inc., AlphaTec

Equity: Artisight Inc., MarchAI, Delvi

Employment: Eikon Therapeutics (wife)



NYU Langone Health

MODEL EVIDENCE

SYSTEM EVIDENCE

PATIENT-BENEFIT EVIDENCE

STAGE 1

1. BENCH CAPABILITY

Can the model perform the task?

REAL-WORLD SAMPLED DATA

Vishwanath et al., 2026



GREATER CLINICAL GROUNDING

SYNTHETIC DATA

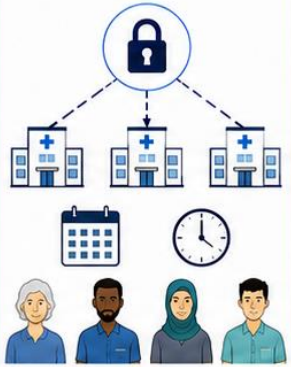
Arora et al., 2025



STAGE 2

2. EXTERNAL VALIDATION

Does performance travel across sites, time, and subgroups?



Zech et al., 2018 • Wong et al., 2021

STAGE 3

3. CLINICAL-TASK EVALUATION

Does it help clinicians perform the task—including in simulated environments?

Reader studies • simulated cases • virtual workflows



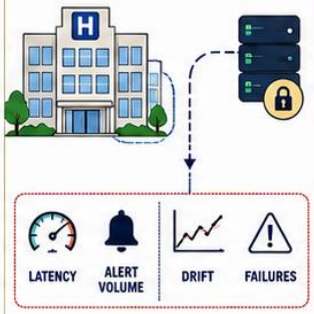
Tschandl et al., 2020 • Goh et al., 2024

STAGE 4

4. STEALTH OBSERVATIONAL STUDY

(SILENT DEPLOYMENT)

Will it operate safely here before influencing care?



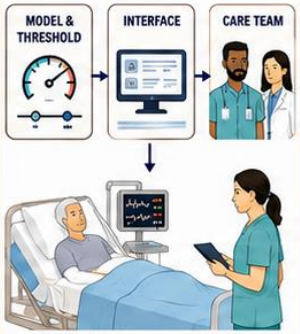
Kwong et al., 2022 • Corbin et al., 2023

STAGE 5

5. COHORT TRIAL

RETROSPECTIVE OR PROSPECTIVE • REAL DEPLOYMENT

Does the deployed model–workflow bundle change care?

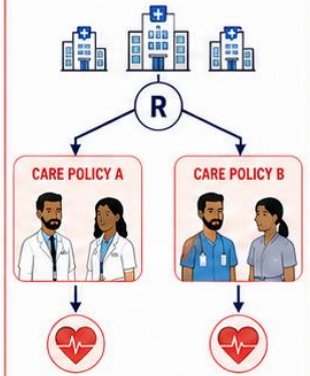


Adams et al., 2022 • Nahass et al., 2026

STAGE 6

6. PRAGMATIC RANDOMIZED TRIAL

Does the care policy improve patient outcomes?



Lin et al., 2024 • Agweyu et al., 2026

VERSION-AWARE SURVEILLANCE

MODEL BEHAVIOR → **OPERATIONAL TRIGGER** → **CLINICIAN ATTENTION** → **CLINICAL ACTION** → **PATIENT OUTCOME**

CLINICAL BENEFIT = MODEL × THRESHOLD × INTERFACE × WORKFLOW × RESPONSE CAPACITY

First, a reality check on the present state of AI in medicine

NYU Langone Health

Past or Present?

0148-396X/90/2705-0837$02.00/0
NEUROSURGERY
Copyright © 1990 by the Congress of Neurological Surgeons

Vol. 27, No. 5, 1990
Printed in U.S.A.

Technical note

Downloaded from http://journals.lww.com/neurosurgery by BhDMf5ePHKbH4TTImqenWAiCYEPWHoBkJebZ+E52crX3VqfZbRe1S0C3eRN4Qi4aO3+O4Mw1ed4GQ== on 12/04/2019

Computer-Aided Intelligence: Application of an Expert System to Brachial Plexus Injuries

Wink S. Fisher III, M.D.

Department of Neurosurgery, Naval Hospital Bethesda, Bethesda, Maryland

When confronted with a patient with a brachial plexus injury, how often as neurosurgeons do we consult an atlas to confirm the anatomy of the brachial plexus and then attempt to establish the location of the lesion? Similar difficulties are encountered with lumbar and sacral plexus lesions. In a project organized to assist the neurosurgeon in this time-consuming task, a computer program that can rapidly determine the site of a lesion in a brachial, lumbar, or sacral plexus injury was created. Using known anatomical pathways (37 clinically relevant upper and 20 lower extremity muscle innervations), and relying solely upon the neurological motor examination, rapid computer-assisted diagnosis is possible. When more than one final common pathway lesion occurs (for example, multiple root avulsions of the brachial plexus), possible lesion sites can be obtained. An interactive dialogue between the user and the program helps to determine the location of the lesion. The program can be run on any IBM-compatible personal computer and is presented as an instrument that provides assistance in cases of complex peripheral nerve injuries, when expert consultants are unavailable. In addition, it can be used as an aid to learning and as a review of basic neuroanatomy. (*Neurosurgery* 27:837–843, 1990)

Key words: Artificial intelligence, Brachial plexus, Computer application, Expert system, Nerve injury

NYU Langone Health

Past or Present?

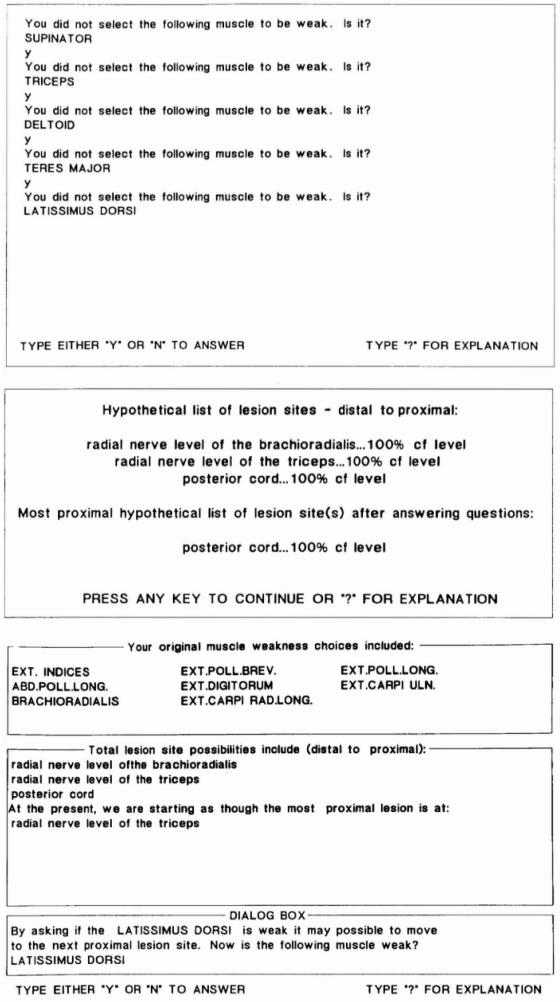
- **Computer assisted diagnosis**
- **Using a Chatbot**
- **Interpretable feedback**
- **… focus on augmented intelligence**

If it said “ChatGPT” this paper could have been written **yesterday**.

It describes ***all of the current trends*** in modern **Medical AI research**.

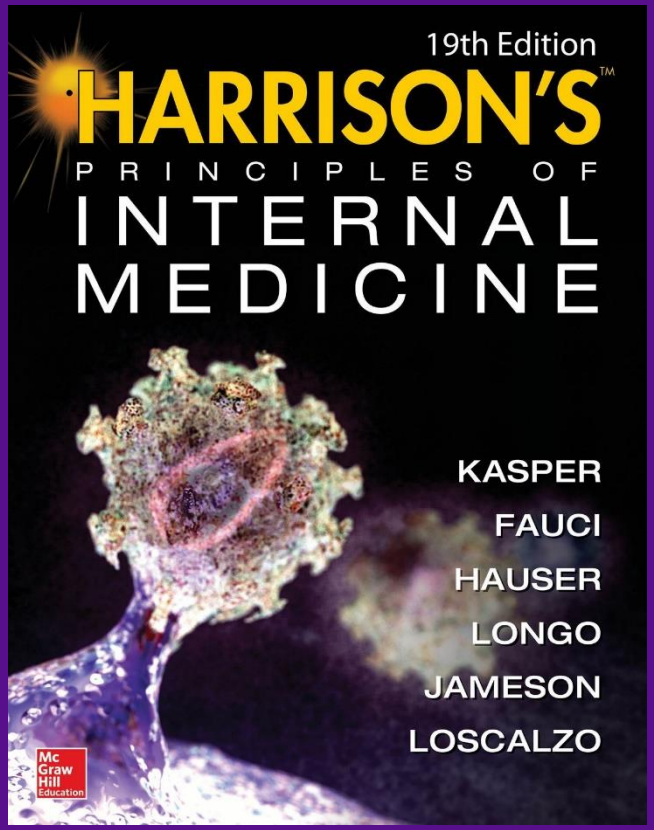
***Physicians have been doing medical AI with chatbots for over 30 years*.**

NYU Langone Health



NYU Langone Health

The original knowledge tool in medicine



A published textbook. A predicting LLM.

BOOK · STATIC AT USE

- Same words for every reader
- Expert-curated; edition + citations visible
- Revised in discrete cycles

LLM · STOCHASTIC AT USE

- Sampled from prompt, context + model version
- Mega-corpus; provenance often diffuse
- Shifts with checkpoint, retrieval + deployment





NYU Langone Health

Defining clinically relevant hypotheses

NYU Langone Health



Sully Chen

nature medicine

Large language models in clinical medicine: an LLM-powered large-scale systematic review

Sully F. Chen, Anton Alyakin, Andreas Seas, Eunice Yang, Joanne J. Choi, Jin Vivian Lee, Amelia Chen, Pranav Warman, Rochelle Bitolas, Robert J. Steele, Daniel Alber, Eric K. Oermann

The reality of our field (artificial intelligence)…

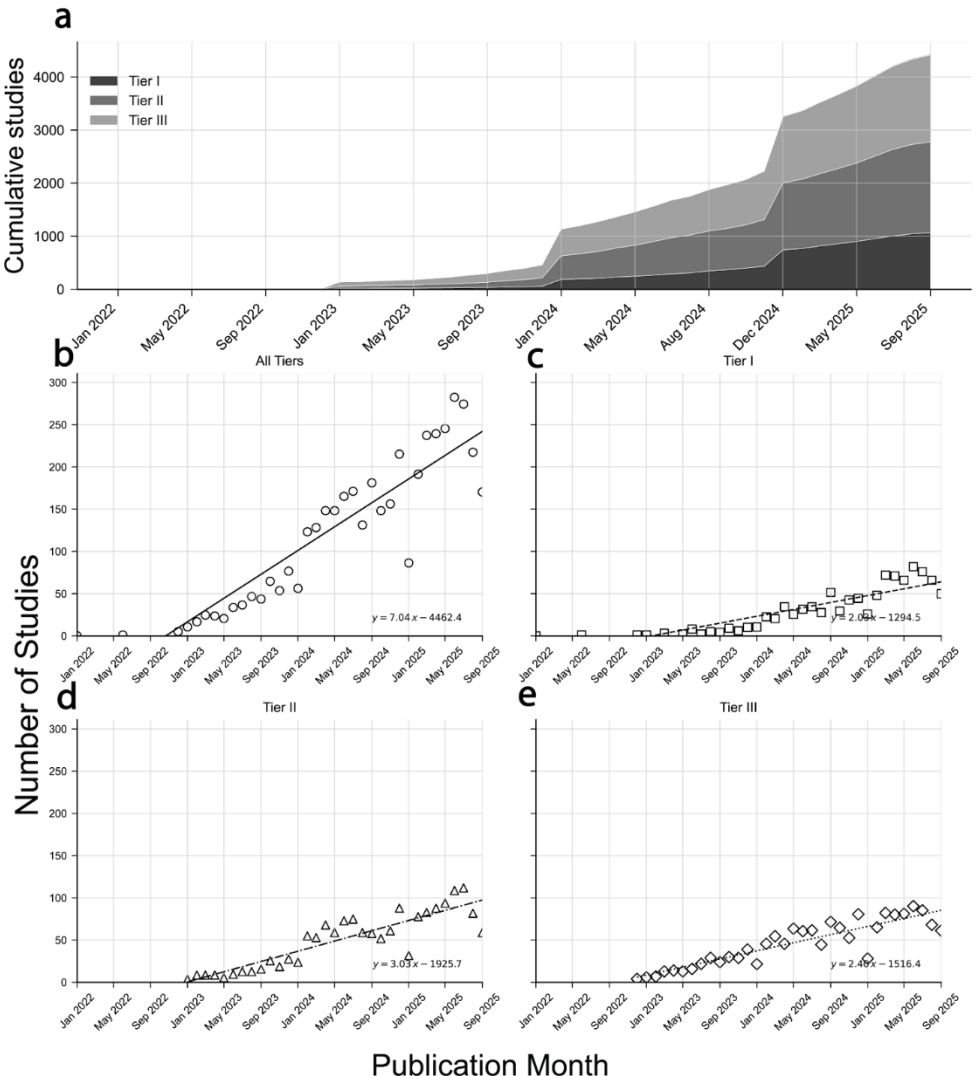
Since 2020, ~5,000 papers published on *just* LLMs in Medicine

