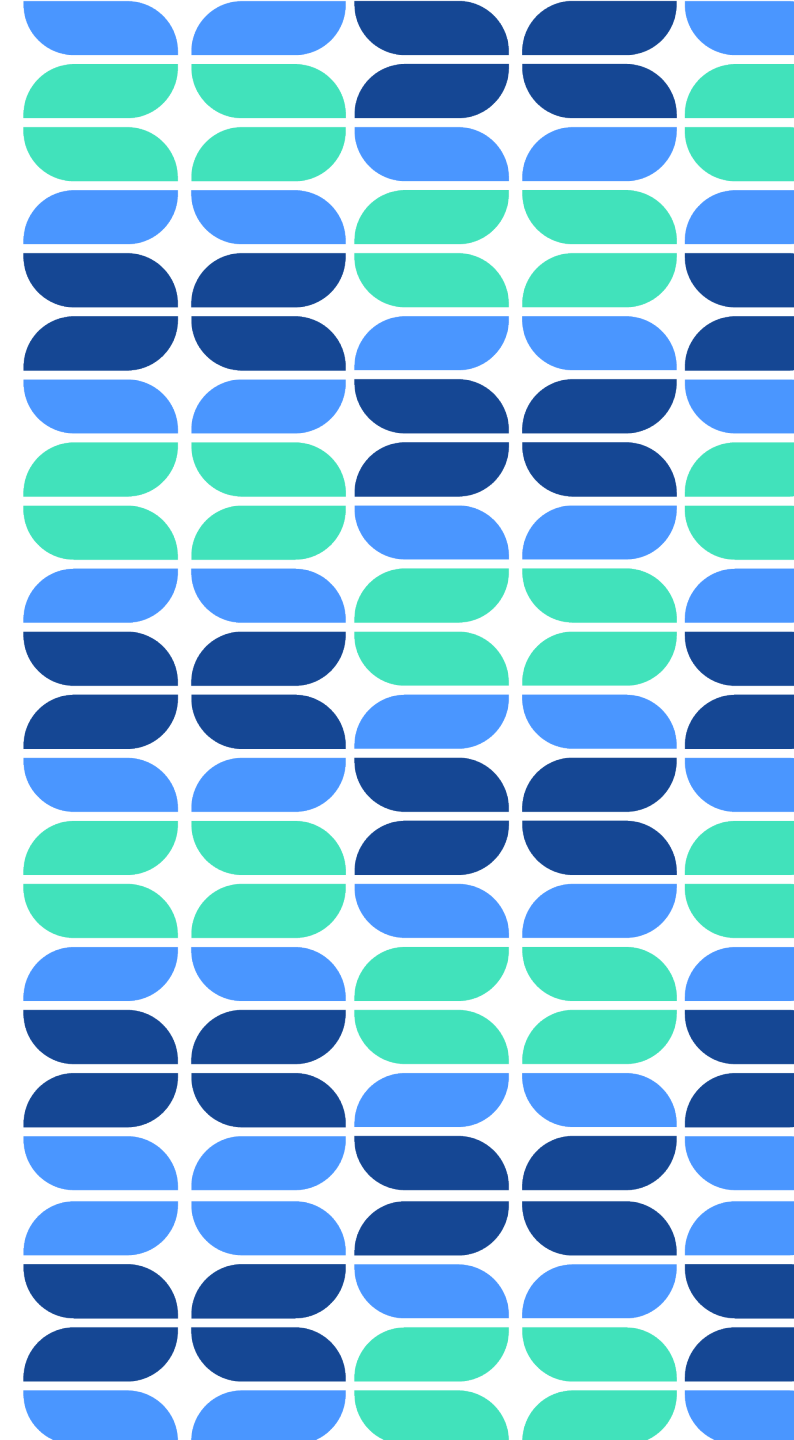




Joint Clinical Meeting

WEDNESDAY, DECEMBER 4, 2024



WELCOME



Joint Clinical Forum on “AI in Health Care”

KARANDEEP SINGH, MD, MMSC

CHIEF HEALTH AI OFFICER FOR UC SAN DIEGO HEALTH, AND JOAN AND IRWIN
JACOBS CHANCELLOR'S ENDOWED CHAIR AND ASSOCIATE PROFESSOR OF
MEDICINE IN BIOMEDICAL INFORMATICS AT UC SAN DIEGO



TACKLING THE HEALTH AI PARADOX

KARANDEEP SINGH, MD, MMSC

Joan and Irwin Jacobs Chancellor's Endowed Chair in Digital Health Innovation
Associate Professor in Biomedical Informatics
Chief Health AI Officer, UC San Diego Health

SPECIAL ARTICLE

A COMPUTER-BASED MEDICAL-HISTORY SYSTEM*

WARNER V. SLACK, M.D.,† G. PHILLIP HICKS, PH.D.,‡ CHARLES E. REED, M.D.,§ AND
LAWRENCE J. VAN CURA, M.S.¶



HAVE YOU EVER HAD
AN ALLERGIC REACTION
TO PENICILLIN

- 1 YES
- 2 NO
- 3 DONT KNOW
- 4 DONT UNDERSTAND

YOUR ANSWER IS ?

