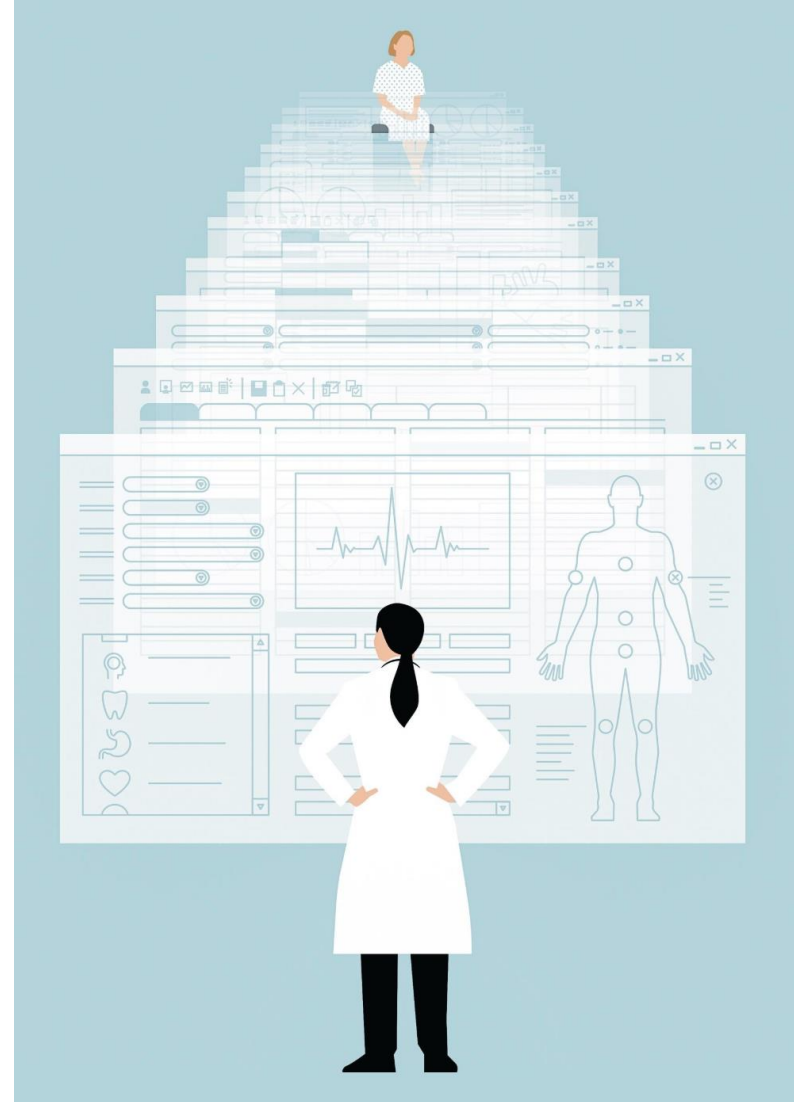




# Digital Empowerment: Using AI to Reduce Physician Burnout

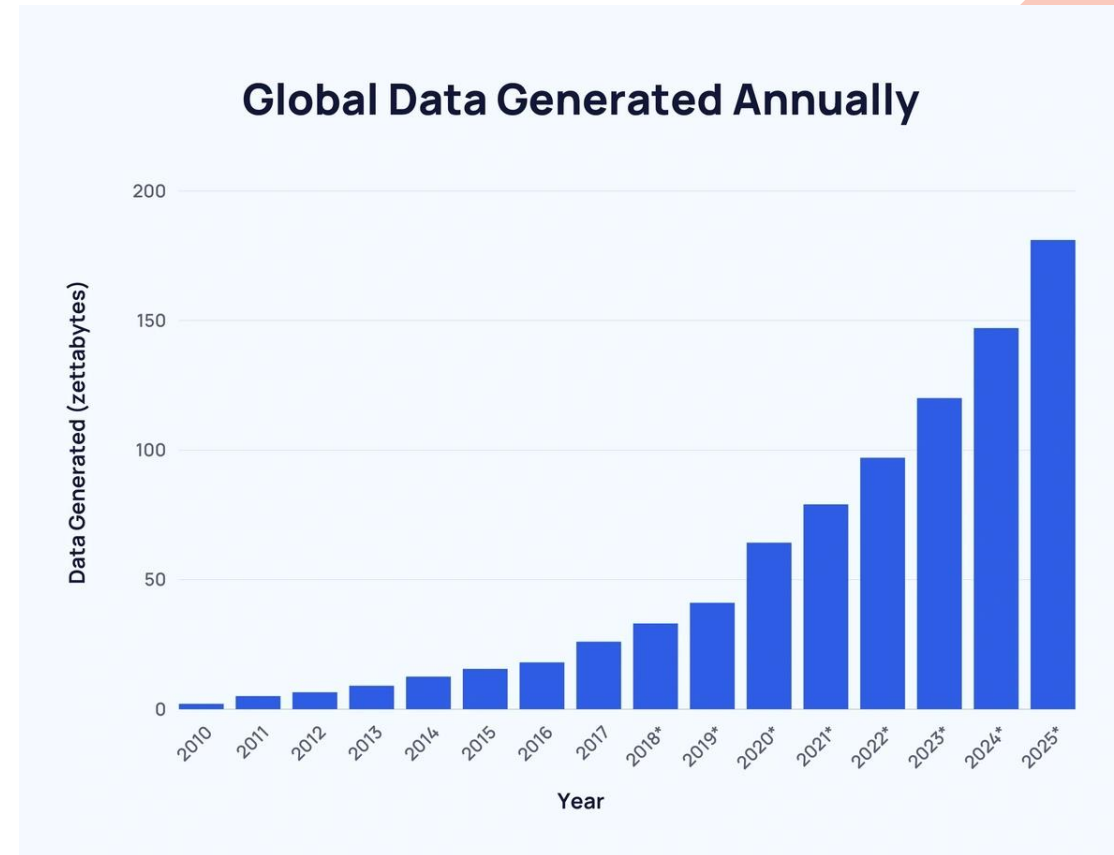
Sidhartha Sinha, MD, AGAF  
Director of Digital Health and Innovation  
Assistant Professor of Medicine  
Division of Gastroenterology & Hepatology  
Stanford University School of Medicine