Majority test on multiple choice medical exams

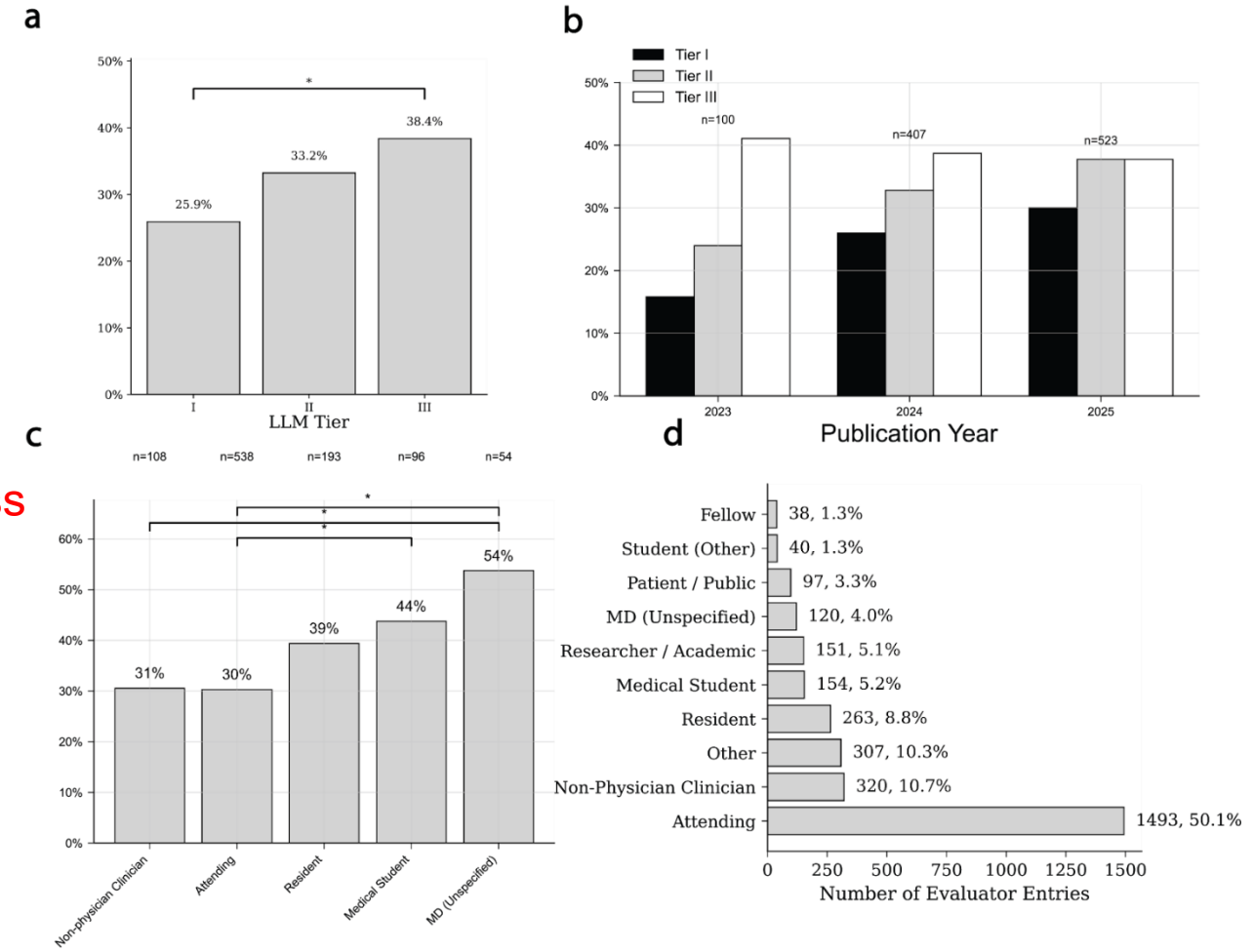
21 *studies* generated actual medical evidence

NYU Langone Health

4 papers are published *per day*



LLM's are less impressive when compared against attendings



NYU Langone Health



Shrutika Singh

scientific reports

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › scientific reports › articles › article

Article | Open access | Published: 26 November 2025

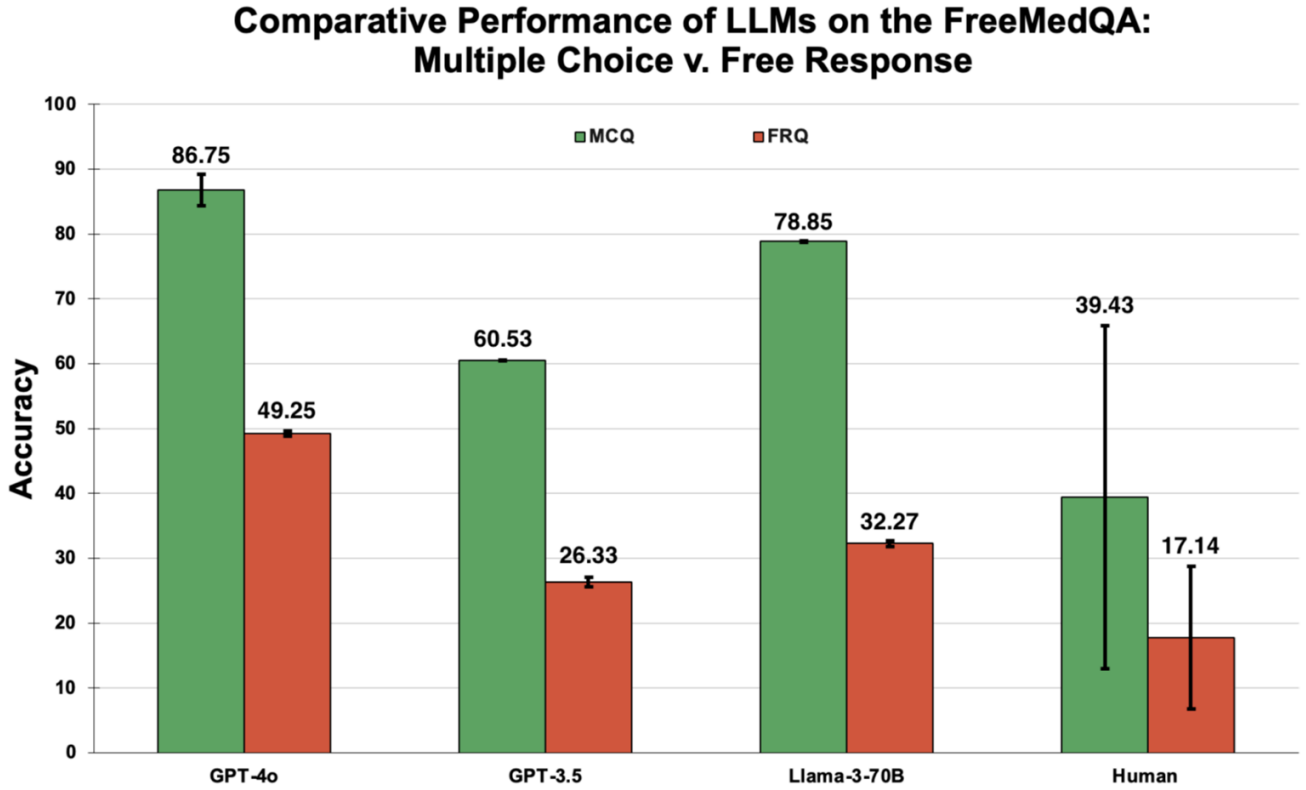
The pitfalls of multiple-choice questions in generative AI and medical education

Shrutika Singh ✉, Anton Alyakin, Daniel Alexander Alber, Jaden Stryker, Ai Phuong S. Tong, Karl Sangwon, Nicolas Goff, Mathew De La Paz, Miguel Hernandez-Rovira, Ki Yun Park, Eric Claude Leuthardt & Eric Karl Oermann ✉

Scientific Reports **15**, Article number: 42096 (2025) | Cite this article

699 Accesses | Metrics

Does it even make sense to use medical exams in modern era?

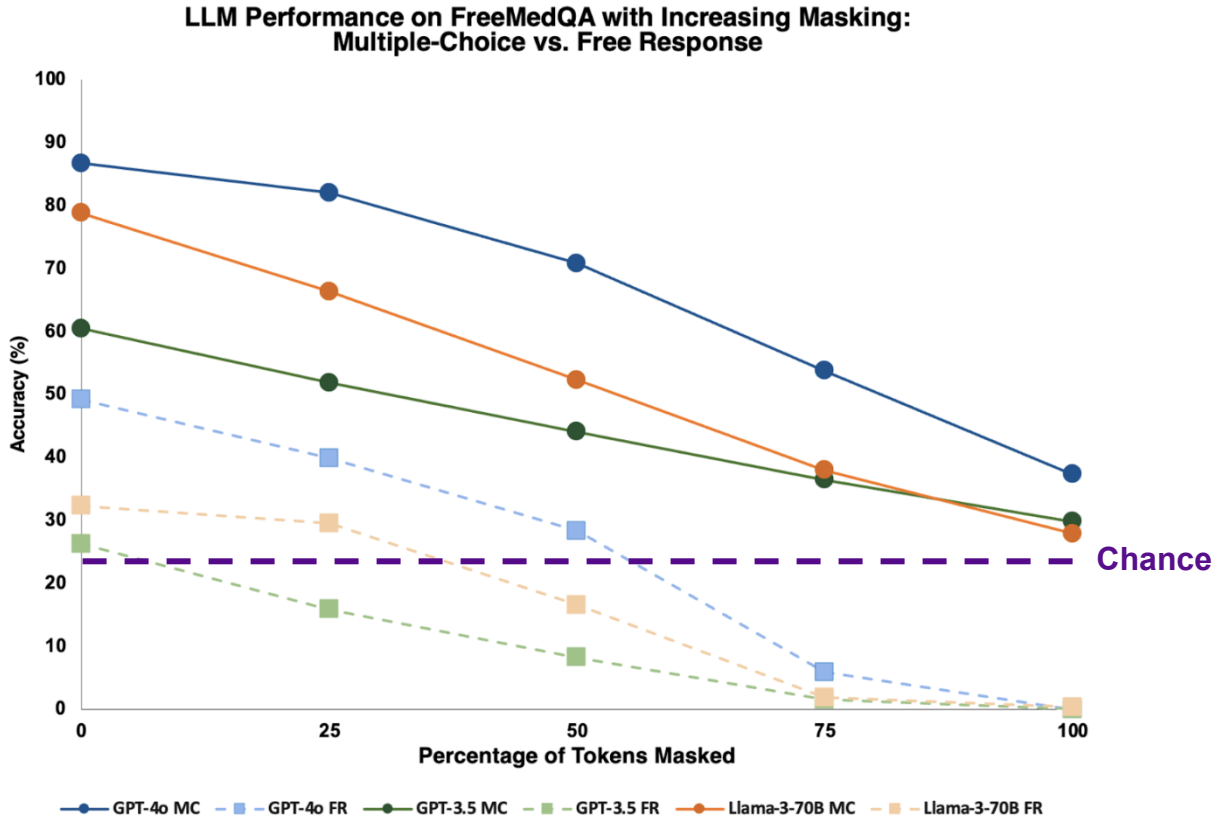


NYU Langone Health

Does it even make sense to use medical exams in modern era?

(1) Memorized in training data?

(2) Exploiting prior in answer distribution?



Base Rate Fallacy in Prior Distribution

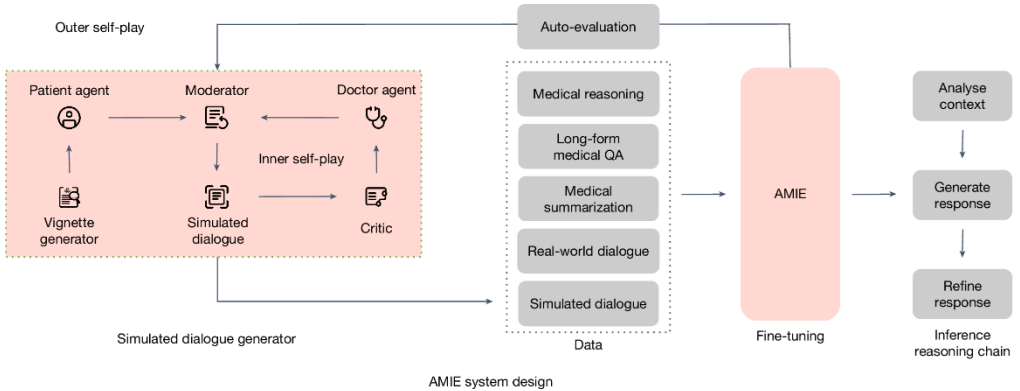
Article

Towards conversational diagnostic artificial intelligence

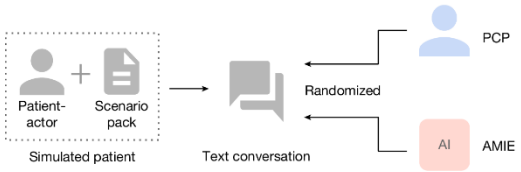
https://doi.org/10.1038/s41586-025-08866-7
Received: 18 January 2024
Accepted: 5 March 2025
Published online: 9 April 2025
Open access

Tao Tu[1,3✉], Mike Schaekermann[1,3✉], Anil Palepu[1,3], Khaled Saab[1], Jan Freyberg[1], Ryutaro Tanno[2], Amy Wang[1], Brenna Li[1], Mohamed Amin[1], Yong Cheng[2], Elahe Vedadi[1], Nenad Tomasev[2], Shekoofeh Azizi[2], Karan Singhal[1], Le Hou[1], Albert Webson[2], Kavita Kulkarni[1], S. Sara Mahdavi[2], Christopher Semturs[1], Juraj Gottweis[1], Joelle Barral[2], Katherine Chou[1], Greg S. Corrado[1], Yossi Matias[1], Alan Karthikesalingam[1,4✉] & Vivek Natarajan[1,4✉]