TABLE 1. Results of Computer-Based Histories of Allergy Compared with Those of Histories Taken by Physicians.

CONDITIONS DETECTED BY HISTORY TAKING	CONDITIONS DETECTED BY BOTH PHYSICIAN & COMPUTER	CONDITIONS DETECTED BY PHYSICIAN ONLY	CONDITIONS DETECTED BY COMPUTER ONLY
Urticaria	0	0	12
Allergic rhinitis	2	0	7
Asthma	4	0	2
Drug allergy	7	0	1

“Any doctor who **can be** replaced by a computer **deserves** to be replaced by a computer.”

– Warner Slack, MD

HELP

58% reduction in per-patient antibiotic costs

30% decrease in antibiotic-related adverse events

Leeds Abdominal Pain

91% accurate in diagnosing surgical abdominal pain as compared to 83% for senior clinicians

INTERNIST-1

Performance matched hospital clinicians on *NEJM* clinical cases but not experts

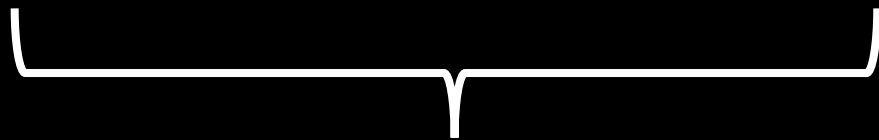
Since then, advancements in...

