AI for Impact: Building a Highly Reliable Health System

Christopher Longhurst, MD, MS

Chief Clinical & Innovation Officer, UC San Diego Health

Executive Director, Jacobs Center for Health Innovation

Professor and Associate Dean, UC San Diego School of Medicine

UC San Diego Health

UC San Diego Health: World-Class Care

~13k
Employees

808
Licensed Beds

36k+
Annual Hospital Discharges

1.3M+
Annual Outpatient Visits

THE LEAPFROG GROUP
TOP TEACHING HOSPITAL 2022

MAGNET RECOGNIZED
AMERICAN NURSES CREDENTIALING CENTER

NCI
Designated Comprehensive Cancer Center

BEST HOSPITALS
U.S. News
HONOR ROLL

Patty Maysent, MPH
Chief Executive Officer
UC San Diego Health

2023 Honor Roll Listing
Mayo Clinic, Rochester, Minn.
UCLA Medical Center, Los Angeles
NYU Langone Hospitals, New York
Cedars-Sinai Medical Center, Los Angeles
New York-Presbyterian Hospital-Columbia and Cornell
Stanford Health Care-Stanford Hospital, Stanford, Calif.
Cleveland Clinic
Massachusetts General Hospital, Boston
Mount Sinai Hospital, New York
Johns Hopkins Hospital, Baltimore
Hospitals of the University of Pennsylvania-Penn Presbyterian
UCSF Health-UCSF Medical Center, San Francisco, Calif.
Northwestern Medicine-Northwestern Memorial Hospital, Chicago
Houston Methodist Hospital
Brigham and Women's Hospital, Boston
Rush University Medical Center, Chicago
Barnes-Jewish Hospital, Saint Louis
UC San Diego Health-La Jolla and Hillcrest Hospitals, San Diego
Vanderbilt University Medical Center, Nashville, Tenn.
North Shore University Hospital at Northwell Health, Manhasset, N.Y.
University of Michigan Health-Ann Arbor
UT Southwestern Medical Center, Dallas

Traditional patient safety efforts have not moved the needle

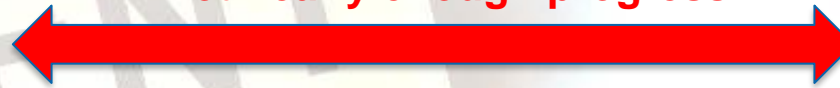
Not nearly enough progress

3.7% of inpatients had adverse events with harm

~3% of inpatients had adverse events and ~50% preventable

~25% of hospitalizations had adverse events and ~25% preventable

Traditional patient safety efforts

Harvard Medical Practice Study
NEJM, 1991
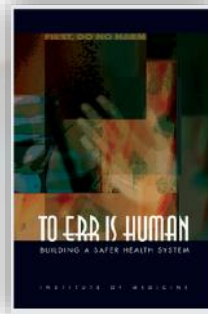
To Err is Human
IOM
1999

The Safety of Inpatient Health Care
NEJM
2023

npj | digital medicine

www.nature.com/npjdigitalmed

COMMENT   **OPEN** can AI help?

Bending the p

David C. Classen, Christo

This paper reviews the curre (AI) techniques to patient safety.
This paper defines patient s care, including diagnostic errors,
misdiagnosis, adverse event ses of AI in patient safety and the
relative adoption of these t the limitations of these AI systems
and the challenges with eva oping a proactive agenda for AI in
healthcare that includes ma

npj Digital Medicine (2023)6

Check for updates

Table 2. Top use cases for the application of AI to specific clinical problems in patient safety.

1. Actionable real time patient safety electronic clinical quality measures
2. Surgical complication prediction
3. Pressure ulcer prediction
4. Hypoglycemia prediction
5. Sepsis prediction
6. Suicide prediction
7. Diabetic eye AI screening
8. Breast imaging cancer screening
9. Chest x-ray imaging AI diagnosis
10. Skin melanoma AI diagnosis
11. Chest x-ray imaging AI cancer screening
12. Patient self-managed electronic safety dashboards

[ADVENTISTHEALTH:INTERNAL]
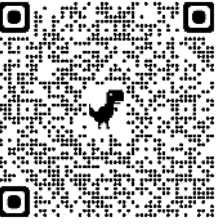
UC San Diego Health

February 2020 – Wuhan Evacuees in San Diego

UC San Diego Health

Artificial Intelligence Enables Rapid COVID-19 Lung Imaging Analysis at UC San Diego Health

With support from Amazon Web Services, health care providers are using AI in a clinical research study aimed at speeding the detection of pneumonia, a condition associated with severe COVID-19

April 07, 2020 | Heather Buschman, PhD

For most patients who have died of COVID-19, the pandemic disease caused by a novel coronavirus, the ultimate cause of death was pneumonia, a condition in which inflammation and fluid buildup make it difficult to breathe. Severe pneumonia often requires lengthy hospital stays in intensive care units and assistance breathing with ventilators — medical devices now in high demand in some cities grappling with a surge of COVID-19 cases.

To quickly detect pneumonia — and therefore better distinguish between COVID-19 patients likely to need more supportive care in the hospital and those who could be followed closely at home — UC San Diego Health radiologists and other physicians are now using artificial intelligence (AI) to augment lung imaging analysis in a clinical research study enabled by Amazon Web Services (AWS).

The new AI capability has so far provided UC San Diego Health physicians with unique insights into more than 2,000 images. In one case, a patient in the Emergency Department

UC San Diego uses AWS cloud to deploy homegrown AI algorithms for COVID-19

The health system's algorithms can detect early pneumonia on medical imaging with the precision of a subspecialist radiologist.