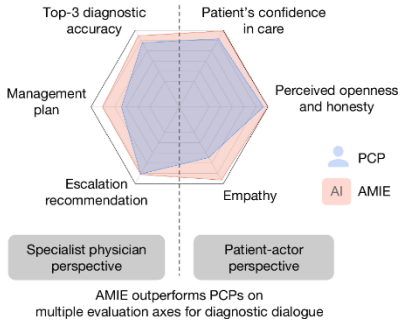
- **More real world – FRQ with multi-turn dialogue…**
- **Is this an AI physician?**
  - Chatbot actually *chats with patients* to arrive at a diagnosis!
- **Compared to *real doctors* using *professional actors*.**



AMIE system design



Randomized study design for remote OSCE



AMIE outperforms PCPs on multiple evaluation axes for diagnostic dialogue

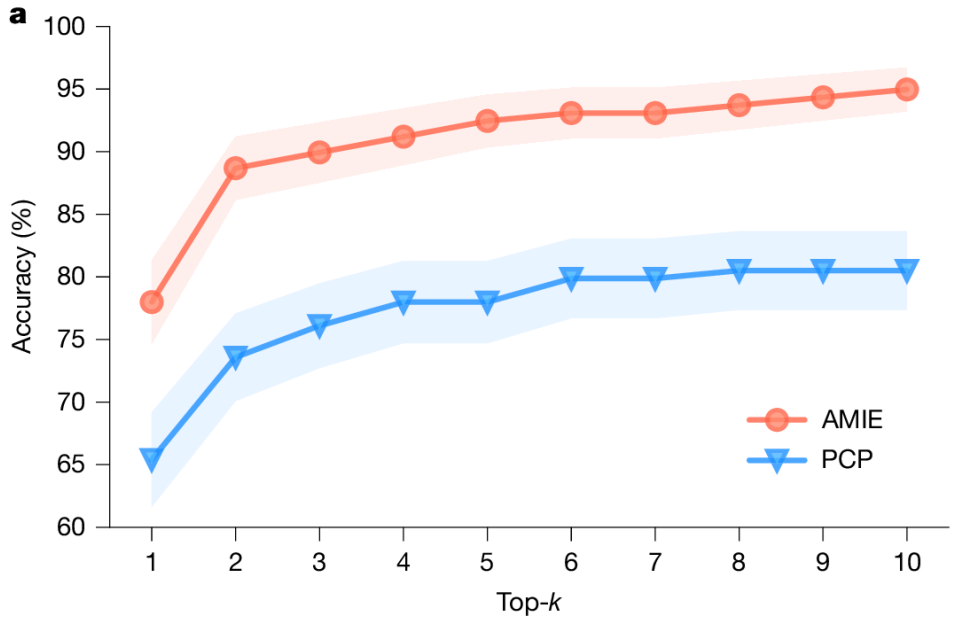
NYU Langone Health

Base Rate Fallacy in Prior Distribution

But what about the distribution of priors?

- **Better than doctors…**
- **But really?**
- **Massive prior shift b/c *not really being used in production* and the *synthetic problem setup is wrong*.**
  - **159 simulated case scenarios (OSCE)**
    - **149 positive cases**
    - **10 negative cases**

p(disease_test) = 0.94 !!!!



NYU Langone Health

Base Rate Fallacy in Prior Distribution

- **Neurological: most common dx in primary care are headaches, back pain, and sleep disorders.**
  - Eval dataset distributions:
    - Low back pain: 0%
    - Headache disorders: 13%
    - GCA: 6.6%
  - Real world distributions:
    - Low back pain: 28%
    - Headache disorders: 16%
    - GCA: 0.01%

Is this a model that does great on rare diseases?

Or a model that will have an unacceptable false positive rate in the real world?

NYU Langone Health

Table SI.10 | Neurology scenarios (N=30).

Location	Ground truth
Canada	Alzheimer's dementia
	Cluster headache
	Delirium tremans
	Essential Tremor
	Giant cell arteritis on background of polymyalgia rheumatica
	Hemorrhagic stroke
	Meningitis
	Migraine
	Migraine
	Multiple sclerosis with optic neuritis
	Optic neuritis which can be associated with multiple sclerosis
	Parkinson's disease
	Space occupying lesion (brain tumour)
	Temporal arteritis +/- Polymyalgia rheumatica
	Trigeminal neuralgia
India	Carpal tunnel syndrome
	Cavernous sinus thrombosis
	Dermatomyositis
	Dystonia
	Lead neuropathy
	Migraine
	Myasthenia gravis
	Neurocysticercosis
	Optic neuritis
	Parkinson's disease
	Posterior reversible encephalopathy syndrome (PRES)
	Tabes dorsalis
	Transient ischemic attack
	Trigeminal Neuralgia
	Wernicke's encephalopathy

Independent Assessment

NYU Langone Health

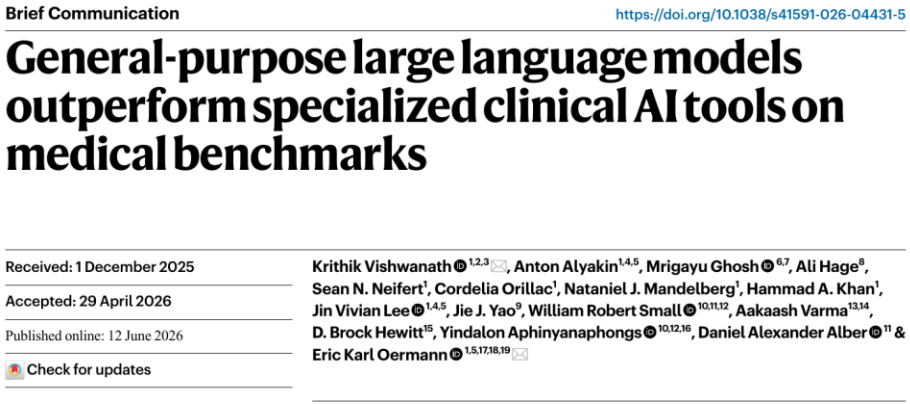
A benchmark tests a *model*.

A clinical study tests a *medical tool*.

NYU Langone Health

Real World Queries…

- Two clinical AI tools already in your practice — OpenEvidence and UpToDate Expert AI
- Against three frontier models: GPT-5.2, Gemini 3.1 Pro, Claude Opus 4.6
- **Independent.**



Brief Communication

https://doi.org/10.1038/s41591-026-04431-5

General-purpose large language models outperform specialized clinical AI tools on medical benchmarks

Received: 1 December 2025

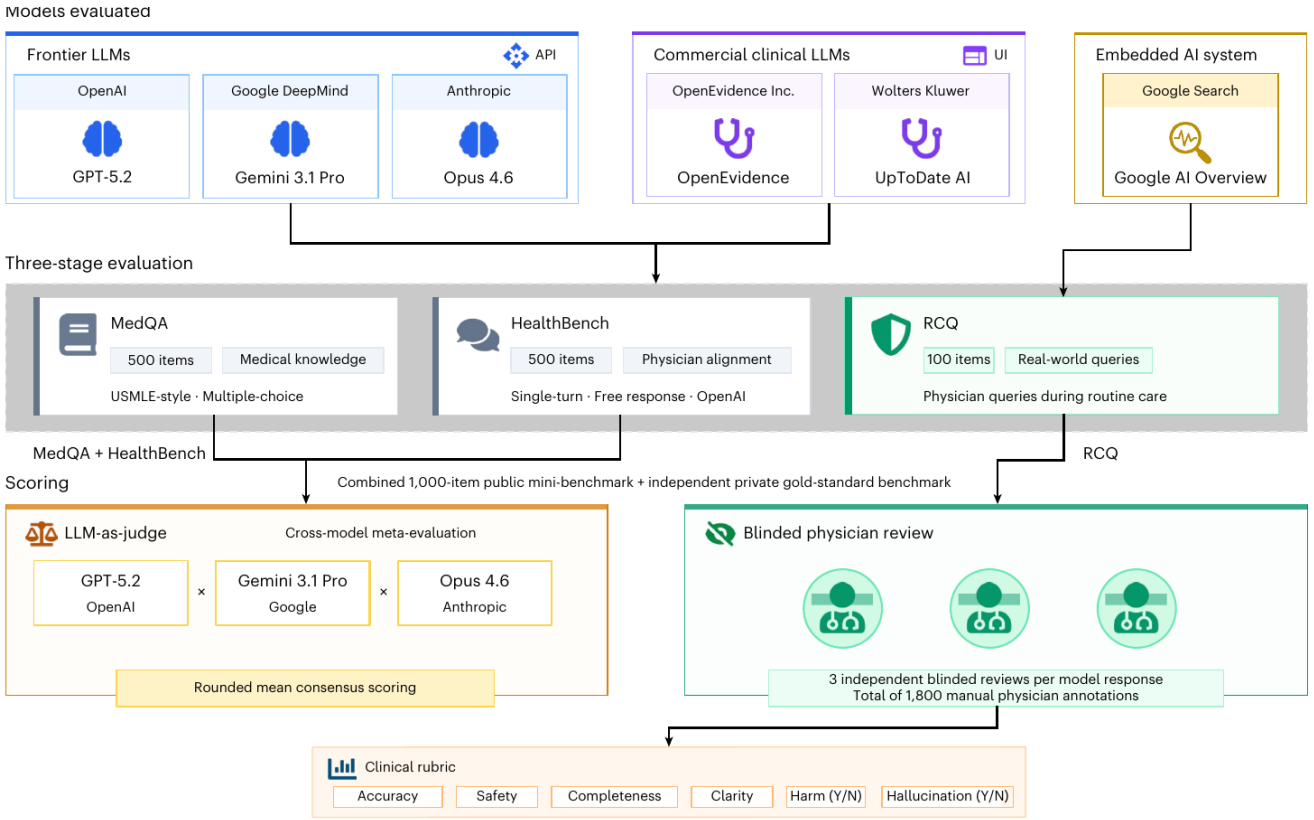
Accepted: 29 April 2026

Published online: 12 June 2026

Check for updates

Krithik Vishwanath[1,2,3], Anton Alyakin[1,4,5], Mrigayu Ghosh[6,7], Ali Hage[8], Sean N. Neifert[1], Cordelia Orillac[1], Nataniel J. Mandelberg[1], Hammad A. Khan[1], Jin Vivian Lee[1,4,5], Jie J. Yao[9], William Robert Small[10,11,12], Aakaash Varma[13,14], D. Brock Hewitt[15], Yindalon Aphinyanaphongs[10,12,16], Daniel Alexander Alber[11] & Eric Karl Oermann[1,5,17,18,19]

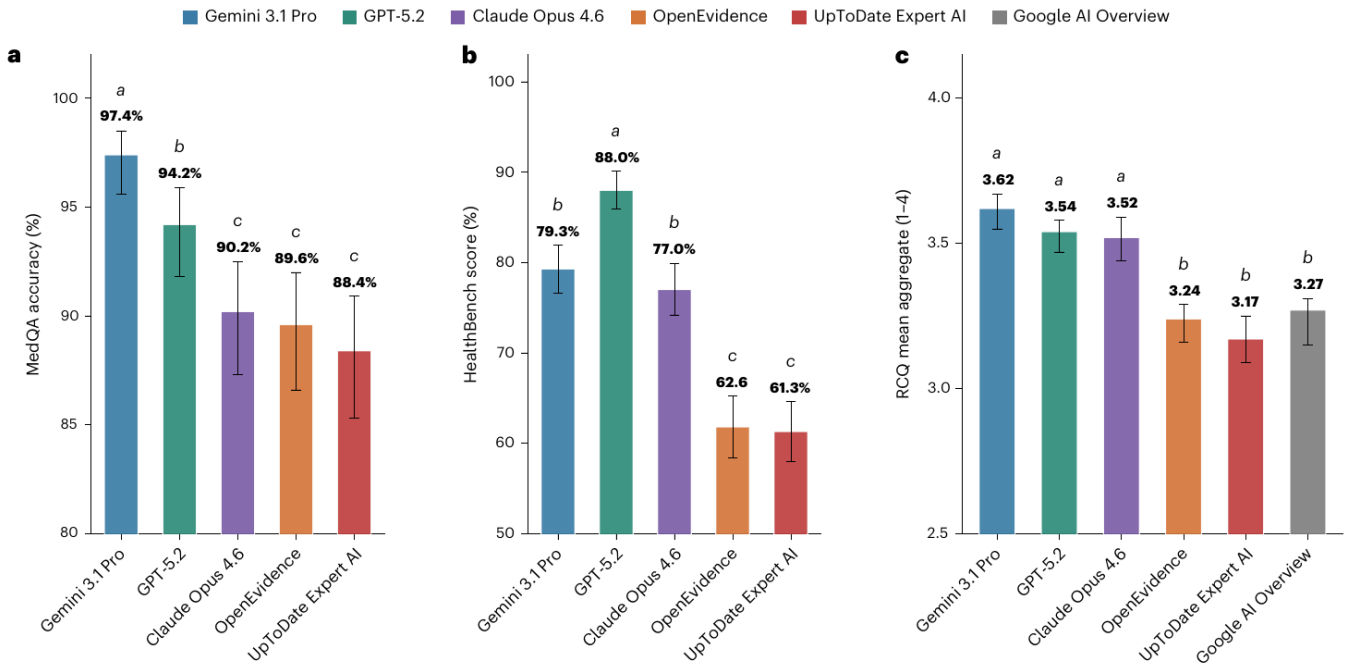
Three stages, each one harder to game



The stage that counts: real clinical queries (RCQ)

- 100 de-identified queries physicians actually sent our HIPAA-compliant instance
- Six systems answered each one; 12 US clinicians scored them blinded to model identity
- Three independent raters per response — 1,800 annotations, 1,704 after refusals
- Four dimensions plus binary harm and hallucination flags; refusals verified by hand
- Friedman → Nemenyi, Holm–Bonferroni throughout, cumulative link models with rater fixed effects

Two tiers — the clinical tools are in the second



What does this really say?