Tree ensembles

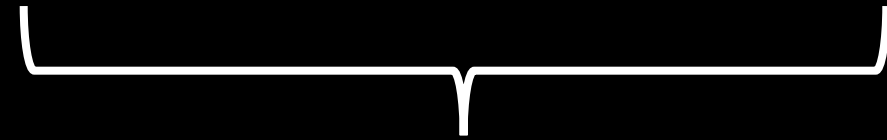
Regression for
prediction

Latent
representations

Deep learning

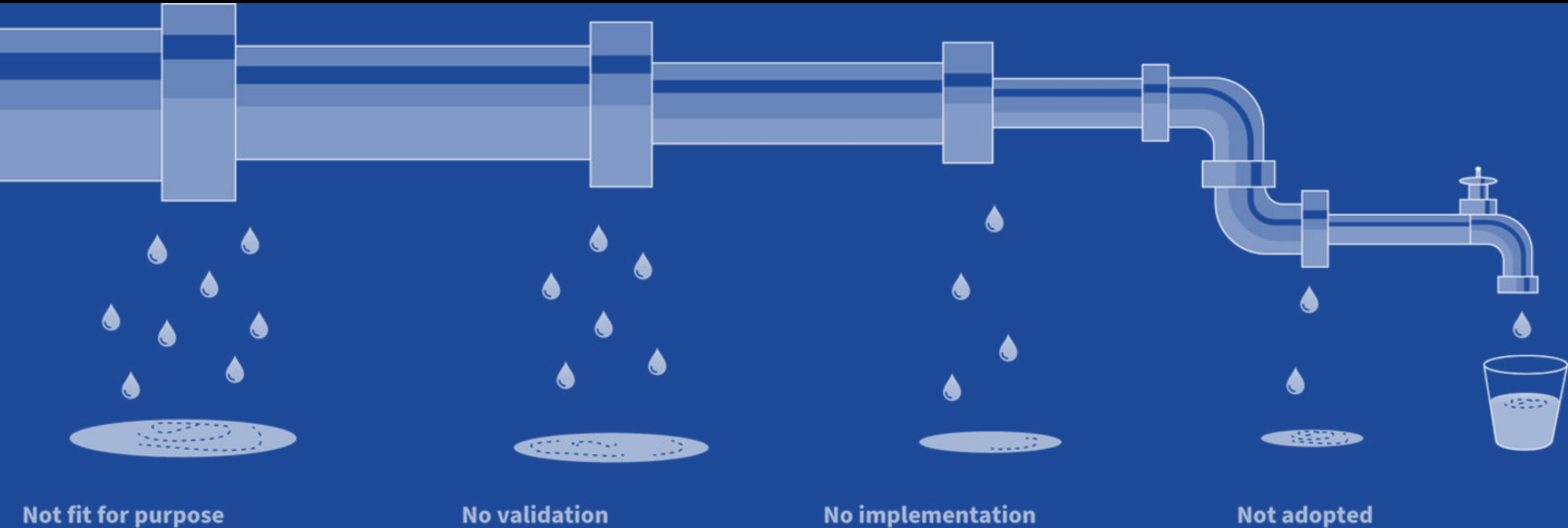


Statistics



Computer
science

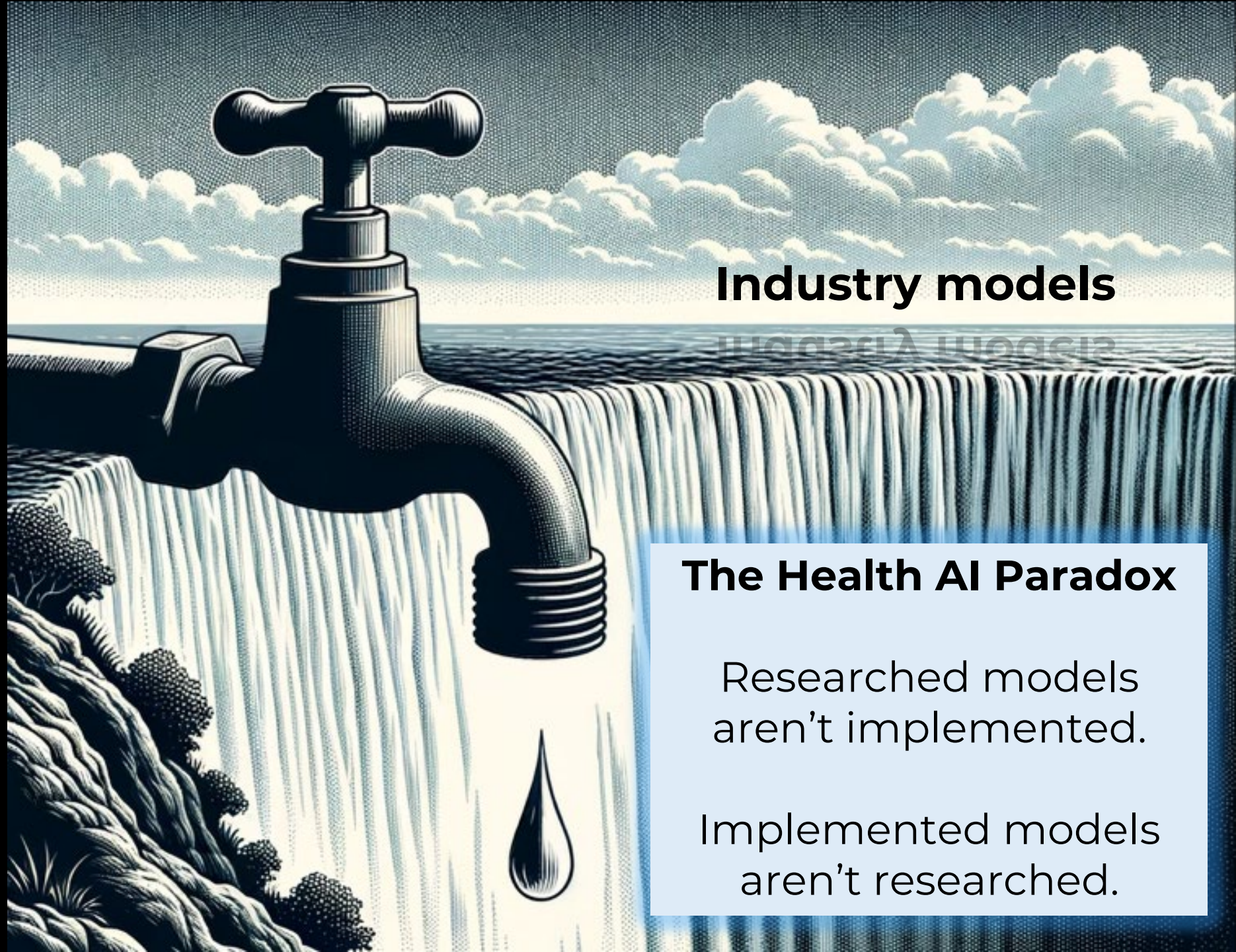
Where are all the AI models?



Florien S. van Royen et al. *European Respiratory Journal* 2022.



Research models



Industry models

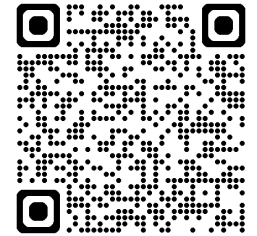
The Health AI Paradox

Researched models
aren't implemented.

Implemented models
aren't researched.

External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients

Andrew Wong, MD; Erkin Otles, MEng; John P. Donnelly, PhD; Andrew Krumm, PhD; Jeffrey McCullough, PhD; Olivia DeTroyer-Cooley, BSE; Justin Pestrue, MEcon; Marie Phillips, BA; Judy Konye, MSN, RN; Carleen Penozza, MHSA, RN; Muhammad Ghous, MBBS; Karandeep Singh, MD, MMSc



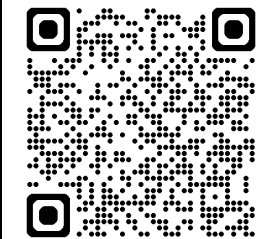
[DOI: 10.1056/Aloa2300032](https://doi.org/10.1056/Aloa2300032)

ORIGINAL ARTICLE

Evaluation of Sepsis Prediction Models before Onset of Treatment

Fahad Kamran , Ph.D.,¹ Donna Tjandra , M.S.,¹ Andrew Heiler , M.B.A.,² Jessica Virzi , M.S.N.,³ Karandeep Singh , M.D.,^{3,4} Jessie E. King , M.D., Ph.D.,⁵ Thomas S. Valley , M.D., M.Sc.,^{6,7} and Jenna Wiens , Ph.D.^{1,3}

Received: July 10, 2023; Revised: November 9, 2023; Accepted: November 15, 2023; Published: February 7, 2024



TOM SIMONITE BUSINESS JUN 21, 2021 11:00 AM

An Algorithm That Predicts Deadly Infections Is Often Flawed

A study found that a system used to identify cases of sepsis missed most instances and frequently issued false alarms.