By Bill Siwicki | June 24, 2020 | 11:57 AM
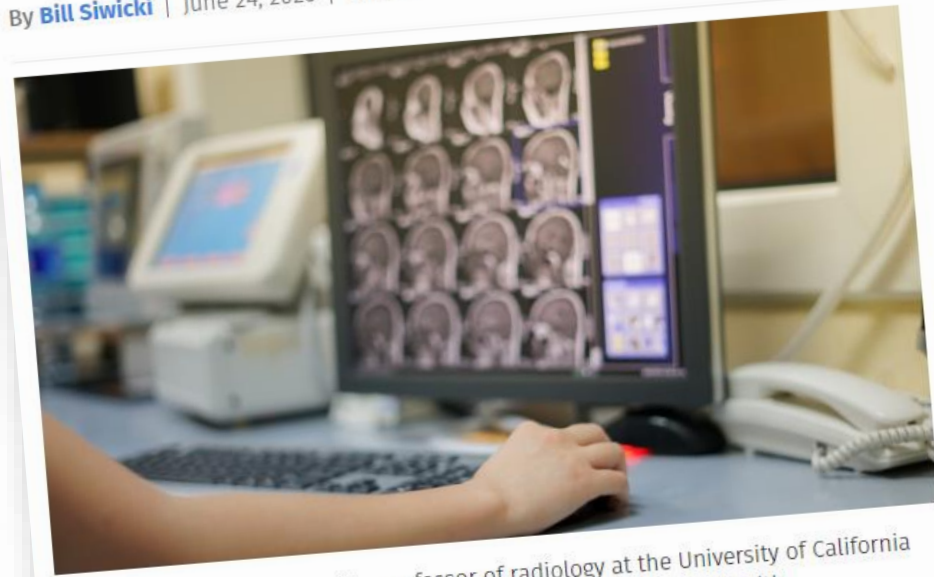
Dr. Albert Hsiao is an associate professor of radiology at the University of California San Diego School of Medicine and a radiologist at UC San Diego Health.

UC San Diego Health

JACEP OPEN
WILEY

BRIEF RESEARCH REPORT

Infectious Disease

Deployment of artificial intelligence for radiographic diagnosis of COVID-19 pneumonia in the emergency department

Christian Dameff, MD
Emergency Medicine

Morgan Carlile MD[1] | Brian Hurt MD, MS[2] | Albert Hsiao MD, PhD[2] | ... Dameff MD, MS[1,3]

"Of the 5,125 total visits and 1,960 chest radiographs obtained in the ED during the study period, 1,855 were analyzed by the algorithm. Among these, emergency physicians were surveyed for their experiences on 202. Overall, **86% either strongly agreed or somewhat agreed that the intervention was easy to use in their workflow. 20% of respondents reported that the algorithm impacted clinical decision making.**"

Correspondence
Christian Dameff, MD, MS, UC San Diego Health Department of Emergency Medicine, 200 W Arbor Drive MC #8676, San Diego, CA 92103, USA.
Email: cdameff@health.ucsd.edu

UC San Diego Health
UC San Diego Health

Artificial Intelligence Functionalities During the COVID-19 Pandemic

Milad Ahmadi Marzaleh PhD[1], Mahmoudreza Peyravi PhD[1], Shahrokh Mousavi MD[2], Fatemeh Sarpourian PhD[3], Milad Seyedi BSC[4] and Naseh Shalyari PhD[1]

[1]Department of Health in Disasters and Emergencies, Health Human Resources Research Center, School of Health Management and Information Sciences, Shiraz University of Medical Sciences, Shiraz, Iran; [2]Student Research Committee, School of Medicine, Shiraz University of Medical Sciences, Shiraz, Iran; [3]Student Research Committee, Department of Health Information Technology, School of Health Management and Information Sciences, Shiraz University of Medical Sciences, Shiraz, Iran and [4]Department of Computer Engineering, Sanandaj Branch, Islamic Azad University, Sanandaj, Iran

Abstract

Background: The coronavirus disease 2019 (COVID-19) pandemic has led us to use virtual solutions and emerging technologies such as artificial intelligence (AI). Recent studies have clearly demonstrated the role of AI in health care and medical practice; however, a comprehensive review can identify potential yet not fulfilled functionalities of such technologies in pandemics. Therefore, this scoping review study aims at assessing AI functionalities in the COVID-19 pandemic in 2022.

Methods: A systematic search was carried out in PubMed, Cochran Library, Scopus, Science Direct, ProQuest, and Web of Science from 2019 to May 9, 2022. Researchers selected the articles according to the search keywords. Finally, the articles mentioning the functionalities of AI in the COVID-19 pandemic were evaluated. Two investigators performed this process.

Results: Initial search resulted in 9123 articles. After reviewing the title, abstract, and full text of these articles, and applying the inclusion and exclusion criteria, 4 articles were selectd for the final analysis. All 4 were cross-sectional studies. Two studies (50%) were performed in the United States, 1 (25%) in Israel, and 1 (25%) in Saudi Arabia. They covered the functionalities of AI in the prediction, detection, and diagnosis of COVID-19.

Over **9000** articles…

…just **4** studies w/ clinical outcomes!!