March 27, 2026



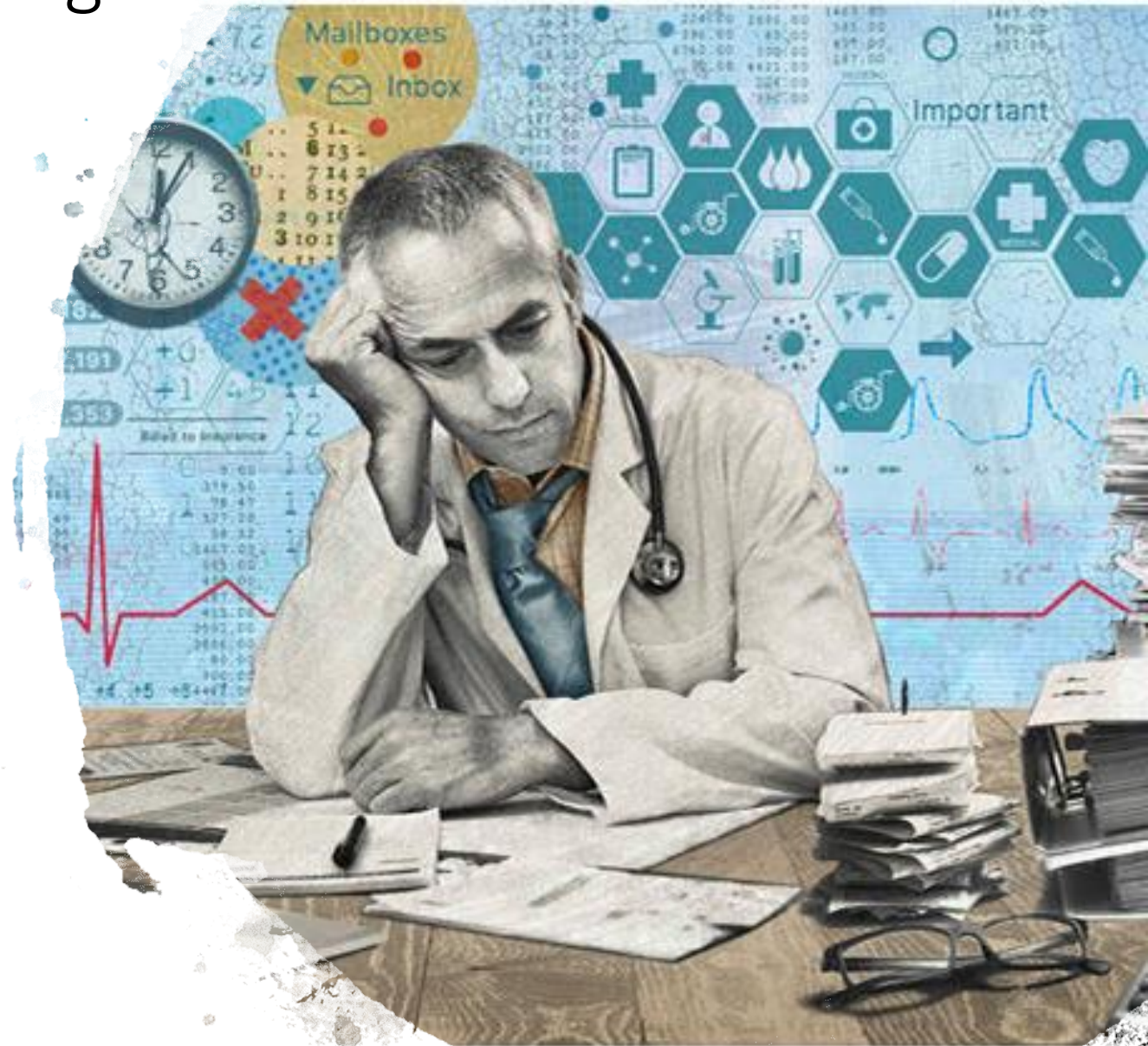
# Framing the Problem

- Ever-growing amounts of medical data
  - EHRs have become unwieldy dumping repositories
  - Impaired ability to review/extract information efficiently
- Physician burnout
  - Linked to EHR screen time
  - Associated with increased errors, attrition, and lower patient satisfaction
- Patients perceive encounters as less personal



# When Data Drowns the Dialogue

- Physicians spend about twice as much time using EMRs as they spend with patients
- Reviewing patient records takes up the majority of this EMR time
- Inefficient extraction and documentation burden reduce time for patient-centered communication