STAT+

A STAT INVESTIGATION

Epic's AI algorithms, shielded from scrutiny by a corporate firewall, are delivering inaccurate information on seriously ill patients



By Casey Ross July 26, 2021

Health Health Care Medical Mysteries Science Wellness

HEALTH

Sepsis prediction tool used by hospitals misses many cases, study says. Firm that developed the tool disputes those findings.

By Erin Blakemore

June 26, 2021 at 8:00 a.m. EDT

Better models

Research Letter

ONLINE FIRST

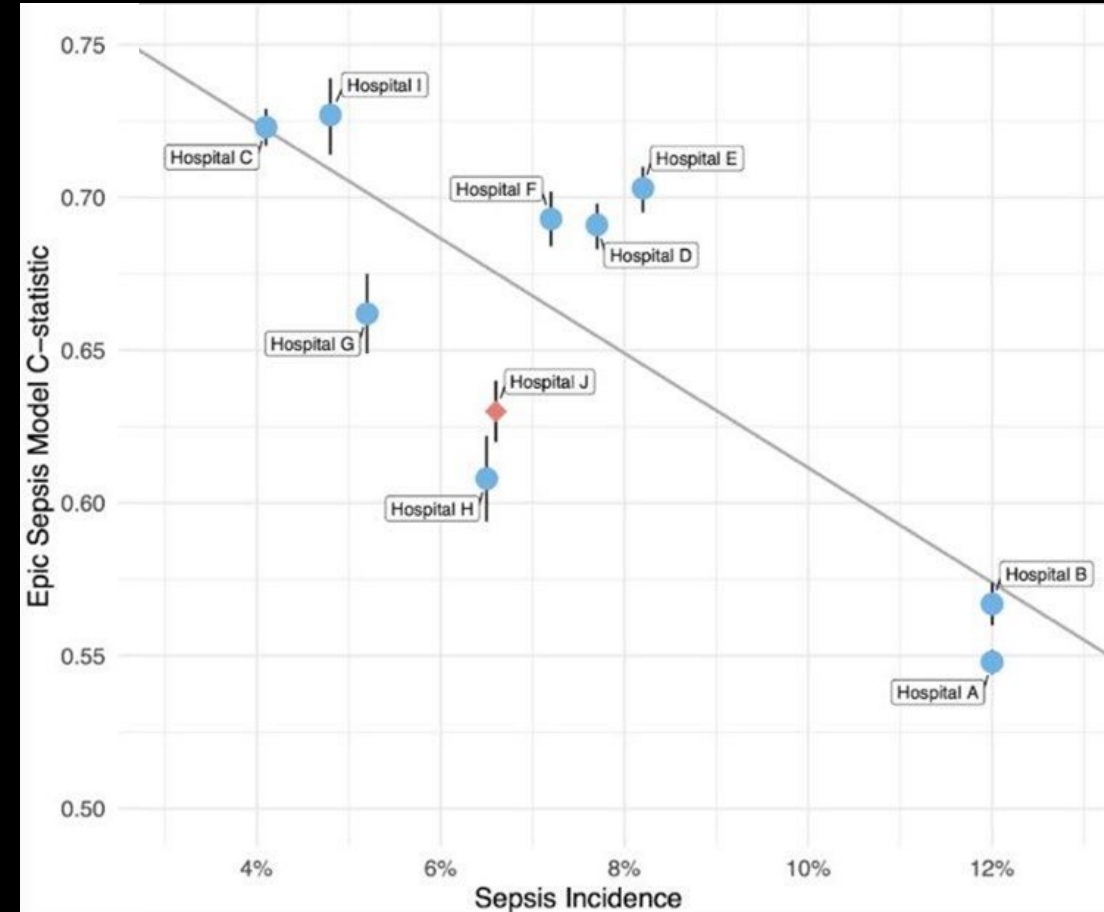
April 3, 2023

Factors Associated With Variability in the Performance of a Proprietary Sepsis Prediction Model Across 9 Networked Hospitals in the US

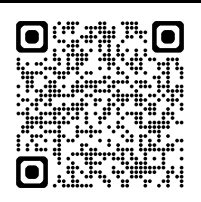
Patrick G. Lyons, MD, MSc^{1,2}; Mackenzie R. Hofford, MD³; Sean C. Yu, PhD³; Andrew P. Michelson, MD¹; Philip R. O. Payne, PhD³; Catherine L. Hough, MD, MSc⁴; Karandeep Singh, MD, MMSc⁵

[» Author Affiliations](#) | [Article Information](#)

JAMA Intern Med. Published online April 3, 2023. doi:10.1001/jamainternmed.2022.7182



Worse models





- A proprietary model was developed to predict the likelihood of sepsis in hospitalized patients and was implemented at hundreds of hospitals around the country. An independent study showed that the model predictions underperformed relative to the designer's claims while also causing 'alert fatigue' by falsely alerting likelihood of sepsis.[i]