[ADVENTISTHEALTH:INTERNAL]

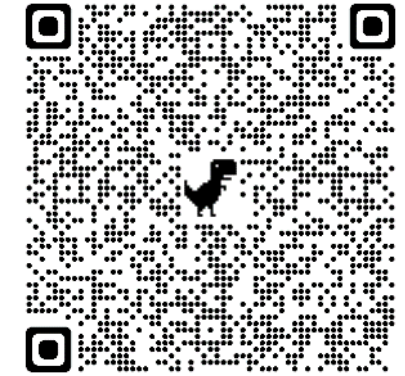UC San Diego Health
UC San Diego Health

We need more clinical outcome evaluation for AI
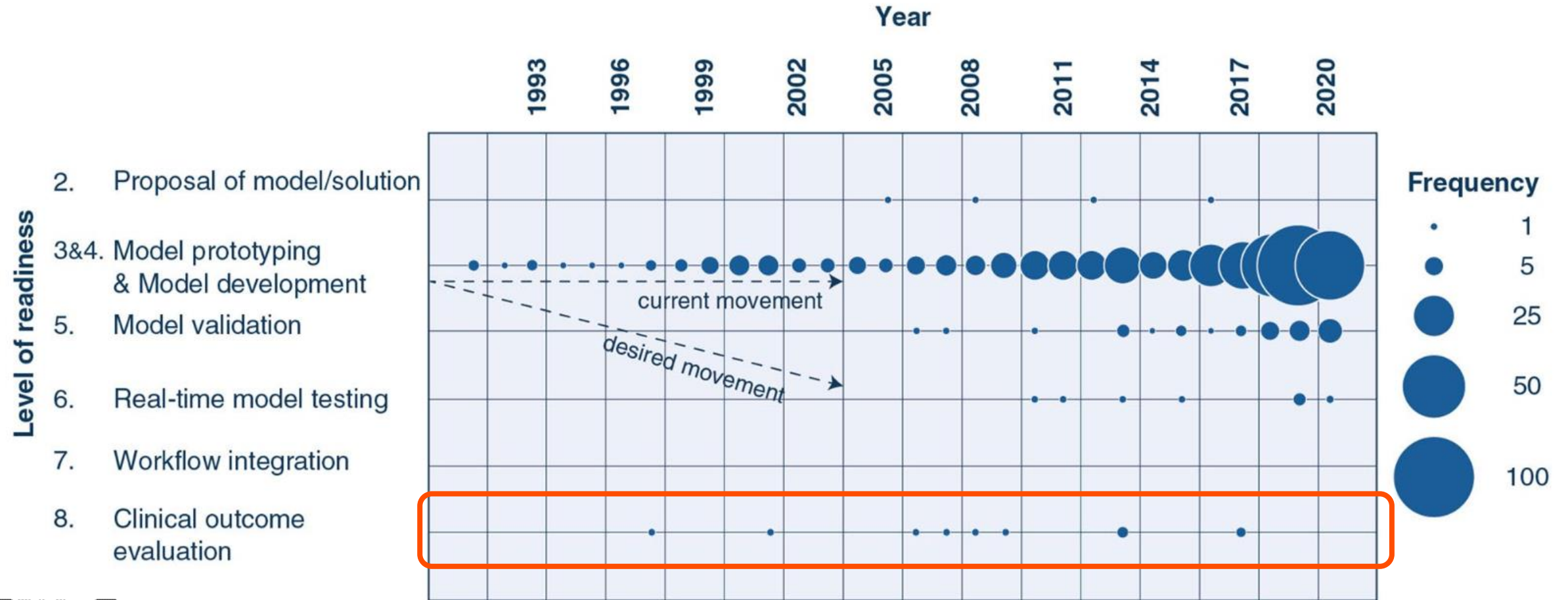
van de Sande D, et al. *Intensive Care Med*. 2021 Jul;47(7):750-760.
doi: 10.1007/s00134-021-06446-7.

JAMA, October 2023

RESEARCH LETTER

Analysis of Devices Authorized by the FDA for Clinical Decision Support in Critical Care

The use of predictive clinical decision support (CDS) devices (ie, those that use machine learning [ML] or artificial intelligence [AI]) has the potential to improve outcomes in critical care, but a clear regulatory framework is lacking.[1] Recent guidance from the U[...]

Administr[...]

gests most [...]

cal illness w[...]

cause of th[...]

nature of t[...]

Editor's Note page 1401

+

Supplemental content

formed by these devices. However, growi[...] the clinical impact of predictive CDS syste[...] about whether current device regulatory fr[...] oped before advanced statistical learning me[...] available, are sufficient to ensure effectiven[...]

On September 22, 2021, the FDA r[...] database of authorizations for medical de[...] or AI. We sought to identify devices that offer CDS in a critical care setting and characterize the evidence cited in their authorization.

Methods | We extracted data from the AI and ML database as of December 15, 2022, and augmented those data through the OpenFDA interface.[3,4] The 2 most common FDA pathways for CDS device approval are the 510(k) pathway, which requires demonstration of substantial equivalence to a previously authorized device (hereafter, a predicate) and typically does not require submission of clinical data, and the de novo pathway, which indicates that a novel device has low to moderate risk and

identified 10 that might inform care for patients with critical illness (**Table 1**). Of these, only 3 included citations of published data, 4 mentioned a safety assessment, and none men-

> Of 521 authorizations in the FDA AI/ML database…only 3 included citations of published data, 4 mentioned a safety assessment, and none mentioned an evaluation of performance bias. **We found no studies examining the clinical impact on care processes or patient outcomes for these device authorizations."**

npj | digital medicine

www.nature.com/npjdigitalmed

COMMENT   **OPEN**

Bending the p... can AI help?