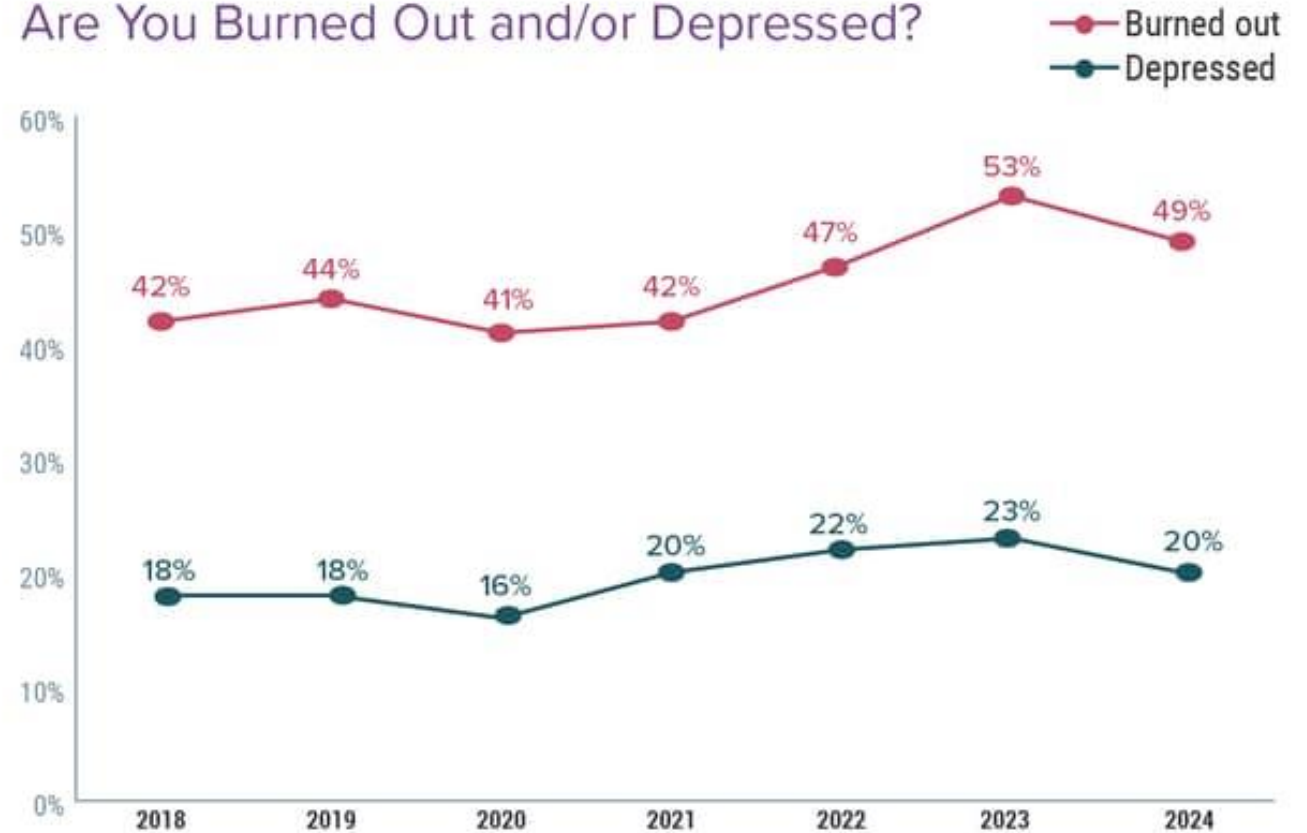
1 Sinsky, C. et al. *Annals of internal medicine* (2016).

2 Wang, J. K., et al. *PloS one* (2019).

# Data overload → ↑ Burnout ↔ ↓ Patient Satisfaction

- Inbox message volumes surged >150% post-COVID
- Clinicians with high volumes had almost 4X the odds of burnout compared to those with the fewest
- 4+ hours per day spent addressing inbox, “pajama time”
- Need: Digital solutions that refocus the encounter on patients...moving from **digitization to transforming** the patient encounter

Are You Burned Out and/or Depressed?



Years shown refer to years report was published. Some respondents said they were both burned out and depressed.

# Automated In basket Replies

## Artificial Intelligence–Generated Draft Replies to Patient Inbox Messages

Patricia Garcia, MD; Stephen P. Ma, MD, PhD; Shreya Shah, MD; Margaret Smith, MBA; Yejin Jeong, BA; Anna Devon-Sand, MPH; Ming Tai-Seale, PhD, MPH; Kevin Takazawa, BBA; Danyelle Clutter, MBA; Kyle Vogt, BA; Carlene Lugtu, MCIM; Matthew Rojo, MS; Steven Lin, MD; Tait Shanafelt, MD; Michael A. Pfeffer, MD; Christopher Sharp, MD

- AI-generated draft replies integrated into Epic MyChart
- 5-week single group study of 162 clinicians across GI, hepatology, and primary care settings
- 12,844 patient messages assessed

The screenshot shows a patient message and a generated draft reply. The patient message is from George Adams to Triage Nurses (supporting Drew Walker, MD) and asks for advice regarding a new granddaughter who is frequently sick and the potential impact on the patient's aortic valve. The AI-generated draft reply is from Dr. Walker, congratulating George on becoming a grandfather and providing advice on hygiene and recommending a follow-up appointment. The interface includes a 'Generated Draft Reply' header, an 'AI Pioneering' badge, a 'Learn More' link, and buttons for 'Start with Draft' and 'Start Blank Reply'. A feedback section at the bottom asks 'Did you find the draft reply helpful?' with thumbs up and thumbs down options.