- Frontier models took the top tier on all three stages; clinical tools took the second
- **Physicians at an academic medical center in Manhattan from Neurosurgery, Dermatology, Orthopedics, and Internal Medicine at the inpatient point of care rated frontier models as being superior.**
- Distributions matter…
- It's likely that some models are better for certain tasks and certain populations

MODEL EVIDENCE

SYSTEM EVIDENCE

PATIENT-BENEFIT EVIDENCE

STAGE 1

1. BENCH CAPABILITY

Can the model perform the task?

REAL-WORLD SAMPLED DATA

Vishwanath et al., 2026



GREATER CLINICAL GROUNDING

SYNTHETIC DATA

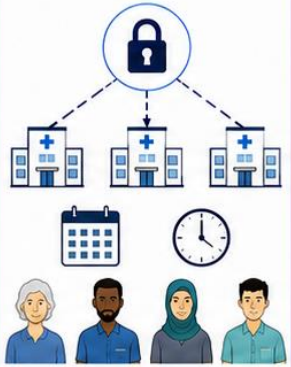
Arora et al., 2025



STAGE 2

2. EXTERNAL VALIDATION

Does performance travel across sites, time, and subgroups?



Zech et al., 2018 • Wong et al., 2021

STAGE 3

3. CLINICAL-TASK EVALUATION

Does it help clinicians perform the task—including in simulated environments?

Reader studies • simulated cases • virtual workflows



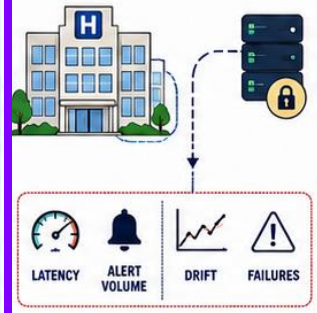
Tschandl et al., 2020 • Goh et al., 2024

STAGE 4

4. STEALTH OBSERVATIONAL STUDY

(SILENT DEPLOYMENT)

Will it operate safely here before influencing care?



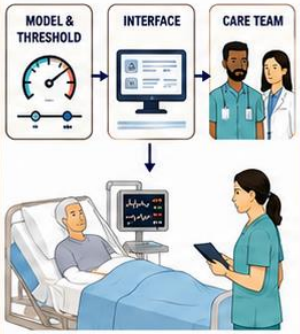
Kwong et al., 2022 • Corbin et al., 2023

STAGE 5

5. COHORT TRIAL

RETROSPECTIVE OR PROSPECTIVE • REAL DEPLOYMENT

Does the deployed model–workflow bundle change care?

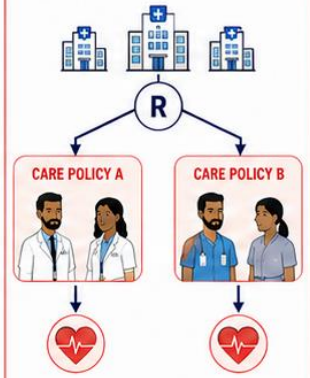


Adams et al., 2022 • Nahass et al., 2026

STAGE 6

6. PRAGMATIC RANDOMIZED TRIAL

Does the care policy improve patient outcomes?



Lin et al., 2024 • Agweyu et al., 2026

VERSION-AWARE SURVEILLANCE

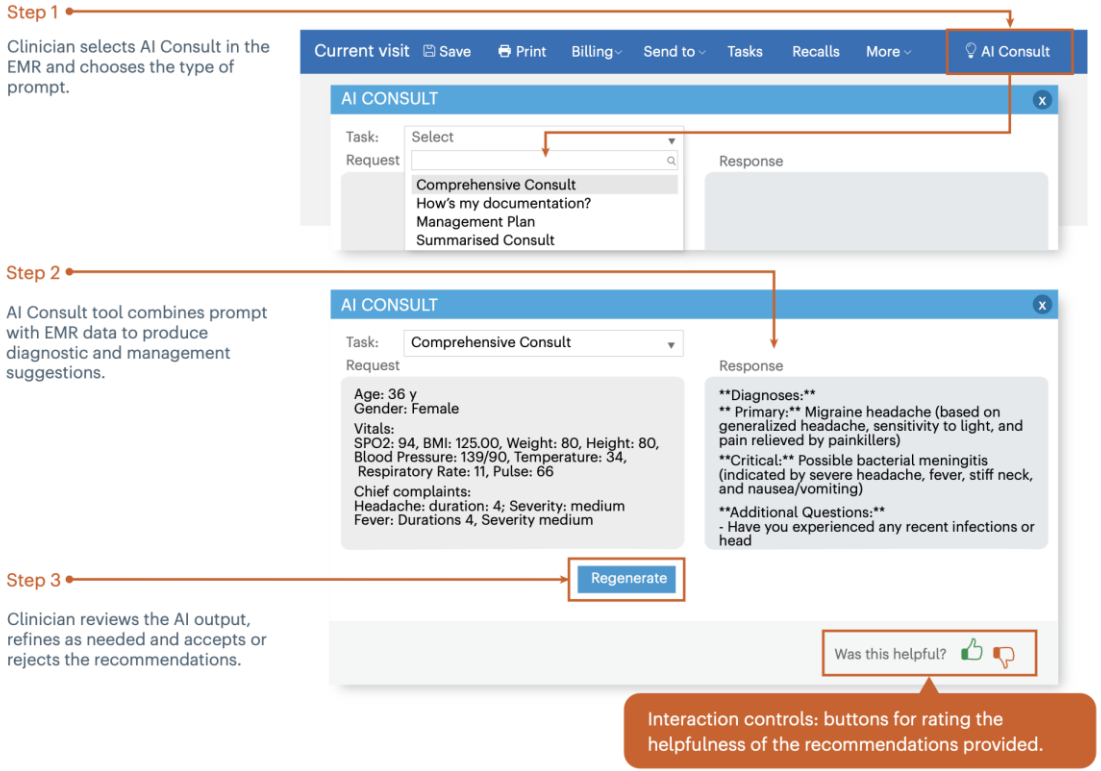
MODEL BEHAVIOR → OPERATIONAL TRIGGER → CLINICIAN ATTENTION → CLINICAL ACTION → PATIENT OUTCOME