David C. Classen [1] ✉, Christo...

This paper reviews the curre... e (AI) techniques to patient safety.
This paper defines patient s... care, including diagnostic errors,
misdiagnosis, adverse event... ses of AI in patient safety and the
relative adoption of these t... the limitations of these AI systems
and the challenges with eva... oping a proactive agenda for AI in
healthcare that includes ma...

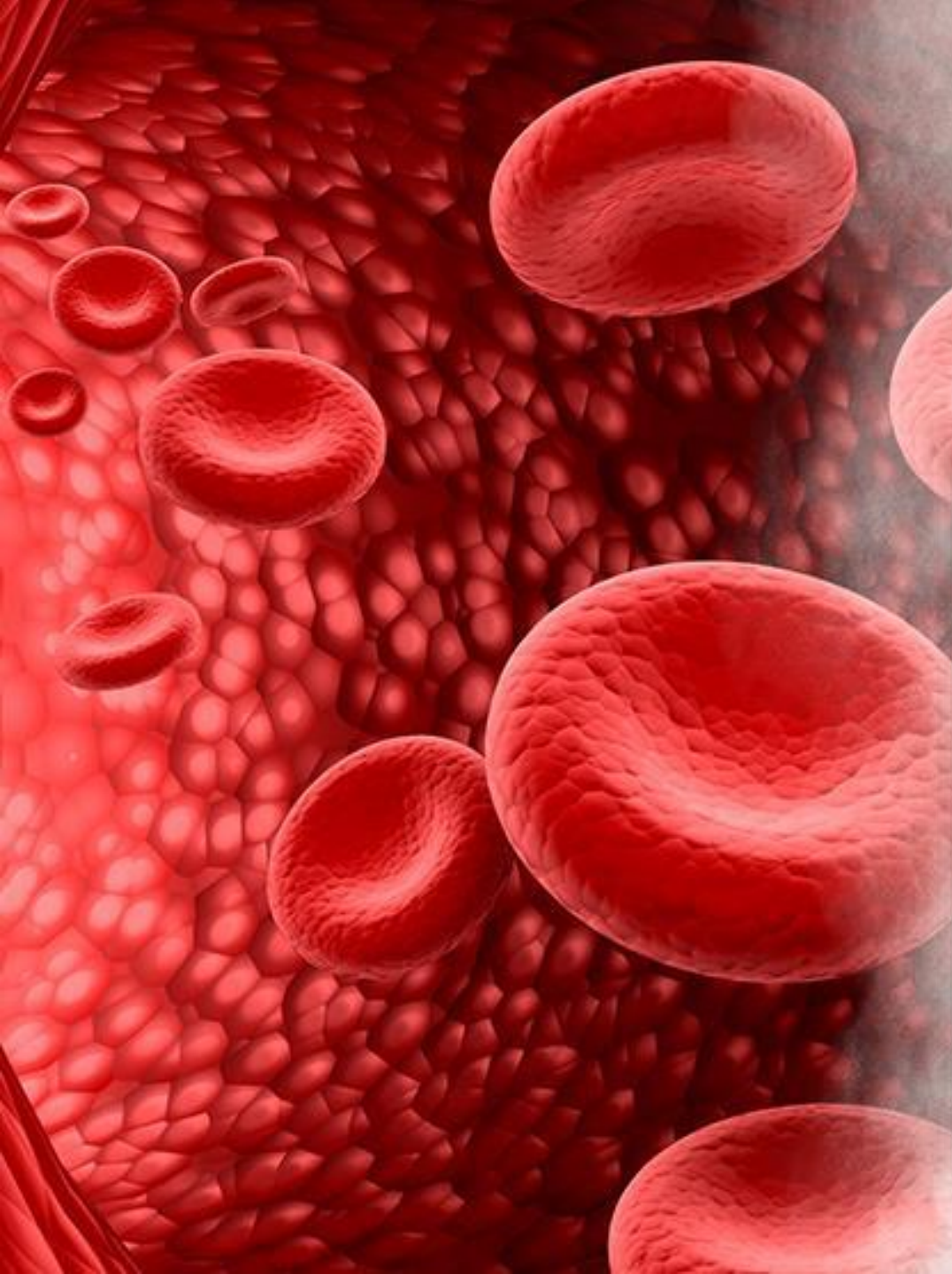*npj Digital Medicine* (2023)6...

Check for updates

Table 2. Top use cases for the application of AI to specific clinical problems in patient safety.

1. Actionable real time patient safety electronic clinical quality measures
2. Surgical complication prediction
3. Pressure ulcer prediction
4. Hypoglycemia prediction
5. Sepsis prediction ⟵
6. Suicide prediction
7. Diabetic eye AI screening
8. Breast imaging cancer screening
9. Chest x-ray imaging AI diagnosis
10. Skin melanoma AI diagnosis
11. Chest x-ray imaging AI cancer screening
12. Patient self-managed electronic safety dashboards

UC San Diego Health

EACH YEAR, ABOUT 1.7 MILLION ADULTS IN AMERICA DEVELOP SEPSIS. AT LEAST 350,000 DIE DURING THEIR HOSPITALIZATION OR ARE DISCHARGED TO HOSPICE.

GET AHEAD OF SEPSIS
KNOW THE RISKS. SPOT THE SIGNS. ACT FAST.