**New grandpa** (Newest Message First)

George Adams → P Triage Nurses (supporting Drew Walker, MD) Just Now (4:55 PM)

Hi Dr. Walker, I forgot to tell you I'm a new grandpa! She's adorable but also is frequently sick. Do I need to be concerned that being around her germs will hurt my aortic valve?

**Generated Draft Reply** AI Pioneering

Hello George,

Congratulations on becoming a new grandpa! It's wonderful to hear about your adorable granddaughter. Regarding your concern about your aortic valve, it's important to maintain good hygiene practices, such as washing your hands frequently and avoiding close contact if she is sick. However, if you have specific concerns about your aortic valve or your overall health, I would recommend scheduling an appointment to discuss this further.

Please don't hesitate to reach out if you have any more questions or concerns.

Best,  
Dr. Walker

[Learn More](#) [Start with Draft](#) [Start Blank Reply](#)

Did you find the draft reply helpful?


This was helpful This was not helpful

# Automated In basket Replies Shows Early Promise

- Significant task load reduction (Task Load Index)
- Significant improvement in burnout and work exhaustion (Professional Fulfillment Index)
- Utilization varied: 20% of drafts adopted by clinicians (nurses 29% and pharmacists 44%)
- However, no changes in overall reply time were found when comparing pre-pilot vs pilot periods
- Demonstrates feasibility and benefit of LLMs in inbox/in basket workflows
- Questions: Do patients prefer this? Should patients know that messages were generated by AI? Overreliance on AI?...

# Ambient AI Scribes

## Ambient artificial intelligence scribes: physician burnout and perspectives on usability and documentation burden

Shreya J. Shah, MD<sup>1,2,\*</sup>, Anna Devon-Sand, MPH<sup>2</sup>, Stephen P. Ma , MD, PhD<sup>1</sup>, Yejin Jeong, BA<sup>2</sup>, Trevor Crowell, BA<sup>2</sup>, Margaret Smith, MBA<sup>2</sup>, April S. Liang, MD<sup>1</sup>, Clarissa Delahaie, BAS<sup>3</sup>, Caroline Hsia, MEng<sup>3</sup>, Tait Shanafelt, MD<sup>1,4</sup>, Michael A. Pfeffer, MD<sup>1,3</sup>, Christopher Sharp, MD<sup>1</sup>, Steven Lin, MD<sup>1,2</sup>, Patricia Garcia, MD<sup>1</sup>

- Pilot of Nuance's DAX Copilot ambient AI scribe
- 48 physicians across Stanford faculty and community practices
- 3-month pilot: automatic note drafting from clinical conversations



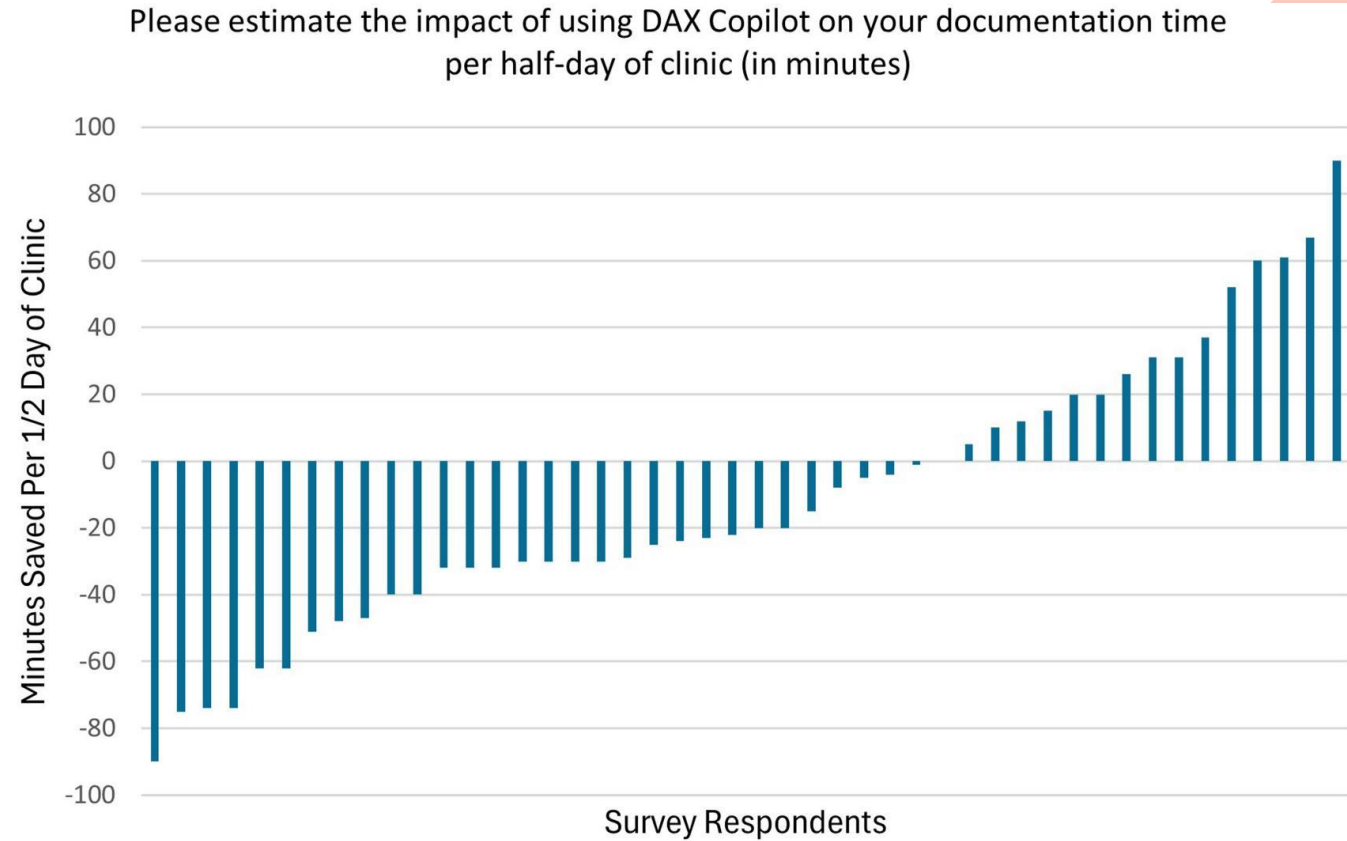
# Ambient AI Scribes Reduce Provider Burden