CLINICAL BENEFIT = MODEL × THRESHOLD × INTERFACE × WORKFLOW × RESPONSE CAPACITY

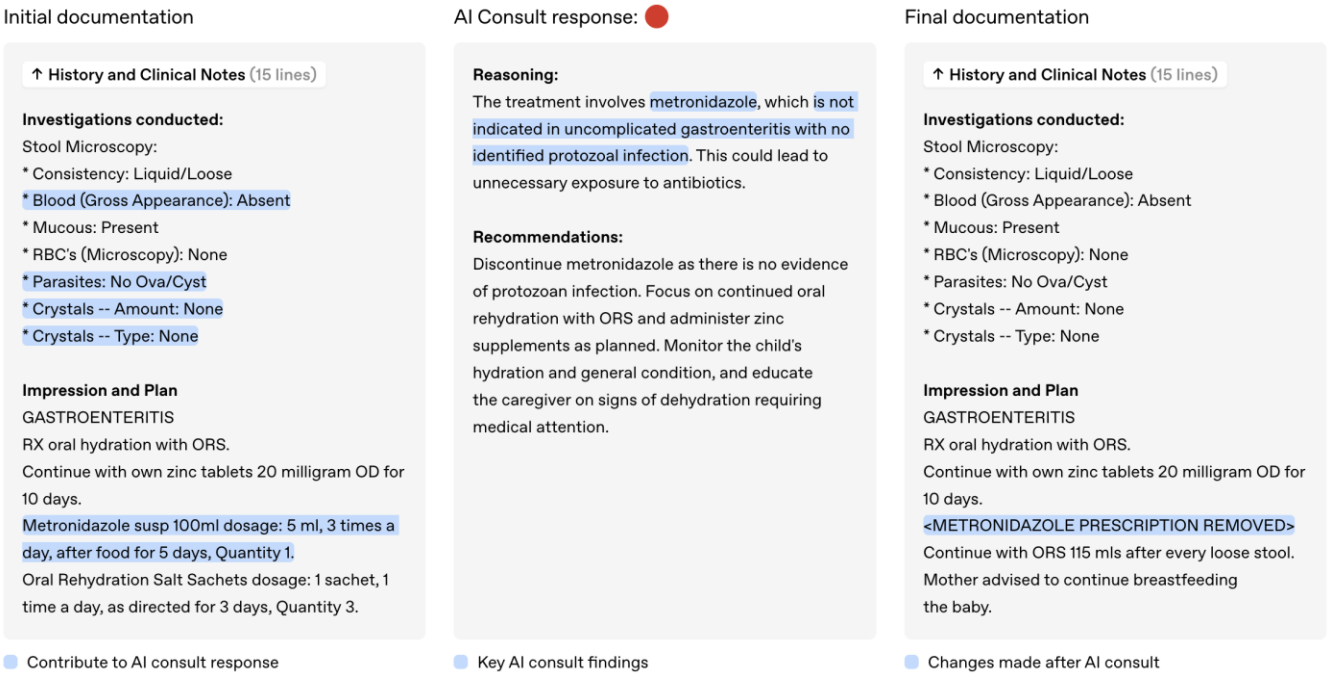
CASE 1 · PENDA AI

NYU Langone Health

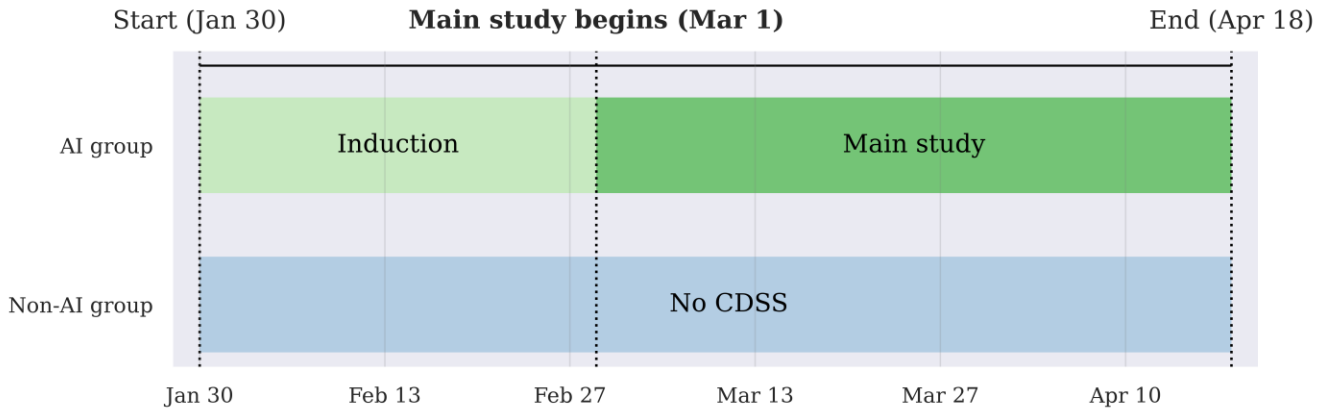
Clinical chatbot integrated into care



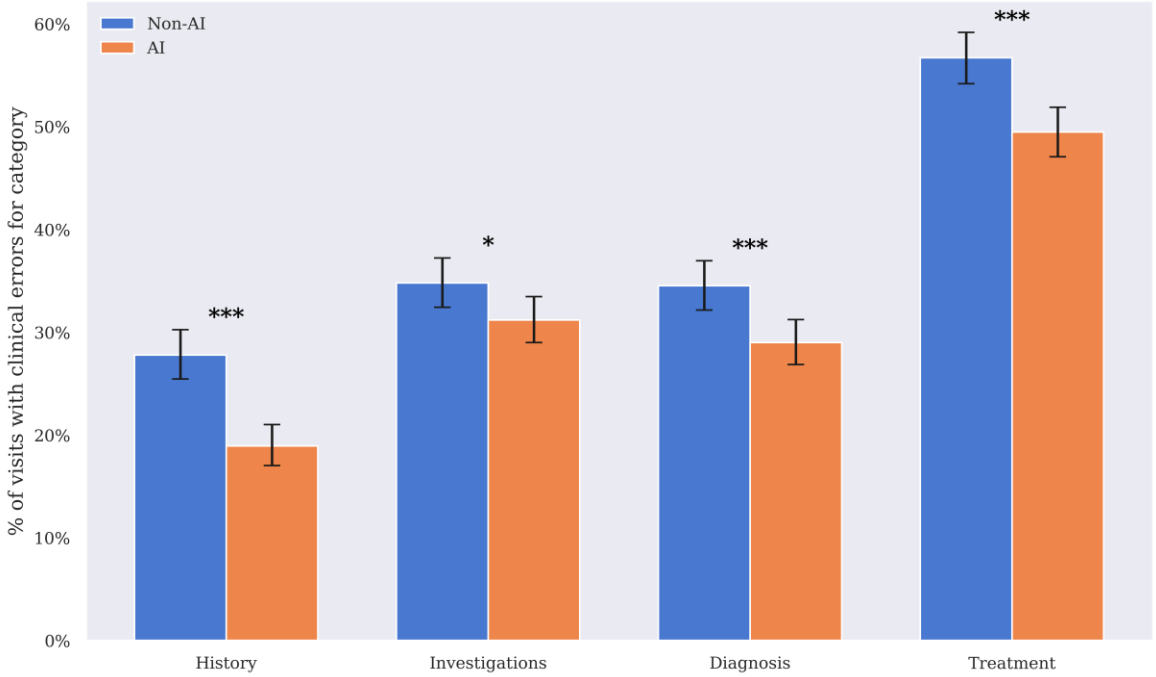
Clinical chatbot integrated into care



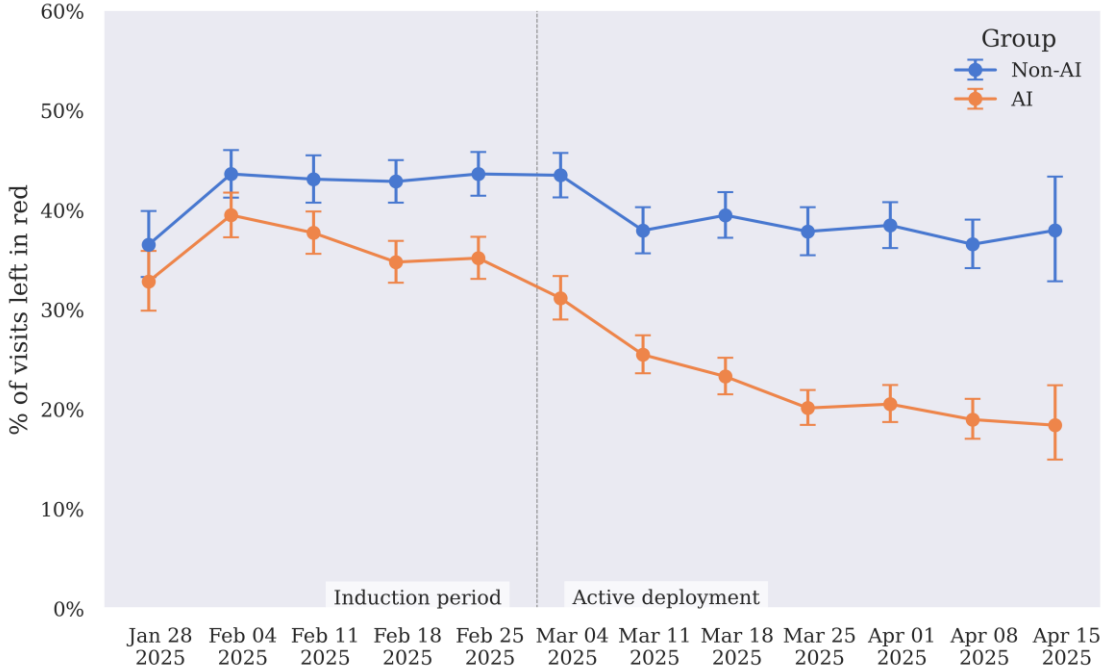
Routine-care evidence from a pragmatic QI stud



AI assistance reduced chart-reviewed errors across four domains



The larger signal arrived with active implementation



The cohort study established *process* benefit

WHAT IT ESTABLISHED

- Use inside routine clinical work
- Nearly 40,000 visits across 15 clinics
- Independent review of 5,666 charts
- Fewer errors across four domains
- An implementation-sensitive mechanism

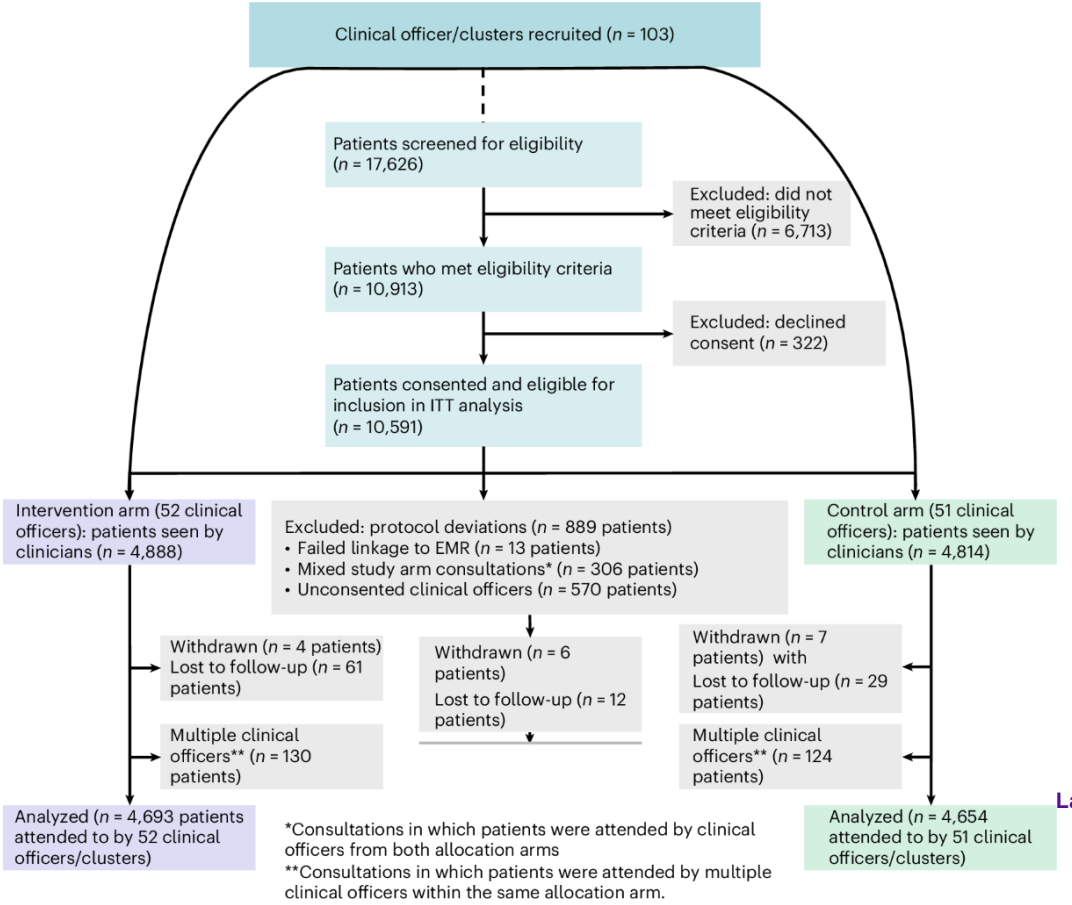
WHAT REMAINED OPEN

- No randomized causal contrast
- Evolving prompts and workflow
- Fair interrater reliability
- Limited patient follow-up
- No demonstrated change in recovery or harm

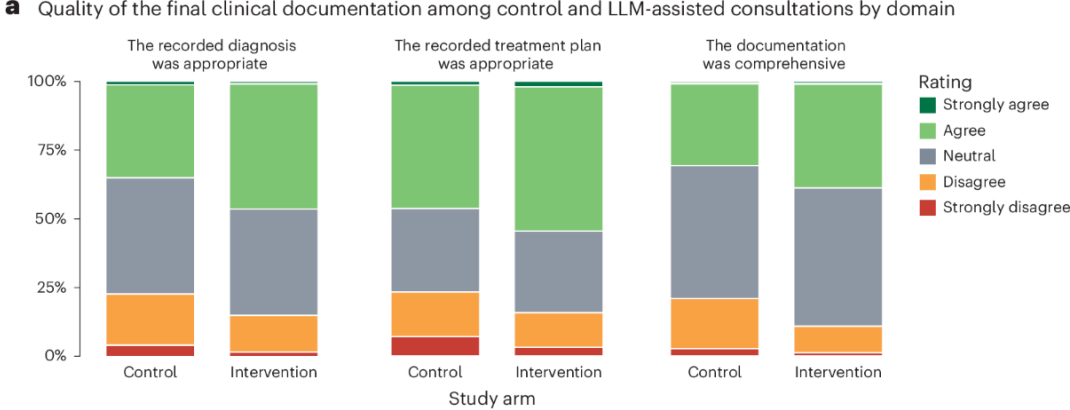
PENDA · THE CLUSTER-RANDOMIZED TEST

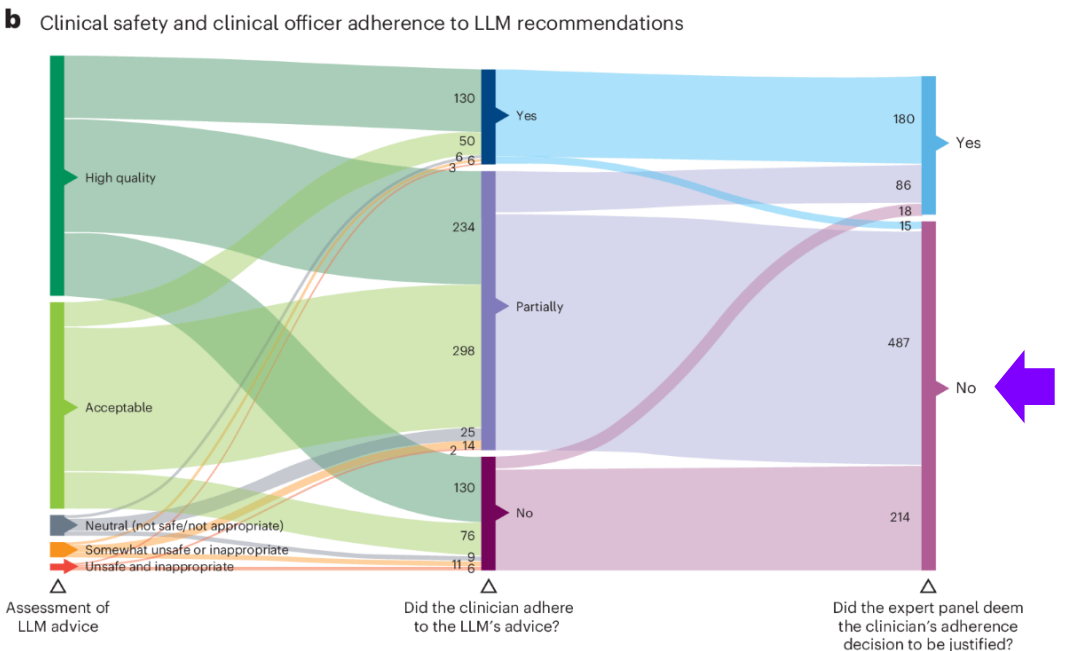
NYU Langone Health

The same tool class was tested across 16 facilities



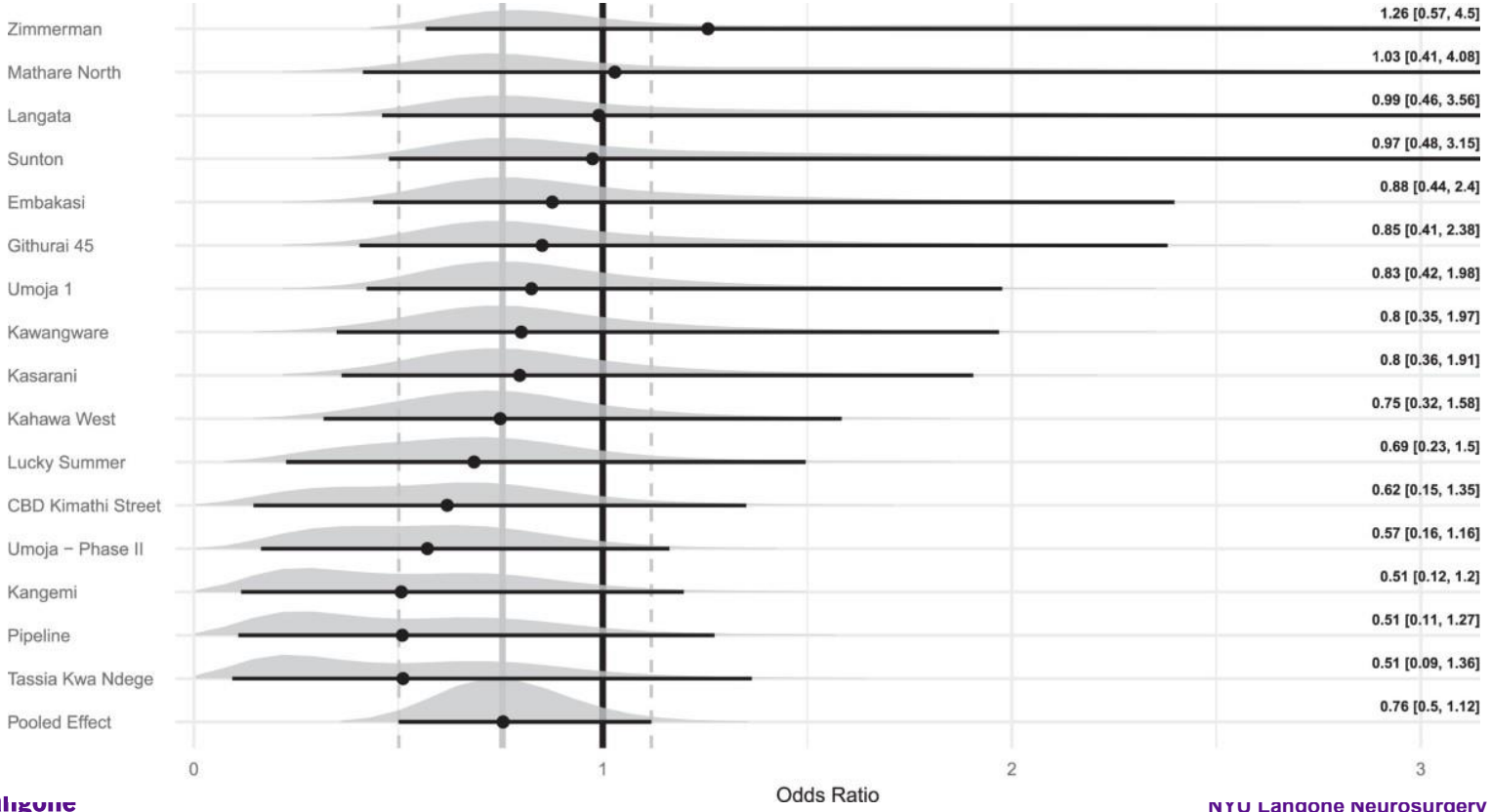
Randomization replicated the process-quality gains





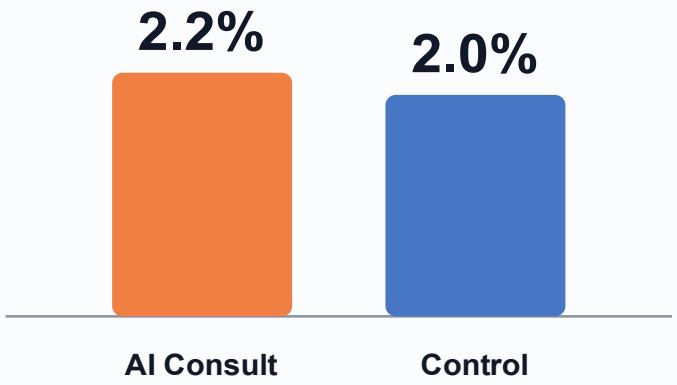
NYU Langone Health

Patient effects varied across sites



The primary endpoint and patient outcome was missed

14-DAY TREATMENT FAILURE



ADJUSTED ESTIMATE

aOR 0.77

95% CI 0.55–1.08

P = .13

Risk difference −5 / 1,000

95% CrI −13 to +1

The trial did not establish patient benefit; its interval still permits a modest benefit.