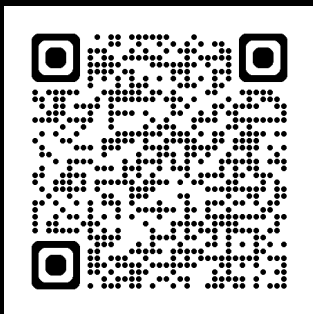
SAFE AND EFFECTIVE SYSTEMS

YOU SHOULD BE PROTECTED FROM UNSAFE OR INEFFECTIVE SYSTEMS



▶ OSTP

▶ BLUEPRINT FOR AN AI BILL OF RIGHTS



HEALTH TECH

STAT+

In new guidance, FDA says AI tools to warn of sepsis should be regulated as devices

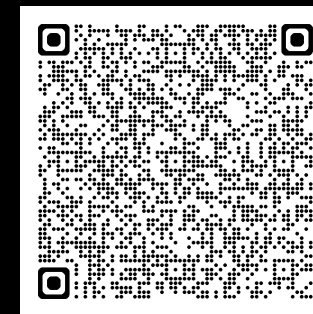


By Casey Ross Sept. 27, 2022

Reprints



ALEX HOGAN/STAT



Implementing DeepMind's state-of-the-art acute kidney injury model

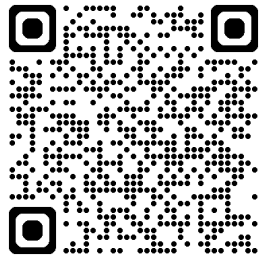
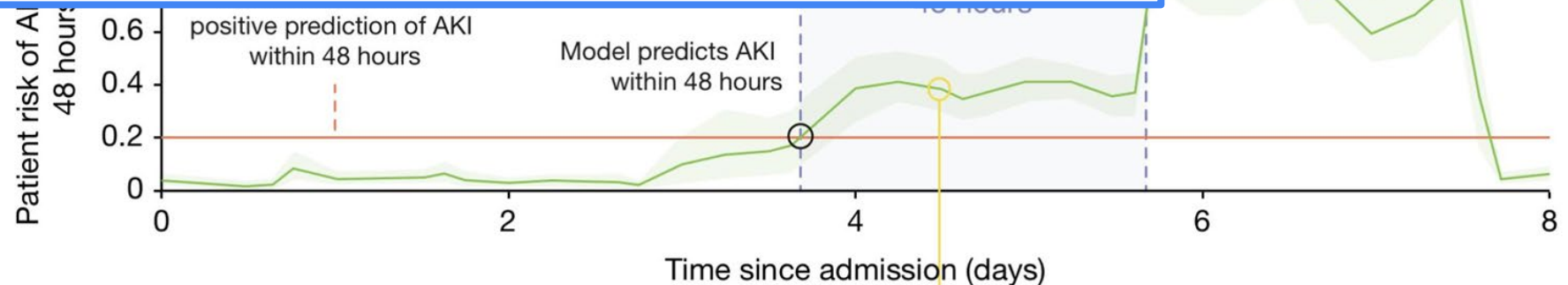
LETTER

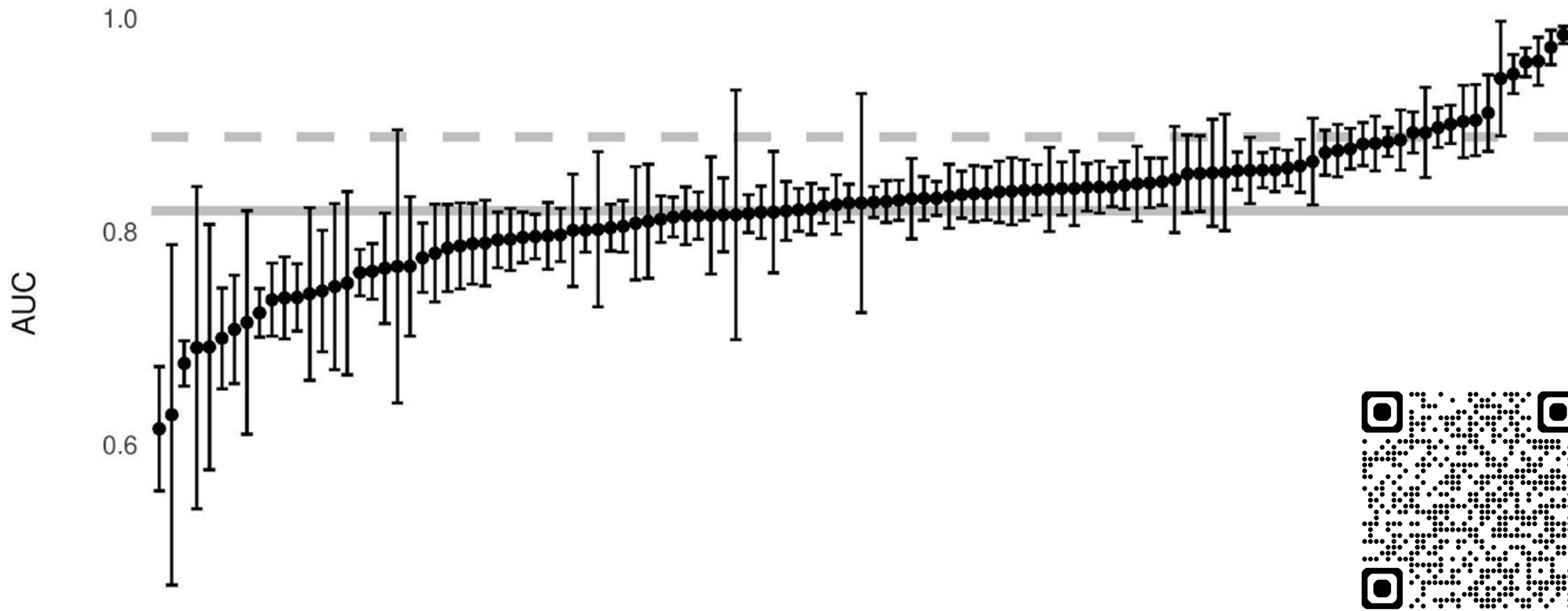
<https://doi.org/10.1038/s41586-019-1390-1>