Physician task load (PTL), Burnout, system usability score (SUS)					
Metric	Pre-pilot mean (SD)	Post-pilot mean (SD)	Cohen's <i>d</i>	<i>T</i> statistic	<i>P</i> -value
PTL	68.98 (17.35)	44.56 (20.62)	1.28	7.35	.000
Burnout <sup>a</sup>	7.47 (2.41)	5.53 (2.10)	0.86	5.12	.000
SUS	58.09 (15.37)	69.01 (16.24)	0.69	-3.20	.003

- Self reported task load scores dropped ,  $p < 0.001$
- Burnout (PFI-WE) improved,  $p < 0.001$
- System Usability Scores, assessing how accessing medical data changed, also improved,  $p = 0.003$

# Ambient AI Scribes Can Save Time...

- Median perceived efficiency gain: 20 minutes per half-day clinic
- Physicians noted improved documentation quality and accuracy
- Common issues with context, nuance, integration, tech failures












# The Evidence Matures: First Rigorous RCT

ORIGINAL ARTICLE



## **NEJM** Ambient AI Scribes in Clinical Practice: A **AI** Randomized Trial

**Authors:** Paul J. Lukac, M.D., M.B.A., M.S. , William Turner, B.S. , Sitaram Vangala, M.S. , Aaron T. Chin, M.D. , Joshua Khalili, M.D. , Ya-Chen Tina Shih, Ph.D. , Catherine Sarkisian, M.D., M.S.H.S. , Eric M. Cheng, M.D., M.S. , and John N. Mafi, M.D., M.P.H.  [Author Info & Affiliations](#)

Published November 26, 2025 | NEJM AI 2025;2(12) | DOI: 10.1056/AIoa2501000 | **VOL. 2 NO. 12**

Copyright © 2025

- 238 physicians, 14 specialties; 1:1:1 randomized to DAX Copilot, Nabla, or control
- **Documentation time:** Nabla -9.5% vs. control (95% CI -17.2% to -1.8%; P=0.02); DAX -1.7% (P=0.66)
- **Burnout (Mini-Z):** DAX +2.83 (95% CI +1.28 to +4.37), Nabla +2.69 (95% CI +1.14 to +4.23); task load and work exhaustion improved for both vs. control
- Burnout benefit confirmed for both; documentation efficiency was found to be tool-specific at UCLA