Replotted from the Penda cluster-randomized trial, Nature Medicine 2026.

Process benefit confirmed, Patient benefit uncertain

QUALITY-IMPROVEMENT EVIDENCE

- Real-world use at scale
- Fewer chart-reviewed errors
- Stronger signal with active implementation
- **Supports workflow viability and process quality**
- **Does not establish patient benefit**

CLUSTER-RANDOMIZED EVIDENCE

- **Process-quality gains replicated**
- **14-day treatment failure not shown**
- Did not establish patient benefit
- **Design matters – powered for a 50% effect size!!!**
- *Maybe just ran the wrong study?*

NYU Langone Health

CASE 2 · EDI + RAPID RESPONSE

NYU Langone Health

EDI is not the best model for patient decline

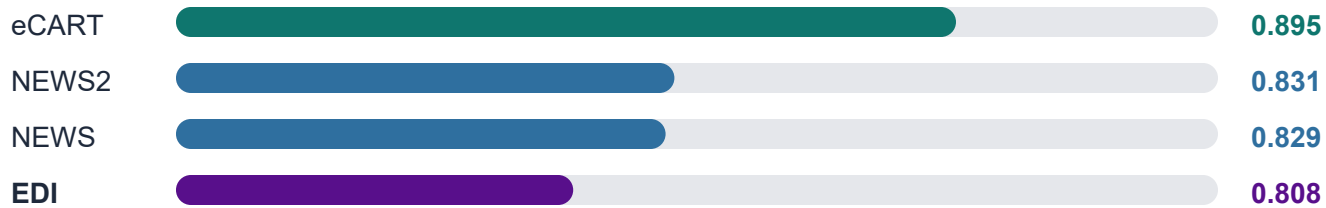
MODEL
Ordinal logistic regression

INPUTS
31 routinely collected measures

CADENCE
Recomputed every 15 minutes

FOOTPRINT
187 Epic instances

EXTERNAL 7-HOSPITAL DISCRIMINATION (AUROC)



EDI discriminates worse than NEWS (pen and paper quality model)

Model calibrated against physician expectations – not too many alarms

SENSITIVITY

49.0%

A local tradeoff among performance, feasibility, and alert burden

RED-ALERT THRESHOLD

EDI ≥60

Chosen and retuned against rapid-response workload

SPECIFICITY

98.9%

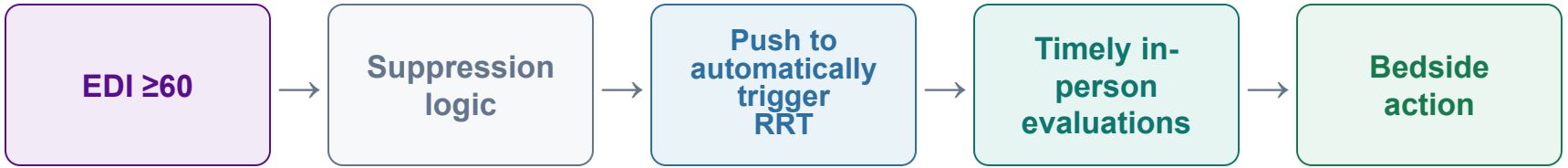
Median 1–10 alerts per hospital per day

SUPPRESS WHEN

- Already in ICU
- Comfort care
- Sepsis/RRT event <6 h
- Prior red alert <48 h
- Procedural area

The threshold allocates scarce clinical attention; it is part of the intervention.

Developed a workflow to turn risk estimates into real rescues

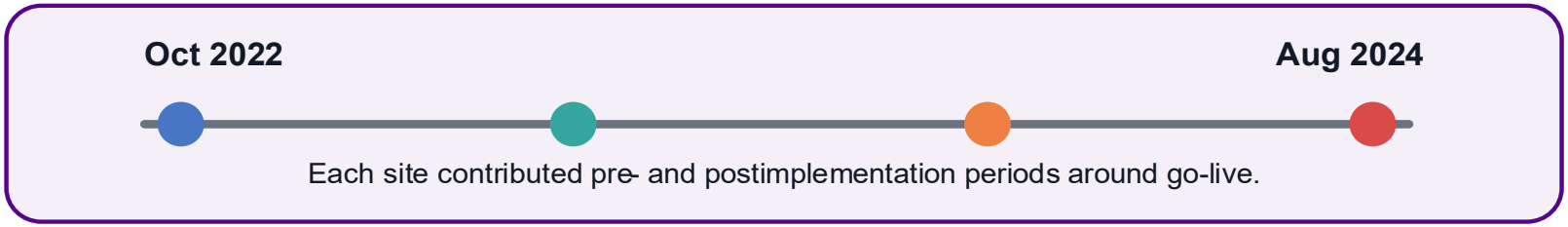


IMPLEMENTATION INFRASTRUCTURE

- Physician + nursing leadership
- Frontline education
- Explicit bedside-response expectation
- Published alert volumes for staffing

A staggered rollout tested the pathway across 11 hospitals

Staggered quasi-experimental implementation



COHORT

23,132

adult med–surg patients

HOSPITALS

11

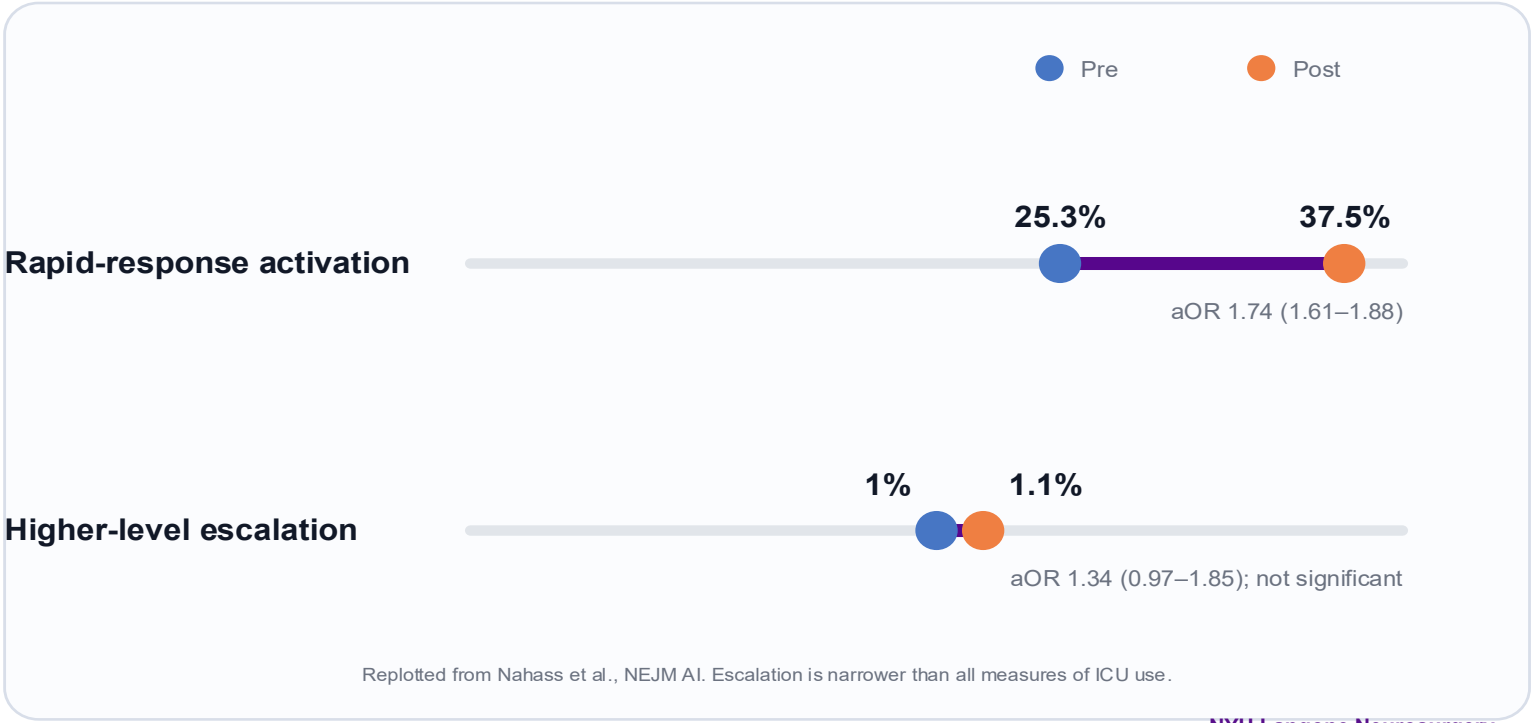
academic + community

MIX

2 academic
6 community teaching
3 nonteaching

Study design and cohort replotted from Nahass et al., NEJM AI.

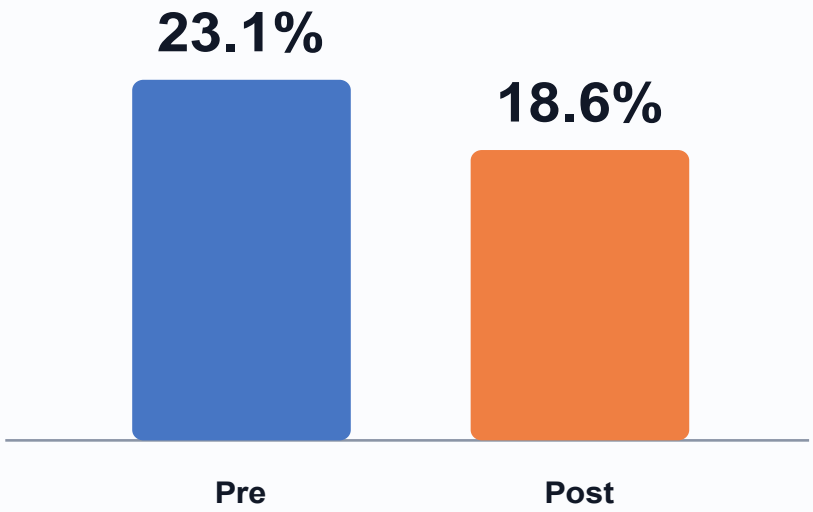
More high-risk patients received bedside RRT evaluation



Replotted from Nahass et al., NEJM AI. Escalation is narrower than all measures of ICU use.

Mortality was lower after implementation

IN-HOSPITAL MORTALITY AMONG HIGH-RISK PATIENTS



ADJUSTED ASSOCIATION

aOR 0.82

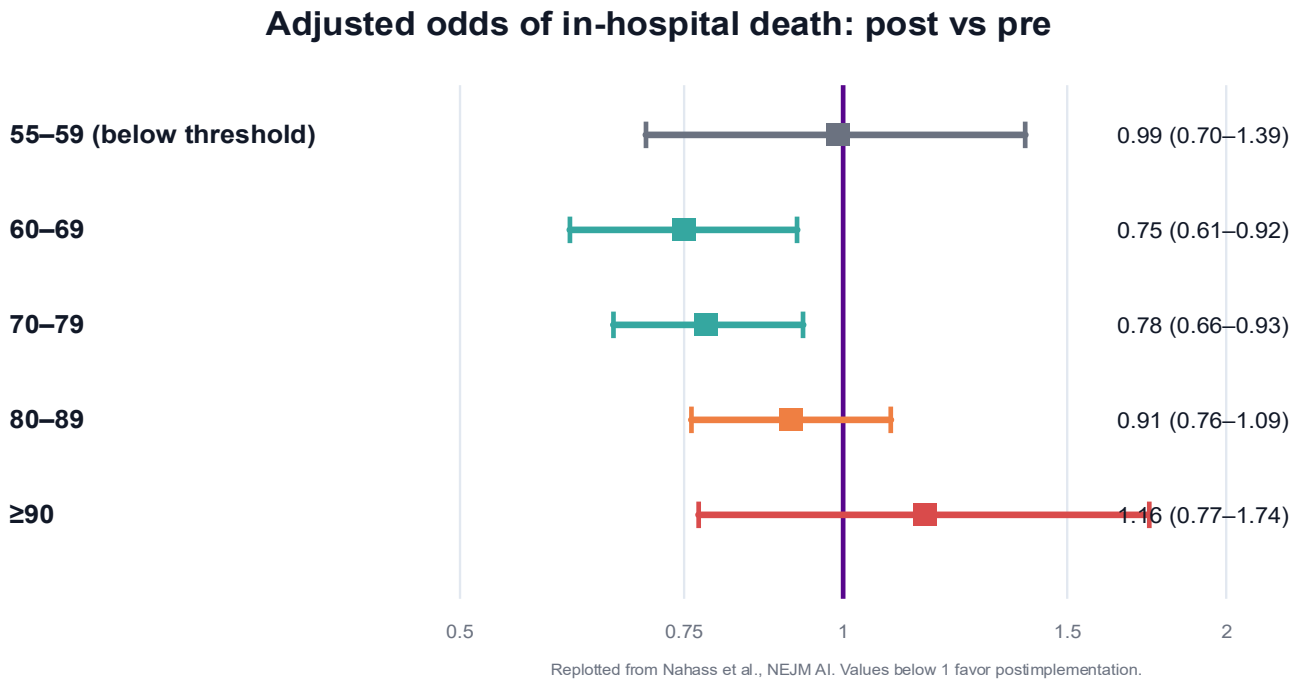
95% CI 0.74–0.91

18% lower adjusted odds

Unadjusted absolute difference: −4.5 percentage points.