Learn more at cdc.gov/sepsis

New England Journal of Medicine AI, 2024

ORIGINAL ARTICLE

Evaluation of Sepsis Prediction Models before Onset of Treatment

Authors: Fahad Kamran, Ph.D. , Donna Tjandra, M.S. , Andrew Heiler, M.B.A. , Jessica Virzi, M.S.N. , Karandeep Singh, M.D. , Jessie E. King, M.D., Ph.D. , Thomas S. Valley, M.D., M.Sc. , and Jenna Wiens, Ph.D.

Author Info & Affiliations

| VOL. 1 NO. 3

"Designers need to consider that model features could inadvertently encode clinical suspicion, thus suggesting strong performance in retrospective studies but failing to identify new cases that a clinician has not yet recognized."

Nature Digital Medicine (2023 and 2024)

npj | Digital Medicine

www.nature.com/npjdigitalmed

ARTICL...

Arti...

"I d...

Supreeth...

Impact of a deep learning sepsis prediction model on quality of care and survival

Aaron Boussina [1,4], Supreeth P. Shashik..., Christopher A. Longhurst [1,2], Kimberly..., Gabriel Wardi [1,3,5]

Sepsis... improve outcomes, but relatively few... administ... assess the impact of a deep-learning... Multidi... before-and-after quasi-experimental st... false al... We included 6217 adult septic patient... COMPO... Advisory (BPA) triggered by COMPOSE... patient... assessment (SOFA) score following se... (ED) we... intervention period (705 days) and the... consist... Bayesian structural time-series appro... 20% ar...

Emergency (1)

(!) Could it be Sepsis?
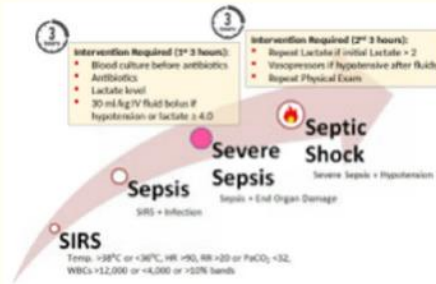
This patient has a Sepsis Risk Score: 90% chance of developing severe sepsis in the next 4 hours.

Consider discussing risk of sepsis with the primary physician or activating Code Sepsis

Top reasons in the past 6 hours
Sepsis Top Causes: Temperature, Heart Rate

The deployment of COMPOSER was significantly associated with a 17% relative decrease in in-hospital sepsis mortality (95% CI, 0.3%–3.5%) and a 10% relative increase in sepsis bundle compliance

UC San Diego Health

Nature Digital Medicine (2024)

npj | digital medicine
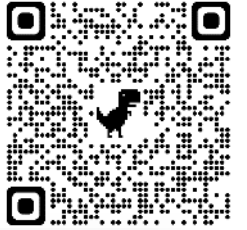
Integrating artificia... healthcare system... algorithm

Boussina et al. recently evaluated a deep learning sepsis prediction model (COMPOSER) in a prospective before-and-after quasi-experimental study within two emergency departments at UC San Diego Health, tracking outcomes before and after deployment. Over the five-month implementation period, they reported a 17% relative reduction in in-hospital sepsis mortality and a 10% relative increase in sepsis bundle compliance.

This editorial discusses the importance of shifting the focus towards evaluating clinically relevant outcomes, such as mortality reduction or quality-of-life

Table 1 | Considerations for implementing AI algorithms into healthcare systems

Theme	Key considerations	How each issue was addressed by Boussina et al.[7]
Data	Are the data needed for the algorithm readily available and in an extractable format?	COMPOSER routinely collected clinical information, including laboratory and vital signs. Data elements were extracted via FHIR standards.
Infrastructure	Can relevant data be extracted in real time?	Data were extracted at hourly intervals to ensure availability for prediction.
	Are there adequate infrastructure and computing resources available to host a cloud-based analytics and storage platform?	The platform was hosted via Amazon Web Services.
Interface	How will the clinical team be made aware of these predictions (i.e., is a custom dashboard integrated into the electronic health record required)?	Predictions were integrated into an Epic flowsheet via an HL7v2 outbound message. A Best Practice Advisory was triggered for patients at high risk of developing sepsis.
	How can the clinical team understand how the algorithm made this prediction (i.e., model explainability)?	A relevance score was generated for each feature, which measured the gradient of the risk score with respect to all input features multiplied by the input features. The features with the highest positive relevance scores were displayed in the flowsheet.
End-users	Which clinical team member(s) are most appropriate to receive the risk prediction?	Nurses were chosen to receive the alert as they cared for a specific roster of patients and frequently opened their patients' charts.
	What is their level of trust in the AI algorithm?	A multidisciplinary team was created to guide implementation. Nurses were surveyed to identify their needs. Regular feedback and educational sessions on COMPOSER were provided to nurses during the implementation phase.
	What happens if the compliance of clinical end-users degrades over time?	Although not specified, the authors acknowledged the importance of continuous education to optimize human-AI collaboration.

Examples of how these issues were addressed by Boussina et al. are provided

> AI researchers should **shift the focus from measuring just algorithm performance metrics such as accuracy to meaningful outcomes.** As a healthcare and AI community, we have a responsibility to deliver on these clinically relevant metrics, and **funding agencies and journals alike should be encouraged to prioritize such studies.**"

PRENOSIS

Prenosis Anno...
Marketing Aut...
ImmunoScore™...ficial
Intelligence Di... Tool
For Early Sepsi...

The Sepsis ImmunoScore™ is the first-ever AI S...