# Multicenter Evidence: Scale and Consistency

JAMA  
Network | **Open**

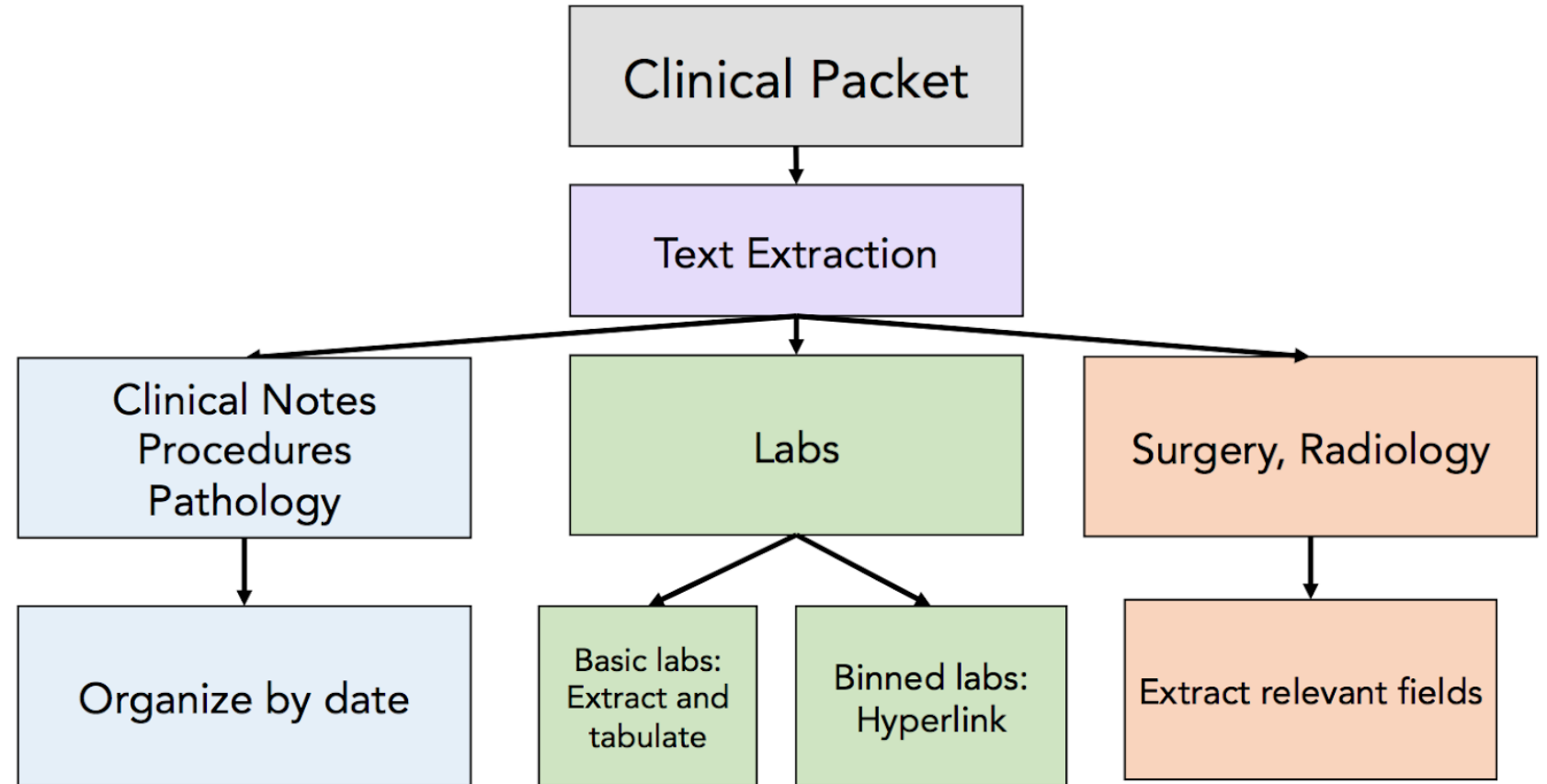
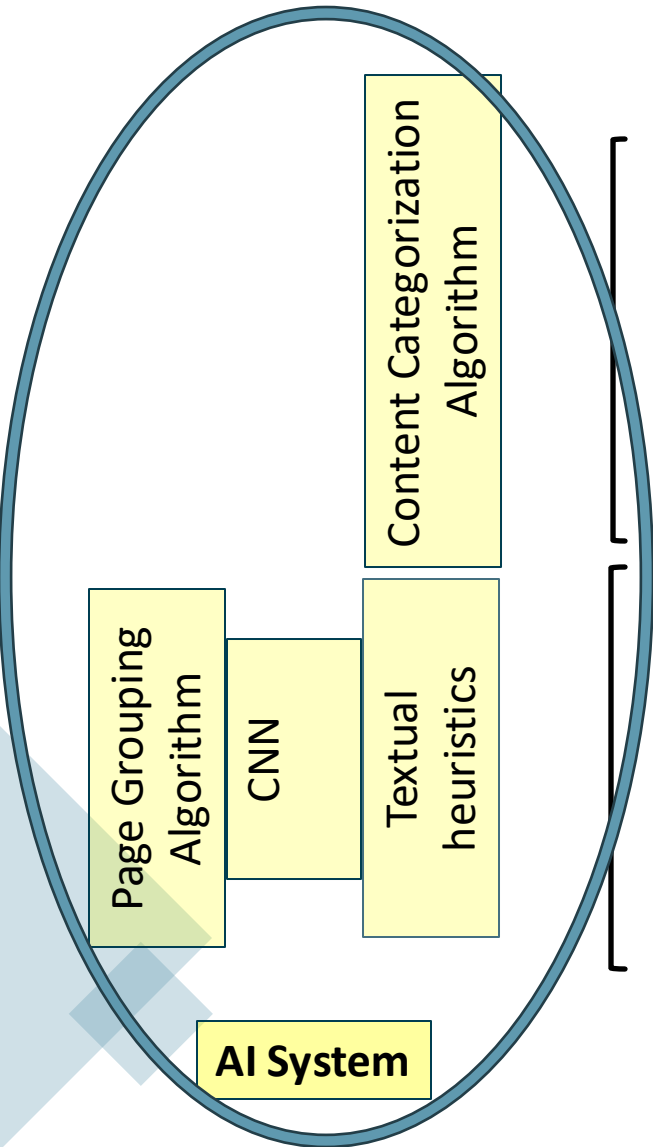
**Original Investigation** | Health Policy

## **Use of Ambient AI Scribes to Reduce Administrative Burden and Professional Burnout**

Kristine D. Olson, MD, MSc<sup>1,2</sup>; Daniella Meeker, PhD<sup>3</sup>; Matt Troup, PA-C<sup>4</sup>; [et al](#)

- 263 physicians/APPs across 6 health systems
- Burnout (Mini-Z): dropped from 51.9% to 38.8% ( $P < 0.001$ ) after 30 days
- 82% reported improved work satisfaction; significant improvements in reduced cognitive task load and after-hours documentation, and increased perception of focused attention on patients during encounters
- Generalizability: consistent signal across 6 academic and community systems — first multicenter evidence; limitation: no control group