Association after a nonrandomized staggered rollout; not an 18% causal risk reduction.

Rescue occurred in “reasonable” cases



Replotted from Nahass et al., NEJM AI. Values below 1 favor postimplementation.

Implementation fidelity tracked with outcomes—but remains post hoc

IMPLEMENTATION FIDELITY = NATURAL CONTROL

- Mortality fell at 9 of 11 sites
- Two exceptions had incomplete delivery
  - No 24/7 rapid-response coverage
  - Team never joined the notification group

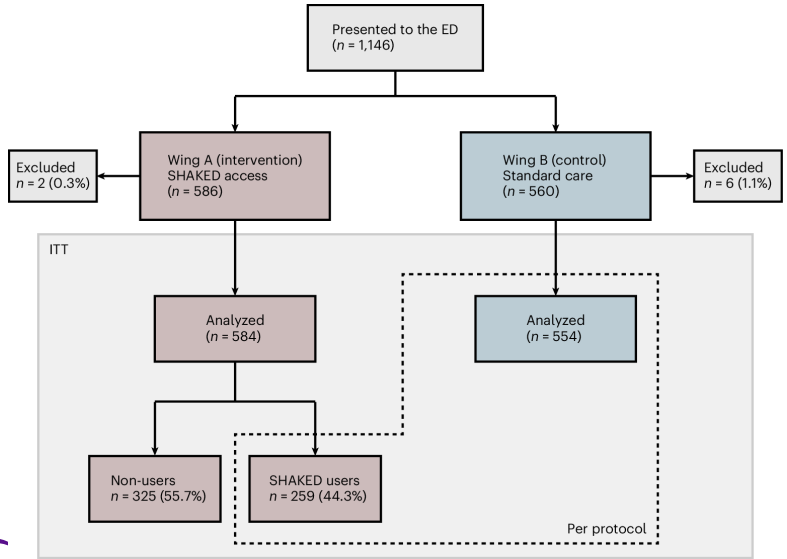
LIMITS

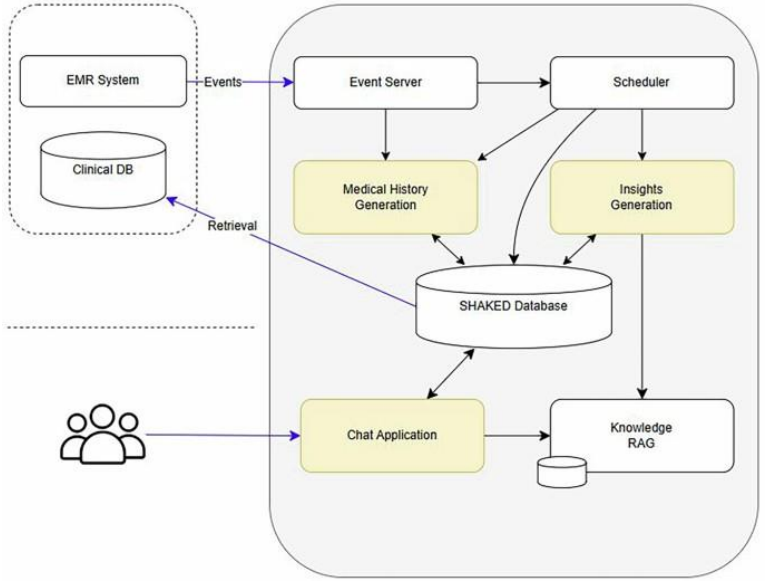
- Staggered pre/post design
- Unequal seasonal exposure
- Inconsistent response-timing data
- **Fidelity observations are ecological - not proof of mechanism** ***(but as a whole, the bundle seems to have worked!)***

CASES 3+4 · SHAKED & NYUCONSULT

NYU Langone Health

SHAKED: Prospective evaluation of a large language model clinical decision support system in the emergency department

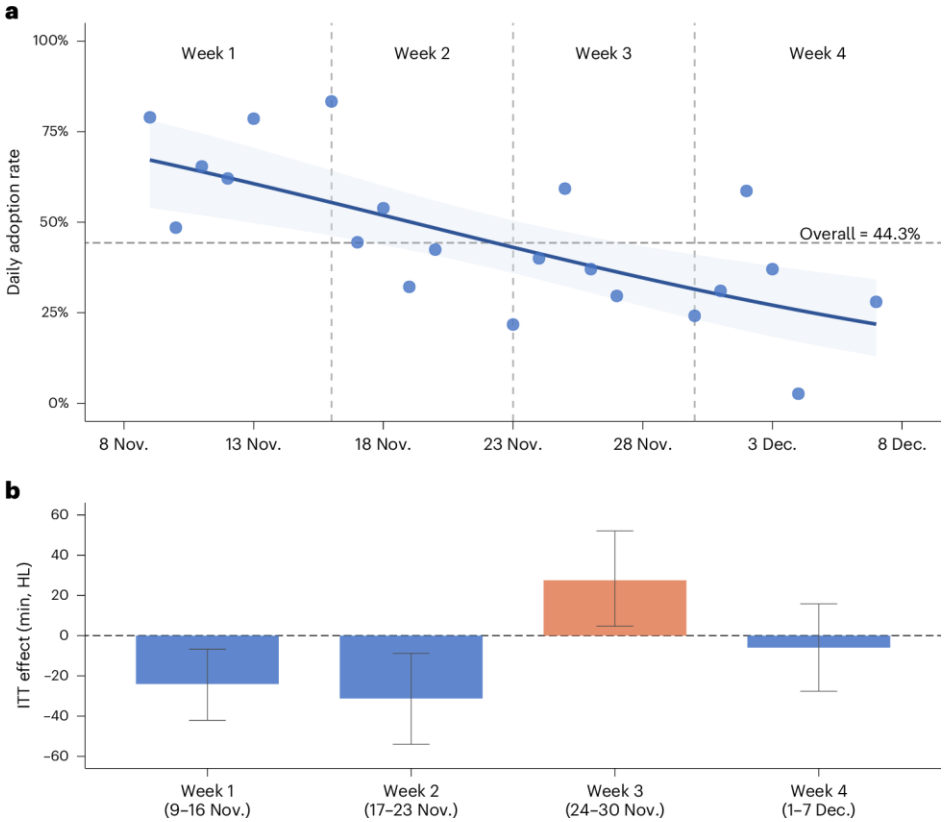




Leibovitch, L., Ahituv, A., Gorenshtein, A. *et al.* Prospective evaluation of a large language model clinical decision support system in the emergency department. *Nat Med* (2026). https://doi.org/10.1038/s41591-026-04601-5

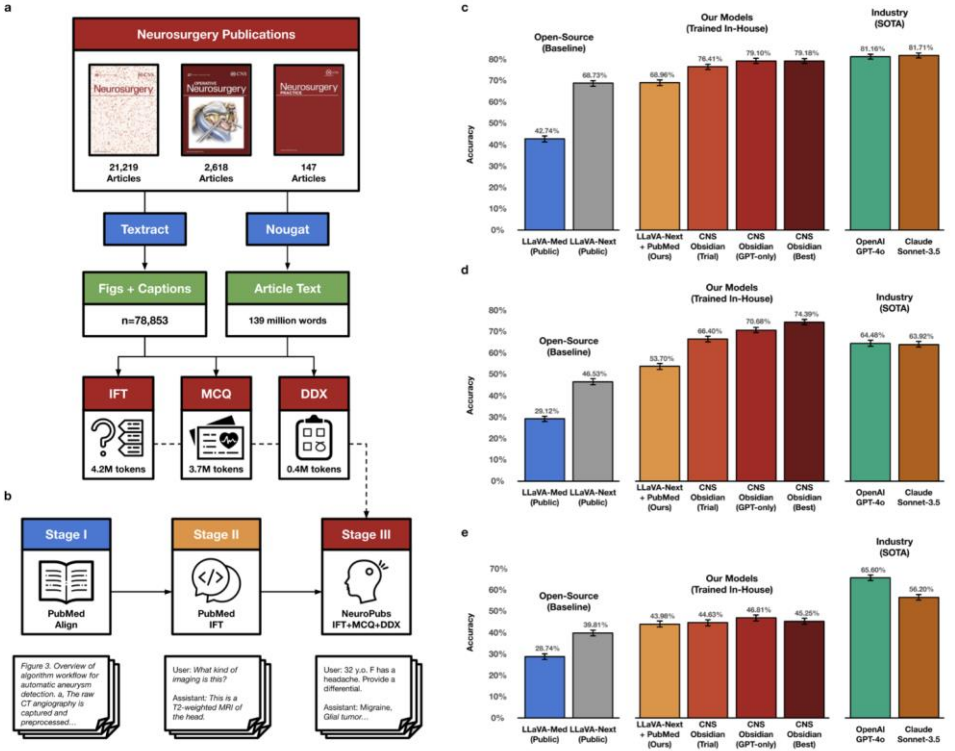
SHAKED: It worked and no one used

- Expert review: 99% DDx accuracy
- 100% test recommendation appropriateness
- 100% factually consistent summaries
- ITT showed trend towards decreasing consultation time.
- **Declined in use from 67.9% to 29.7% over 4 weeks.**



Leibovitch, L., Ahituv, A., Gorenshtein, A. *et al.* Prospective evaluation of a large language model clinical decision support system in the emergency department. *Nat Med* (2026). https://doi.org/10.1038/s41591-026-04601-5

CNS-OBSIDIAN: A language model for neurosurgical consults

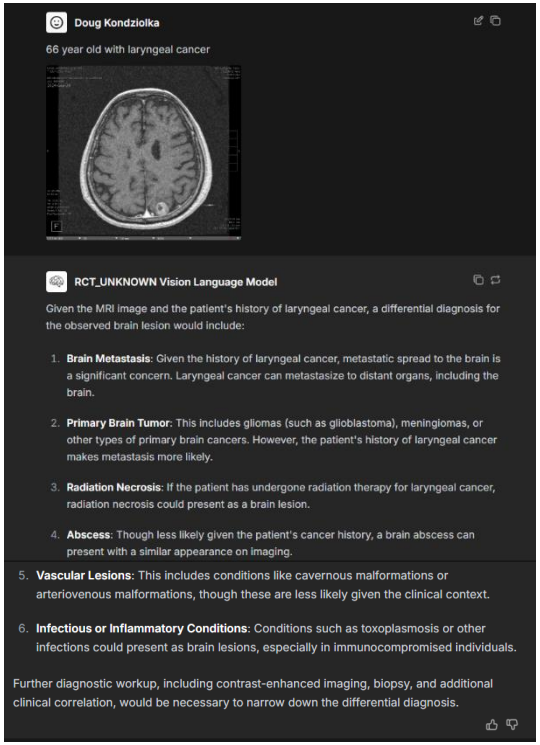


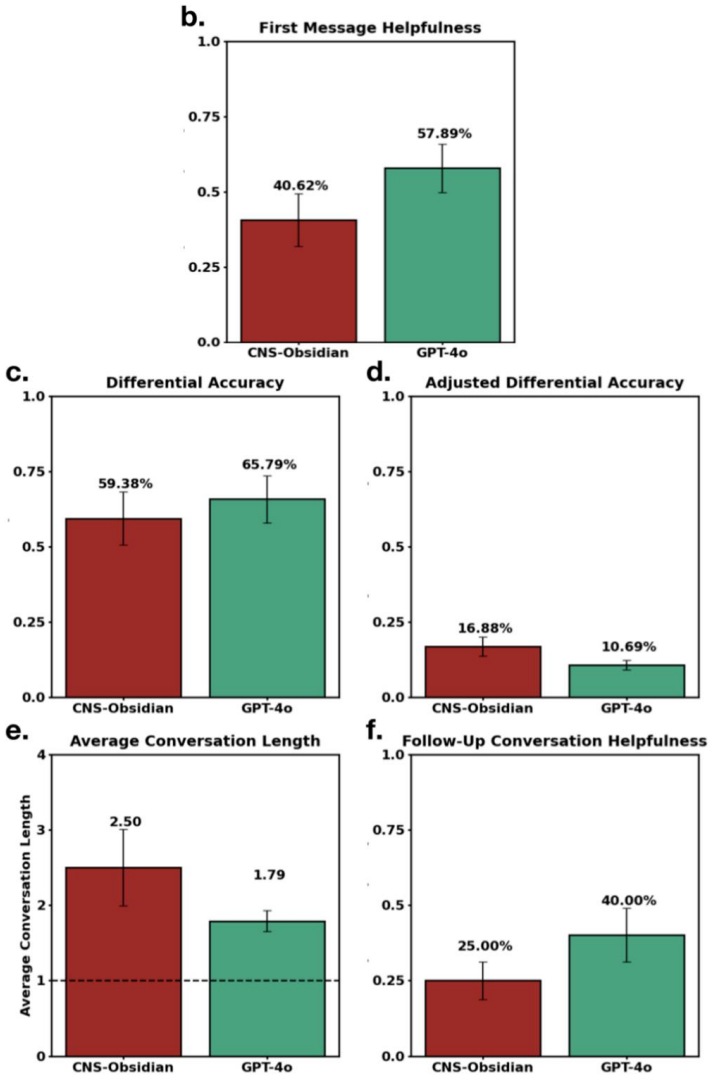
Alyakin A, Stryker J, Alber DA, Lee JV, Sangwon KL, Duderstadt B, Save A, Kurland D, Frome S, Singh S, Zhang J, Yang E, Park KY, Orillac C, Valliani AA, Neifert S, Liu A, Patel A, Livia C, Lau D, Laufer I, Rozman PA, Hidalgo ET, Riina H, Feng R, Hollon T, Aphinyanaphongs Y, Golfinos JG, Snyder L, Leuthardt EC, Kondziolka D, Oermann EK. CNS-Obsidian: A Neurosurgical Vision-Language Model Built From Scientific Publications. Neurosurgery. 2026 May 19. doi: 10.1227/neu.0000000000004070. Epub ahead of print. PMID: 42153721.