Shamim Nemati @ShamimNemati · 20h

2/5 Conditions for use: There needs to exist a clinical suspicion for sepsis (ordering of blood culture), and three plasma proteins (procalcitonin, interleukin-6, and C-reactive protein) and routinely measured EHR parameters to calculate the ImmunoScore for sepsis.

💬 1   🔁   ♡ 1   �'ıl 54   🔖 ↥

Shamim Nemati @ShamimNemati · 20h

3/5 The three biomarkers are not routinely ordered/available in the EHR. Therefore, once sepsis is suspected by the provider and a blood culture is ordered, additional orders of the three biomarkers are required.

💬 1   🔁   ♡ 1   �'ıl 47   🔖 ↥

Shamim Nemati @ShamimNemati · 20h

4/5 In an ER setting, ordering of blood culture is often followed by administration of antibiotics, therefore it's not clear if providers will wait for a AI risk score to initiate treatments for sepsis. Moreover, all cases without a clinical suspicion of sepsis will be missed!

💬 1   🔁 2   ♡ 1   ⑊ıl 163   🔖 ↥

Shamim Nemati @ShamimNemati · 20h

5/5 I'm not aware of any patient-centered outcome data to demonstrate clinical utility of this AI system (would love to hear more!). This JAMA view point argues for a regulatory refocus on patient outcomes:

UC San Diego Health

JAMA, January 2024

AI IN MEDICINE

VIEWPOINT

Regulate Artificial Intelligence in Health Care by Prioritizing Patient Outcomes

John W. Ayers, PhD, MA
Qualcomm Institute, University of California San Diego, La Jolla; and Altman Clinical Translational Institute, California La Jolla.

Nimit De
Qualcomm University San Diego School of University San Diego

Davey M. Smith, MD, MAS
Altman Clinical Translational Research Institute, University of California San Diego, La Jolla; and Division of Infectious Diseases and Global Public Health,

Artificial intelligence (AI) holds great promise to enhance the quality of health care.[1] Despite the aim of health care regulations to ensure high-quality care, there are few regulations governing the use of AI in health care. This dearth will soon change because a White House ex-

instance, it calls for AI-powered algorithms with more than 10 billion parameters to potentially register their codebase with federal agencies, perform a battery of to-be-determined regulator-selected tests, and report codebase modifications for review. Concurrently,

> "We propose a regulatory strategy for AI that is outcome-centric by requiring companies to demonstrate that AI tools produce clinically important differences in patient outcomes before being brought to market."

and treat problems as "teachable moments" to make rules that mitigate the likelihood of similar problems happening again. These regulations typically mandate compliance with specific procedures. We refer to these as "process-centric" regulations because they manipulate

effective, but despite a growing marketplace, these systems have not been required to produce such evidence. A third-party evaluation of the most widely adopted system, the Epic Sepsis Model, found that among 2552 hospitalized patients who developed sepsis, only 7% who did not already receive early treatment were identified by the

UC San Diego Health

I think…AI Governance Should be a Shared Responsibility
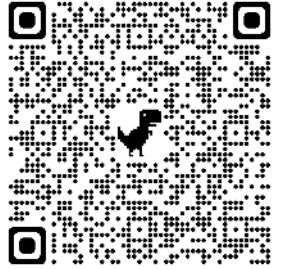
Epic plans to launch AI validation software for healthcare organizations to test, monitor models

By Heather Landi · Apr 3, 2024 8:00am

- "The only way to achieve outcomes is going to be with local workflows and local optimization. It's a shared responsibility between both vendors and local health systems"

- "Dr. Longhurst believes the federal government should go a step further and incorporate responsible AI use into conditions of participation in Medicare. In the same way that hospitals and health systems are audited for quality and safety, organizations also should be required to meet compliance requirements around AI governance, he said."

UC San Diego Health

AI Principles at UC San Diego Health

Artificial Intelligence

Our Statement

1. We believe that AI can enhance human health and well-being, and we are committed to developing and deploying AI solutions that are ethical, responsible, and beneficial for patients and society.
2. We respect the dignity, autonomy, and privacy of each patient, and we design and evaluate our AI systems with their needs, preferences, and feedback in mind.
3. We uphold the highest standards of scientific rigor, transparency, and accountability in our AI research and practice, and we adhere to the ethical codes and regulations of our profession and institution.
4. We foster a culture of collaboration, excellence, and innovation among our AI researchers, practitioners, partners, and stakeholders, and we seek to share our knowledge and expertise with the broader community.
5. We embrace the diversity of our patients, staff, and collaborators, and we strive to create an inclusive and supportive environment that values different perspectives, backgrounds, and experie...

Effectiveness, Responsibility, and Accountability

AI products should be successful for their intended use, beneficial, valid, accurate, and reliable and those who develop and apply tools with responsibility and accountability

Ethics and Inclusiveness

AI products should strive to achieve health equity and fairness by design and operation

Human Factors

AI product design, development, and implementation should involve and prioritize the needs of the diverse population it serves

Promoting Human Well-Being, Safety, Privacy, and Common Good

AI products should protect human well-being, privacy, sustainability, and the environment

---

DECEMBER 14, 2023

Delivering on the Promise of AI to Improve Health Outcomes

🏛 ▸ BRIEFING ROOM ▸ BLOG

Lael Brainard, National Economic Advisor
Neera Tanden, Domestic Policy Advisor
Arati Prabhakar, Director of the Office of Science and Technology Policy

———

As President Biden has said, artificial intelligence (AI) holds tremendous ...ruer than healthcare. The ...dmark Executive Order on ...tion is committed to ...opment and use of AI ...comes for Americans

---

In its blueprint for an AI Bill of Rights, the Biden administration outlined that AI **"testing conditions should mirror as closely as possible the conditions in which the AI will be deployed**." The administration has also recommended that **AI should be monitored for adverse outcomes on an ongoing basis rather than a single, off-site evaluation.**

# *JAMA Internal Medicine*, April 2023

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

IMPORTANCE The rapid exp[...]
messages concomitant with[...]
Artificial intelligence (AI) as[...]
questions by drafting respo[...]