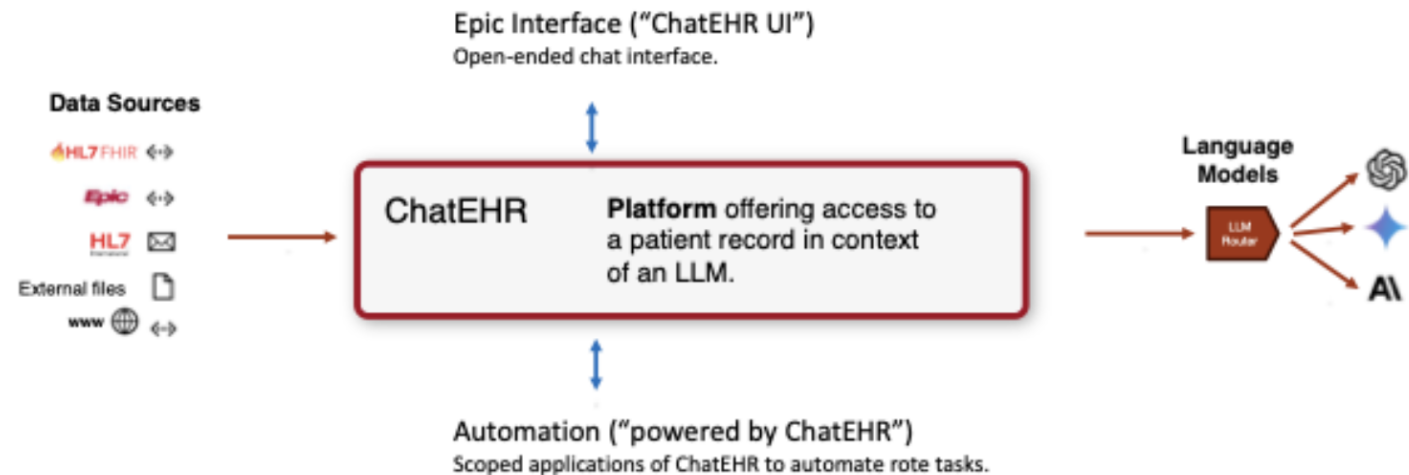
# How to Improve Accessing and Making Decisions with Ever-growing EHR Data?



Chi EA, et al. Development and Validation of an Artificial Intelligence System to Optimize Clinician Review of Patient Records. *JAMA Network Open*. 2021;4(7):e2117391.

# ChatEHR: A Glimpse of the Future EHR Copilot

- ChatEHR (Stanford): Uses LLMs fully integrated into Epic to query patient history, summarize longitudinal data, and answer natural-language questions — e.g., “Show me the last 3 EGD reports”
- Preliminary data: 40% faster chart reviews vs. manual retrieval; supports ChatGPT, Llama, Gemini, Claude
- Represents a shift from EHR-as-filing-cabinet to EHR copilot



# Caution Required in Using AI to Improve the Patient Encounter

## Attestation

**I confirm my understanding of the following key points regarding the use of ChatEHR:**

- **Limitations of ChatEHR:** The outputs generated by ChatEHR are not intended for diagnosing conditions, prescribing treatments, or substituting for professional judgment.
- **Verification of Information:** Users must verify the information provided by ChatEHR for possible errors of accuracy or omission.
- **Use for Documentation:** ChatEHR outputs must be reviewed if used to support documentation. The user remains responsible for the accuracy of their documentation.
- **Feedback and Reporting:** Users are to provide feedback within the ChatEHR user interface and report technical issues using methods described.

**For Clinicians:** I acknowledge that I remain responsible for clinical diagnosis and treatment of my patients at all times. ChatEHR is a tool provided to aid in my practice, and I acknowledge its potential limitations.