CNS-OBSIDIAN

92-day trial randomized controlled trial comparing GPT4o1 and OBSIDIAN on surgical service.

- **70 chats were generated** (0.75/day)
  - 38 randomized to GPT-4o1
  - 32 randomized to Obsidian
- **Non-inferiority of OBSIDIAN to GPT-4o1** in terms of DDx accuracy.
- **However … 959 total consults** seen in that time (7% utilization rate)

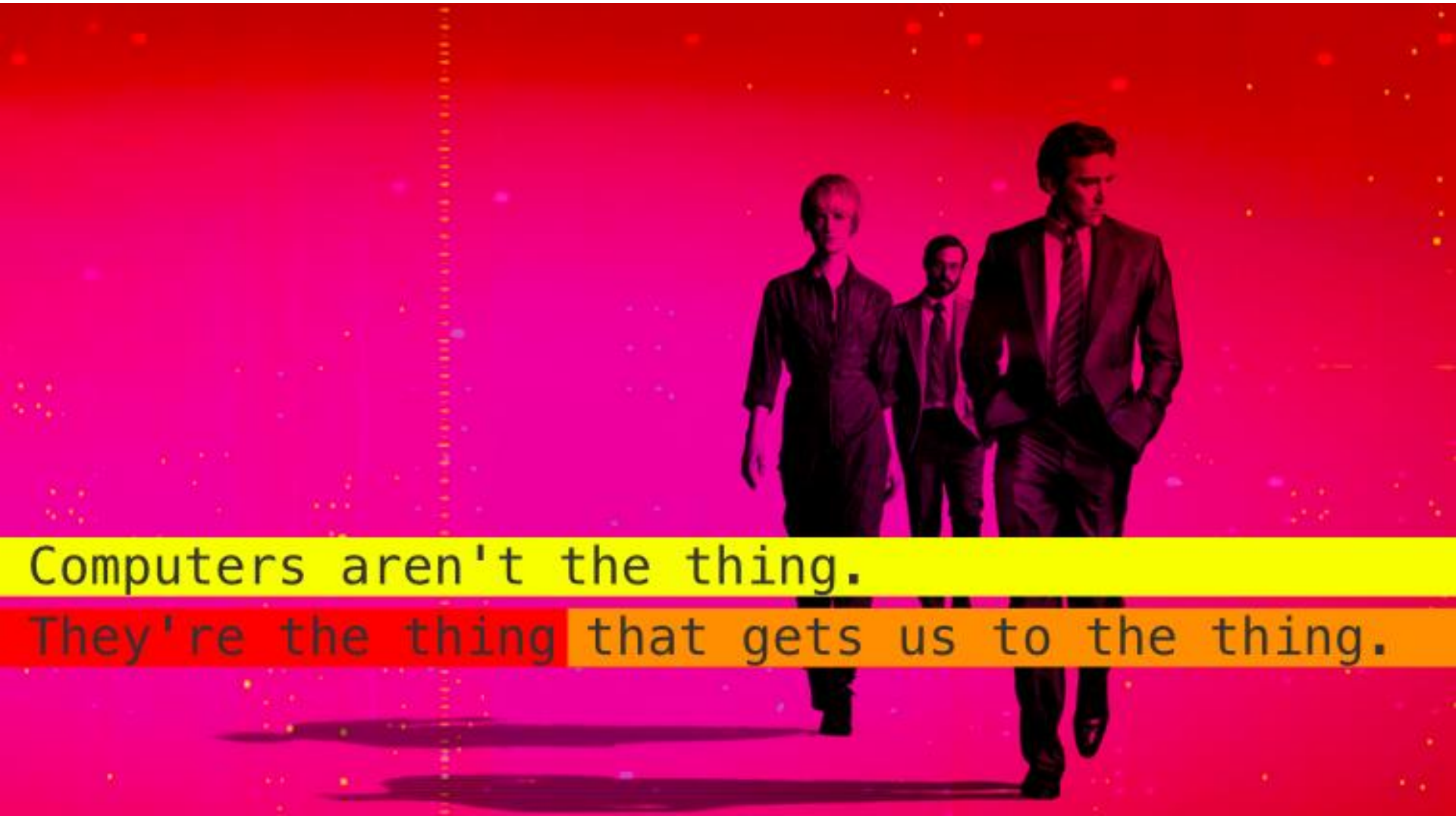




NYU Langone Health

Conclusion = Physicians hate chat interfaces

Alyakin A, Stryker J, Alber DA, Lee JV, Sangwon KL, Duderstadt B, Save A, Kurland D, Frome S, Singh S, Zhang J, Yang E, Park KY, Orillac C, Valliani AA H, Feng R, Hollon T, Aphinyanaphongs Y, Golfinos JG, Snyder L, Leuthardt EC, Kondziolka D, Oermann EK. CNS-Obsidian: A Neurosurgical Vision-Langu 10.1227/neu.0000000000004070. Epub ahead of print. PMID: 42153721.


Computers aren't the thing.
They're the thing that gets us to the thing.

MODEL EVIDENCE

SYSTEM EVIDENCE

PATIENT-BENEFIT EVIDENCE

STAGE 1

1. BENCH CAPABILITY

Can the model perform the task?

REAL-WORLD SAMPLED DATA

Vishwanath et al., 2026



GREATER CLINICAL GROUNDING

SYNTHETIC DATA

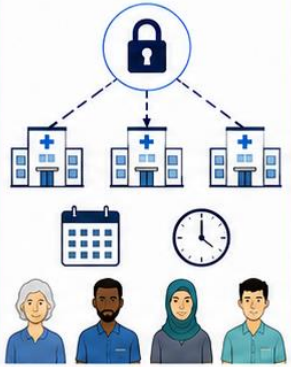
Arora et al., 2025



STAGE 2

2. EXTERNAL VALIDATION

Does performance travel across sites, time, and subgroups?



Zech et al., 2018 • Wong et al., 2021

STAGE 3

3. CLINICAL-TASK EVALUATION

Does it help clinicians perform the task—including in simulated environments?

Reader studies • simulated cases • virtual workflows



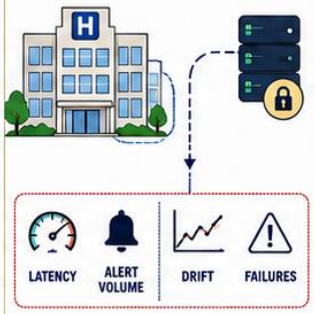
Tschandl et al., 2020 • Goh et al., 2024

STAGE 4

4. STEALTH OBSERVATIONAL STUDY

(SILENT DEPLOYMENT)

Will it operate safely here before influencing care?



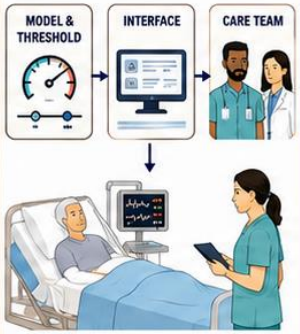
Kwong et al., 2022 • Corbin et al., 2023

STAGE 5

5. COHORT TRIAL

RETROSPECTIVE OR PROSPECTIVE • REAL DEPLOYMENT

Does the deployed model–workflow bundle change care?

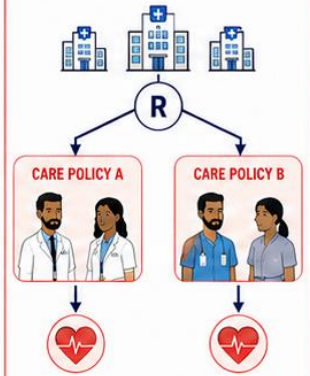


Adams et al., 2022 • Nahass et al., 2026

STAGE 6

6. PRAGMATIC RANDOMIZED TRIAL

Does the care policy improve patient outcomes?



Lin et al., 2024 • Agweyu et al., 2026

VERSION-AWARE SURVEILLANCE

MODEL BEHAVIOR → **OPERATIONAL TRIGGER** → **CLINICIAN ATTENTION** → **CLINICAL ACTION** → **PATIENT OUTCOME**

CLINICAL BENEFIT = MODEL × THRESHOLD × INTERFACE × WORKFLOW × RESPONSE CAPACITY

Closing Thoughts

AI isn't a book, AI is a *medical tool*.

Process benefit does not imply clinical benefit

We need trials of models + workflows, not just models

***How* clinicians use models is *as important* as *what models* they use.**

NYU Langone Health

Thank you to all of our wonderful team and collaborators!

OLAB

Eric Karl Oermann, M.D.
Associate Professor of Neurosurgery, Radiology, and Data Science

Lavender Yao Jiang, B.S.
PhD Student
NYU Center for Data Science
(Prev. CMU)

Chris Xujin Liu, B.S.
PhD Student
NYU Tandon School of Engineering
(Prev. CMU)

Radhika Dua, B.S.
PhD Student
NYU Center for Data Science
(Prev. Google)

Ioana Marinescu, B.S.
PhD Student
NYU Center for Data Science
(Prev. Princeton)

Qingyang Zhu, B.S.
PhD Student
NYU Center for Data Science
(Prev. Shanghai Tech and Alibaba)

Frederick Kwon, M.D., Ph.D.
Resident in Radiology
(Prev. ISMMS)

STUDENTS
Anton Alyakin (Visiting)
Mustafa Nasir-Moin
Karl Sangwon Lee
Daniel Alber
Grace Von Oiste
Ashley Amend Thompson
Eunice Yang
Michelle Murata
Gavin Yang

Visiting Scholars
Sumedha Rai, M.S. (Kepos Capital)
Jan Niehues, Ph.D. (U Aachen)
Beakal Lemeneh, M.S. (Blackrock)
Paz Vives, M.S. (JP Morgan)

Perception and Brain Dynamics Laboratory
Biyu He, Ph.D.
Ayaka Hachisuka, B.S.
Jonathan Shor, M.S.



NYU Neurosurgery
Douglas Kondziolka, M.D.
John Golfinos, M.D.
Daniel Orringer, M.D.
Howard Riina, M.D.
Shannon Ciprut
Allison George Lucas
Qian Qian Du
Sean Neifert
Aly Valliani

Neurodata Hub
Xu Han, M.S
Alvin Pan, M.S.
Jaden Stryker, B.S.



NYU OSR
Dafna Bar-Sagi, Ph.D.



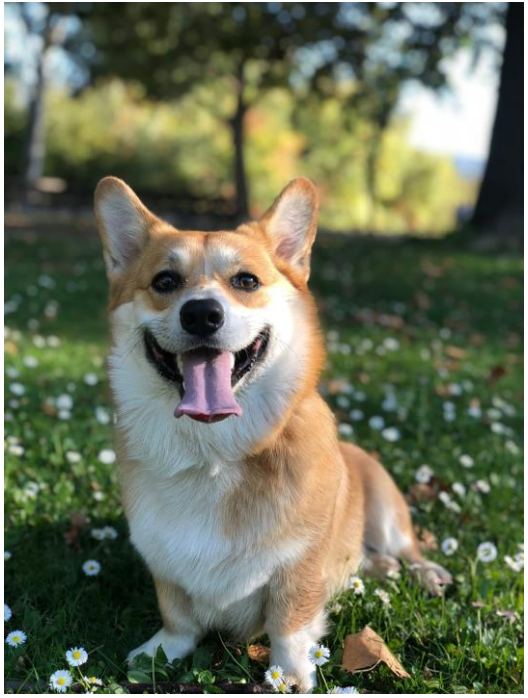
Kyunghyun Cho, Ph.D.
Professor of Data Science and Computer Science

NVIDIA

Anthony Costa, Ph.D.
Mona Flores, M.D.
Katherine E. Link, B.S.



NYU MCIT
Nader Mherabi, CIO NYULH
Yindalon Aphinyanaphongs, M.D.,Ph.D.
Jeff Zhang, M.S.



OTTO

NYU Langone Health

NYU Center for Data Science

NYU NEUROSCIENCE

Carl Feinberg

W. M. KECK FOUNDATION

NIH NATIONAL CANCER INSTITUTE

NYU Langone Radiology—Center for Biomedical Imaging