OBJECTIVE To evaluate the[...]
2022, to provide quality an[...]

DESIGN, SETTING, AND PART[...]
database of questions from[...]
randomly draw 195 exchang[...]
public question. Chatbot re[...]
fresh session (without prior[...]
23, 2022. The original ques[...]
chatbot responses were ev[...]
professionals. Evaluators c[...]
information provided" (*ver*[...]
bedside manner provided"[...]
empathetic, and *very empa*[...]
compared between chatbo[...]

UC San Diego Health
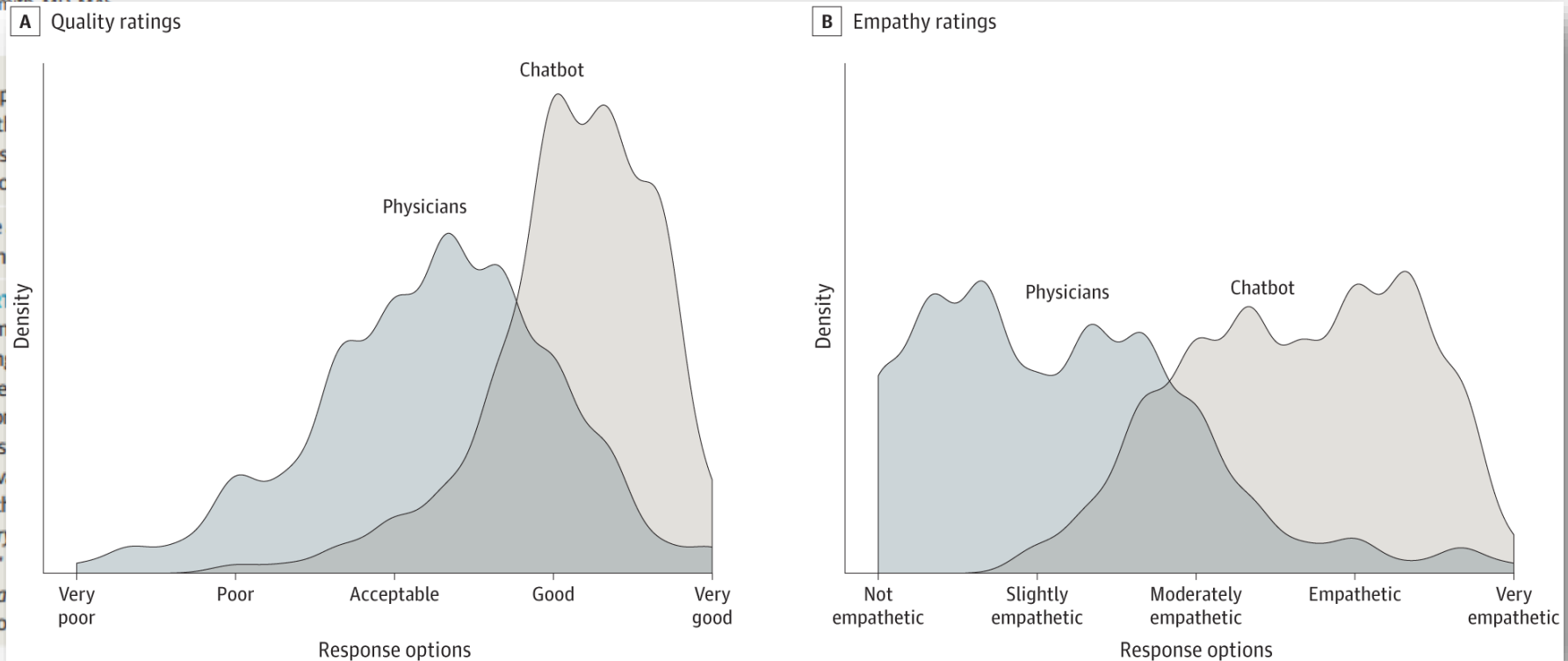
UC San Diego Health

AI Principles in Action – April 2023

HEALTH TECH

HIMSS23: Epic taps Microsoft to integrate generative AI into EHRs with Stanford, UC San Diego as early adopters

By **Heather Landi** · Apr 17, 2023 11:00am

Microsoft   Epic   generative AI   electronic health records (EHRs)

...with the patient at the

Epic

Generated Draft Reply   Pioneering Feature

Hello Barbara,

Your CBC with Auto Diff result from 3/7/2023 is within the normal range and does not indicate any cause for concern. Your provider will contact you if any further action is needed. Your results are available for viewing in MyChart, under the "Test Results" tab.

If you have any further questions or concerns, please do not hesitate to contact us.

Best regards,

Learn more   Start with Draft   Start Blank Reply

Did you find the draft reply helpful?

This was helpful   This was not helpful

Dear

Thank you for your message and for taking an active role in your health. It's great to hear that you're considering the Hepatitis A vaccine. Twinrix is indeed a combination vaccine that covers both Hepatitis A and B. However, since you've already had the Hepatitis B vaccine, you may not need the combination.

I see you have an upcoming appointment and I recommend reviewing at that appointment. It is a 2 shot series so you can do the 1st shot at the appt if you would like.