A clinically applicable approach to continuous prediction of future acute kidney injury

Nenad Tomašev^{1*}, Xavier Glorot¹, Jack W. Rae², Mihaela W. Beaulac¹, Adam W. Senior¹, Clemens Meyer¹, Suman Ravuri¹, Ivan Protsyuk¹, Julien Cornebise^{1,12}, Hugh Montgomery³, Gertraud Malsbenden¹, Demis Hassabis¹, Dominic King¹, Mustafa Sulam¹, Shakir Mohamed^{1,13}

Despite the state-of-the-art retrospective performance of our model compared to existing literature, future work should now prospectively evaluate and independently validate the proposed model to establish its clinical utility and effect on patient outcomes, as well as explore the role of the model in researching strategies for delivering preventative care for AKI.





← Variability across hospitals →

CORRESPONDENCE



The Clinician and Dataset Shift in Artificial Intelligence

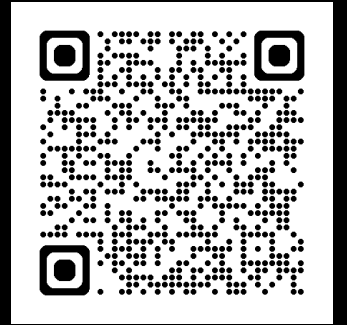


Table 1. Overview of Our Recommended Approach to Recognizing and Mitigating Dataset Shift.*

Dataset Shift Category and Checklist Considerations	Examples of Dataset Shift	Recognition Strategies	Mitigation Strategies
Changes in technology			
Are there new types of data-acquisition devices upstream from the model?	A CAD model developed to predict hip fractures was shown to rely on specific radiographic scanner models and technicians. The adoption of high-sensitivity troponin assays changes clinical interpretation of detectable troponin levels.	Governance committee: For new implementations, check for differences in input-device types between what the model expects and what is being used in the current care environment. For ongoing monitoring, proactively identify when data-acquisition devices or protocols change. Frontline clinicians: Flag when there are changes in data-acquisition protocols.	When new input devices are added, model outputs are checked for validity and models are retrained or tuned if needed.



“Right now, hospitals are overwhelmed by the number of AI models available to them,” Dr. Singh says. To safely use the tools in future, they have to “understand when AI is not working as intended, and prioritize problems based on whether they are solvable rather than simply what AI tools are available.”

A \$22 million donation to UC San Diego Health will establish a mission control center to manage emerging AI



Dr. Christopher Longhurst, medical and digital director at UC San Diego Health, explains plans for a digital command center at the Jacobs Medical Center in La Jolla during a symposium on Friday. (Paul Sisson)

Philanthropists Joan and Irwin Jacobs help pay for center to aggregate streams of digital information



UC San Diego Health

Joan & Irwin Jacobs

Center for Health Innovation

Vision

To be a world leader in the adaptation, development, and implementation of innovative health technologies for the betterment of people's lives

Mission

To cultivate a health technology ecosystem, realizing **innovation at scale with impact** through testing at UC San Diego Health and beyond