Take care,
Marlene May Millen, MD

Part of this message was generated automatically and was reviewed and edited by Marlene May Millen, MD.

UC San Diego Health

JAMA, April 2024

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.

EXPOSURE Access to GenAI-drafted replies for patient messages.

MAIN OUTCOMES AND MEASURES Time spent (1) reading messages, (2) replying to messages, (3) length of replies, and (4) physician likelihood to recommend GenAI drafts. The a priori hypothesis was that GenAI drafts would be associated with less physician time spent reading and replying to messages. A mixed-effects model was used.

RESULTS Fifty-two physicians participated in this QI study, with 25 randomized to the immediate activation group and 27 randomized to the delayed activation group. A contemporary control group included 70 physicians. There were 18 female participants (72.0%) in the immediate group and 17 female participants (63.0%) in the delayed group; the median age range was 35-44 years in the

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 longer replies. Physicians valued AI-generated drafts as a compassionate starting point for their replies and also noted areas for improvement.

Meaning These findings suggest generative AI was not associated with reduced time on writing a reply but was associated with longer read time, longer replies, and perceived value in making a more compassionate reply.

Figure 2. Reply Time per Message (in Seconds) for Each Group and Study Period
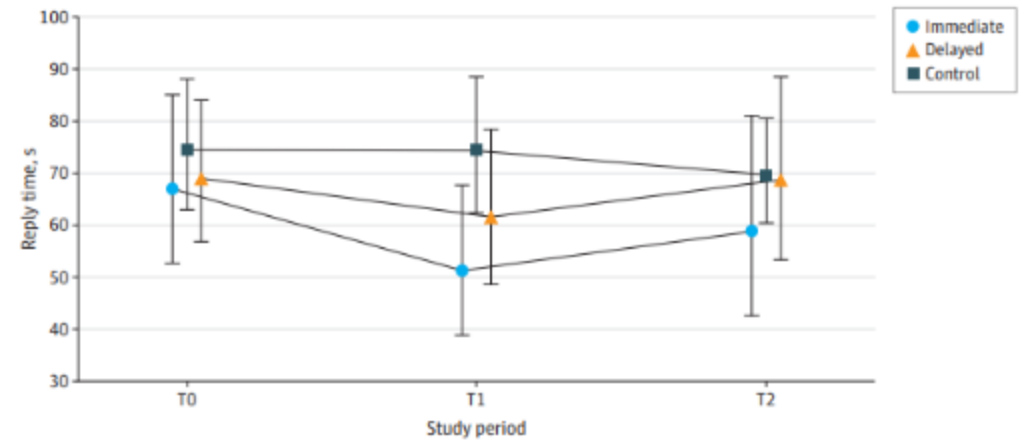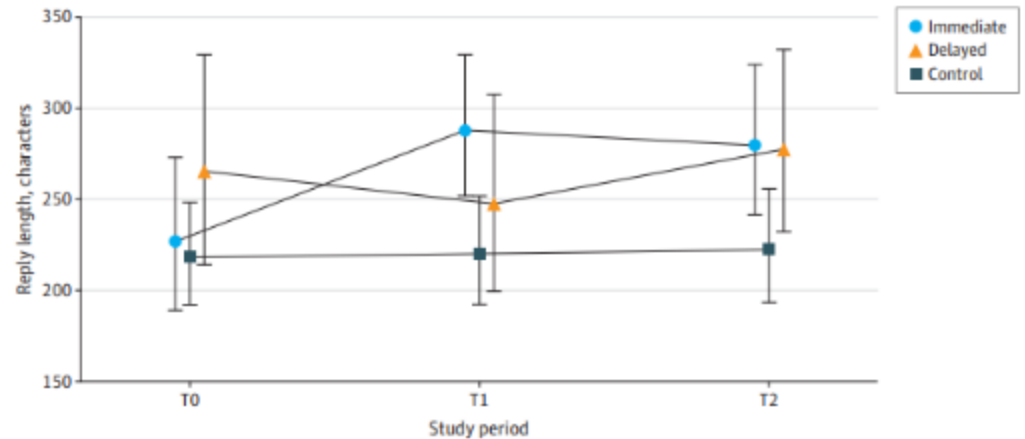

Figure 3. Reply Length for Each Group and Study Period

UC San Diego Health

Overall Efficiency

Practical AI Can Drive Significant Time Savings For Physicians

	Hrs / Wk Today	% Time AI Could Save	Est. AI Hours Saved / Wk
Insurer Appeals	1.8	53%	1.0
Patient Letters Education	3.1	58%	1.8
Faxed Letters Forms	1.7	35%	0.6
EHR Docs Notes	11.5	52%	5.9
Clinical Reference	3.9	51%	2.0
Colleague / Hospital Emails	3.6	35%	1.3
PD: Licensure MOC	2.2	38%	0.8

Source: Doximity user research

12.7
Hours of Work / Week Potentially Saved by AI

UC San Diego Health

Ambient AI Documentation

Key Points

- The health AI paradox is that **implemented algorithms are rarely researched and researched algorithms are rarely implemented**

- The Jacobs Center for Health Innovation at UC San Diego is focused on developing **AI algorithms *and* workflows that impact patient care outcomes**

- All AI efforts are reviewed by our governance to ensure each algorithm is fair, appropriate, valid, effective, and safe (FAVES)

Dr. Longhurst

UC San Diego Health

Thank you!

clonghurst@health.ucsd.edu

@calonghurst