UC San Diego Health

Joan & Irwin Jacobs

Center for Health Innovation

Mission Control

Patient flow,
quality/safety
surveillance,
avoiding
hospitalizations

Digital Health



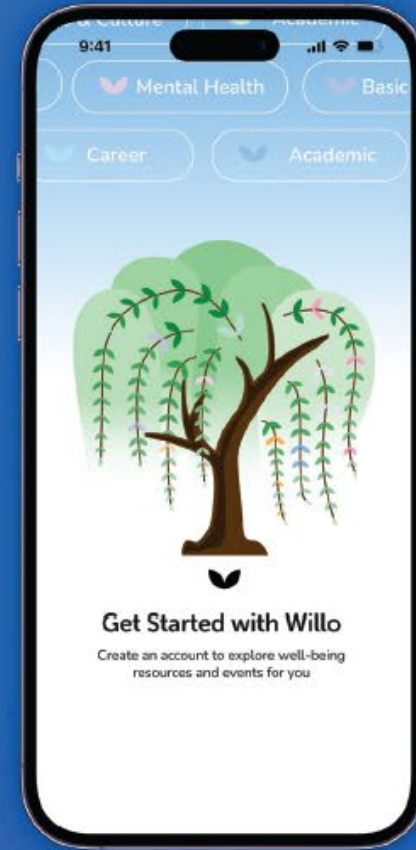
AI Portfolio

Predictive and
generative AI for
admin and
clinical uses

Mission control prototype launched



willo ^{© 2024}



<https://willo.ucsd.edu>

Pushing the field of predictive health AI

Develop and
evaluate
AI models



Usability of AI
workflows



Implement
and evaluate
AI workflows

Pushing the field of predictive health AI

Develop and
evaluate
AI models



Usability of AI
workflows



Implement
and evaluate
AI workflows

TRIPOD+AI
STARD-AI
PROBAST-AI

DECIDE-AI

SPIRIT-AI
CONSORT-AI

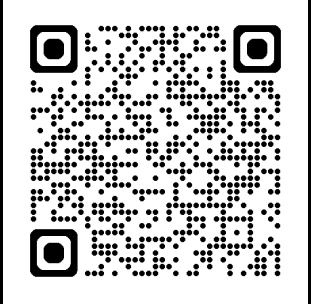
EQUATOR Network reporting guidelines

PERSPECTIVE

A Call for Artificial Intelligence Implementation Science Centers to Evaluate Clinical Effectiveness






Christopher A. Longhurst , M.D., M.S.,¹ Karandeep Singh , M.D., M.M.Sc.,¹ Aneesh Chopra , M.P.P.,²
Ashish Atreja , M.D., M.P.H.,³ and John S. Brownstein , Ph.D.^{4,5}

Received: March 1, 2024; Revised: May 7, 2024; Accepted: May 16, 2024; Published: July 10, 2024



PERSPECTIVE

The Chief Health AI Officer — An Emerging Role for an Emerging Technology

Ashley N. Beecy , M.D.,¹ Christopher A. Longhurst , M.D., M.S.,² Karandeep Singh , M.D., M.M.Sc.,²
Robert M. Wachter , M.D.,³ and Sara G. Murray , M.D., M.A.S.⁴

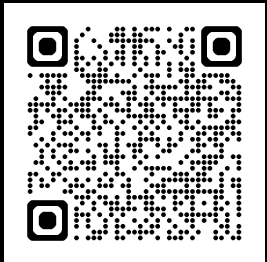
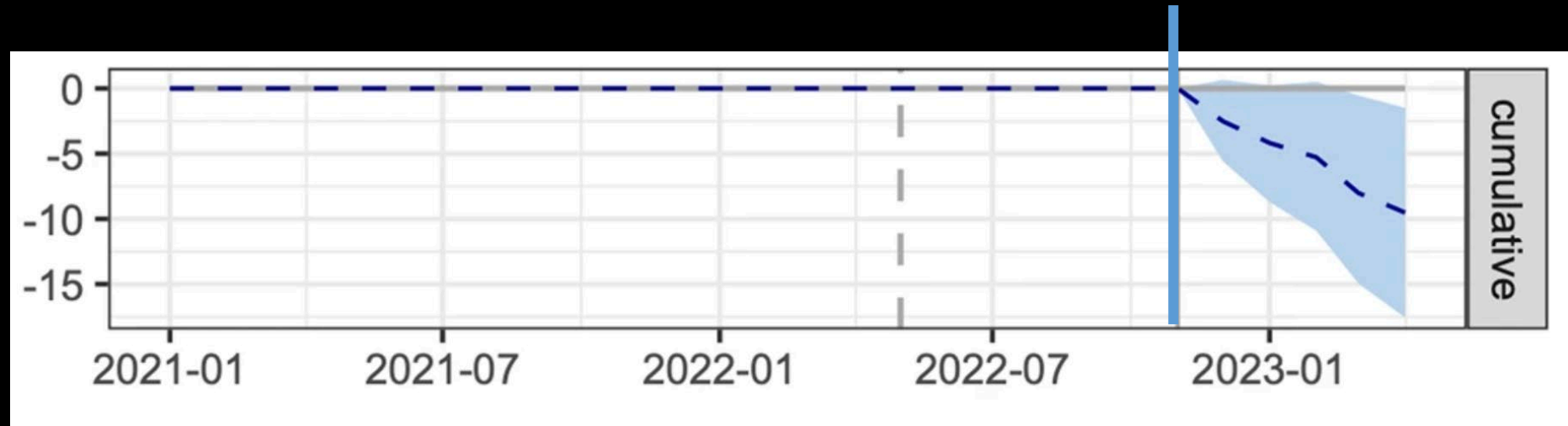
Received: January 31, 2024; Revised: March 18, 2024; Accepted: April 2, 2024; Published: June 17, 2024



Observational study shows fewer sepsis deaths

Sepsis AI model implemented

Observed –
expected
mortality



Pushing the field of generative health AI

Train a foundation model on the entirety of the internet.

Use *foundation* model to generate advice



Implement and evaluate AI workflows

AI Copilots



Incorporate external information, consider multiple perspectives

First randomized study of generative AI to reply to patient messages

JAMA Network | **Open**[™]



Original Investigation | Health Informatics

AI-Generated Draft Replies Integrated Into Health Records and Physicians' Electronic Communication

Ming Tai-Seale, PhD, MPH; Sally L. Baxter, MD, MSc; Florin Vaida, PhD; Amanda Walker, MS; Amy M. Sitapati, MD; Chad Osborne, MD; Joseph Diaz, MD; Nimit Desai, BS; Sophie Webb, MS; Gregory Polston, MD; Teresa Helsten, MD; Erin Gross, MD; Jessica Thackaberry, MD; Ammar Mandvi, MD; Dustin Lillie, MD; Steve Li, MD; Geneen Gin, DO; Suraj Achar, MD; Heather Hofflich, DO; Christopher Sharp, MD; Marlene Millen, MD; Christopher A. Longhurst, MD, MS

Abstract

IMPORTANCE Timely tests are warranted to assess the association between generative artificial intelligence (GenAI) use and physicians' work efforts.

OBJECTIVE To investigate the association between GenAI-drafted replies for patient messages and physician time spent on answering messages and the length of replies.

DESIGN, SETTING, AND PARTICIPANTS Randomized waiting list quality improvement (QI) study from June to August 2023 in an academic health system. Primary care physicians were randomized to an immediate activation group and a delayed activation group. Data were analyzed from August to November 2023.

Key Points

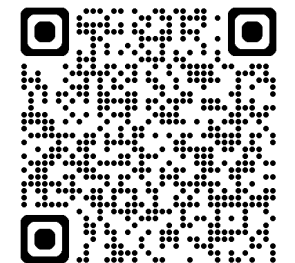
Question Would access to generative artificial intelligence–drafted replies correlate with decreased physician time on reading and replying to patient messages, alongside an increase in reply length?

Findings In this quality improvement study including 122 physicians, generative AI-drafted replies correlated with increased message read time, no change in reply time, and significantly

Use of GenAI resulted in:

21.8% increase in read time

17.9% increase in reply length



First health system to be transparent about use of generative AI

The New York Times

HEALTH | That Message From Your Doctor? It May Have Been Drafted by AI

Share full article

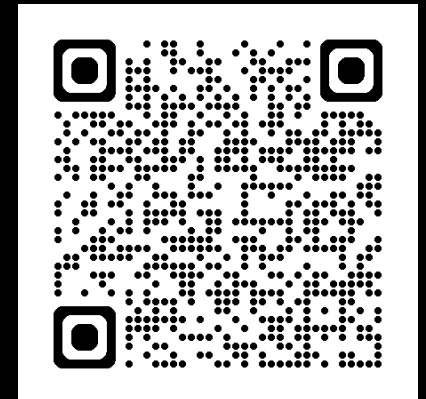


In the absence of strong [federal regulations](#) or widely accepted ethical frameworks, each health system decides how to test the tool's safety and whether to inform patients about its use.

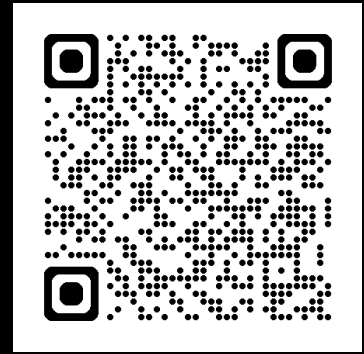
Some hospital systems, like U.C. San Diego Health, put a disclosure at the bottom of each message explaining that it has been “generated automatically,” and reviewed and edited by a physician.

“I see, personally, no downside to being transparent,” said Dr. Christopher Longhurst, the health system's chief clinical and innovation officer.

Patients have generally accepted the new technology, he said. (One doctor received an email saying, “I want to be the first to congratulate you on your A.I. co-pilot and be the first to send you an A.I.-generated patient message.”)



First study using generative AI for quality measurement



Question Measure	Human/AI Agreement
Was severe sepsis present?	98%
When was the last criterion met to establish the presence of severe sepsis?	97%
Is there documentation that the patient or authorized patient advocate refused either a blood draw, IV fluid administration, or IV antibiotic administration within the specified time frame?	100%
Is there physician/APN/PA documentation of comfort measures only, palliative care, or another inclusion term before or within six hours after the presentation of severe sepsis?	99%

Tackling the health AI paradox

Help
researched
models get
implemented.

UC San Diego Health
Joan & Irwin Jacobs
Center for Health Innovation

Help
implemented
models get
researched.

Develop cross-campus collaborations

Enable our faculty, staff, students, and trainees

We are just getting started

Clinical AI

Ambient documentation
HIPAA-compliant GPT

Mission Control AI

Identifying discharge barriers
Predicting length of stay

Operations AI

Augmenting administrative tasks
Real-time operations advice

Gratitude

The late Joan Jacobs
Irwin Jacobs
Pradeep Khosla
Patty Maysent
Chris Longhurst
Amy Sitapati
Nicole May
Jeff Pan
Zea Borok
Marlene Millen
Josh Glandorf



JCHI staff and
students

My family

...and so many
other students,
trainees, staff, and
faculty across
campus for the
warm welcome.



THANK YOU!