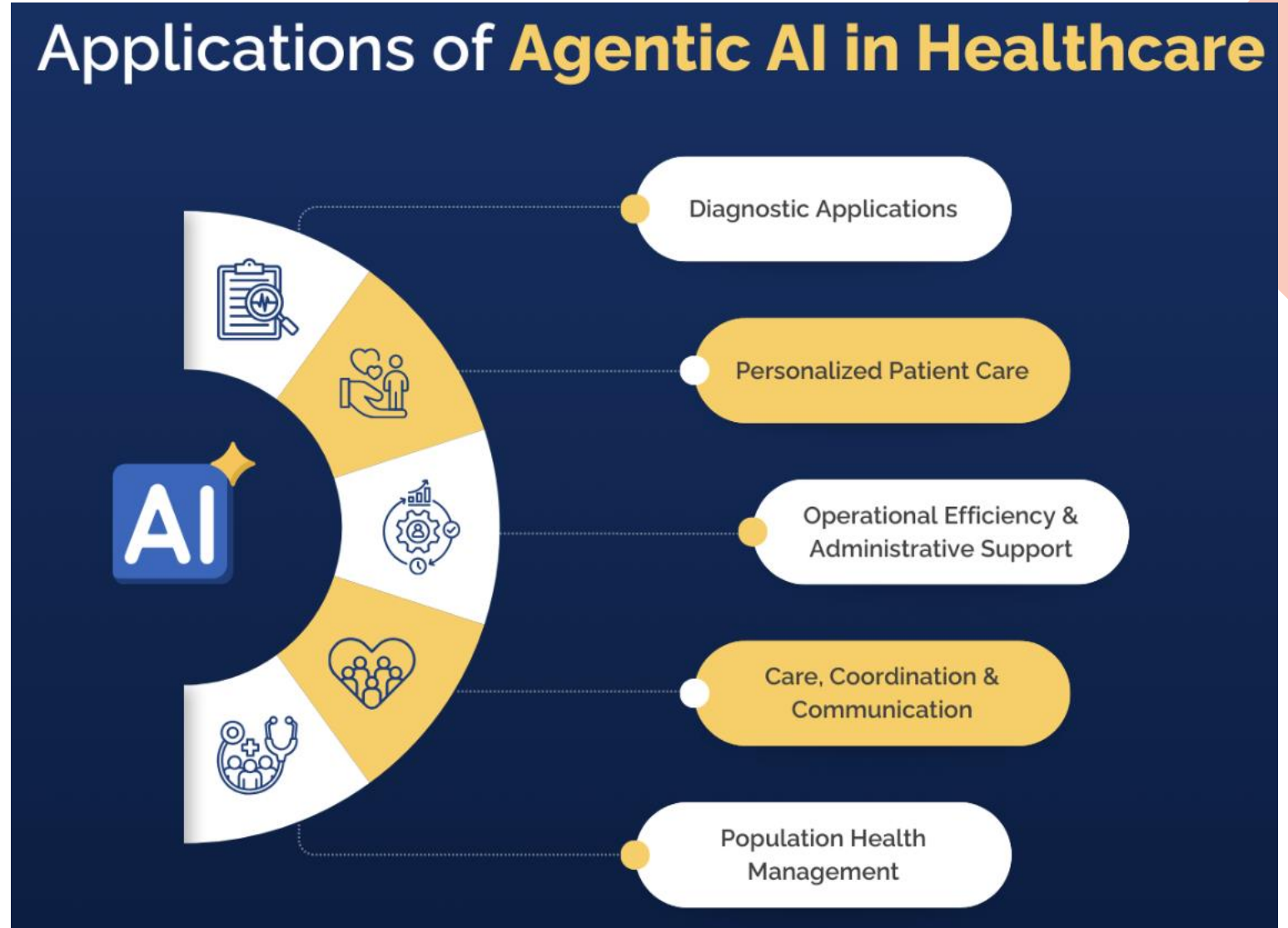
I understand and agree.

**Yes**       **No**

**Submit**

# The Next Frontier: Agentic AI & Workflow Automation

- Vision: Agentic AI handles administrative layers or medical practice: authorizations, refills, routing so physicians can focus on patients
- Physicians spend 14 hrs/week on insurance claims (AMA Survey 2024)
- Agentic AI for PA: Submissions drafted in <90 sec; market grew 10x year-over-year



# The Next Frontier Is Closer than You Think

## Appeal Review Summary

Lumbar Spinal Fusion L4-L5

**Appeal ID:** APL-2024-0284719 | **Member:** UHC-8847291 (62F)

**Procedure:** Lumbar spinal fusion L4-L5 (CPT 22612, 22630)

**Provider:** James Morrison, MD | **Review Due:** January 17, 2025

## Original Denial

### Code: MN-401

**Reason:** Insufficient documentation of failed conservative treatment

**Policy Reference:** Section 4.2.1

### Code: MN-403

**Reason:** Required interventional treatments not documented

**Policy Reference:** Section 4.2.3

**Gap identified:** ESI #3 procedure note was missing from original submission.

### Prompt

Analyze this insurance claim appeal and help me understand why the claim was originally rejected. Cross-reference the patient's medical records and coverage policy to identify whether any supporting justification has been provided. Draft a comprehensive review summary that addresses each denial reason with specific clinical evidence, cites relevant medical necessity criteria, and references any additional applicable information.

### Attachments

Appeal submission  
8kb

PDF

Member medical  
records  
14kb

pdf

# Insurance Claims with LLMs

## Policy Criteria Review

### Failed conservative treatment $\geq$ 6 months

**Evidence:** 9 months documented

**Status:**  Met

### Trial of $\geq$ 2 NSAIDs

**Evidence:** Naproxen, Meloxicam

**Status:**  Met

### Physical therapy $\geq$ 12 sessions

**Evidence:** 18 sessions, documented plateau

**Status:**  Met

### $\geq$ 3 epidural steroid injections

**Evidence:** 3 ESIs with diminishing relief

**Status:**  Met

### MRI within 90 days

**Evidence:** 10/28/24 (51 days prior)

**Status:**  Met

### Functional impairment

#### Prompt

Analyze this insurance claim appeal and help me understand why the claim was originally rejected. Cross-reference the patient's medical records and coverage policy to identify whether any supporting justification has been provided. Draft a comprehensive review summary that addresses each denial reason with specific clinical evidence, cites relevant medical necessity criteria, and references any additional applicable information.

#### Attachments

Appeal submission  
8kb

PDF

Member medical  
records  
14kb

pdf

# Insurance Claims with LLMs

## Recommendation

### ✔ APPROVE APPEAL

1. ESI #3 procedure note now included — addresses primary gap from original denial
2. All policy criteria met with supporting clinical documentation
3. Progressive functional decline despite completed conservative pathway

## Alternative Recommendations

### Peer-to-peer review

**Recommendation:** Not required — documentation complete

### Request additional records

**Recommendation:** Not required — all criteria addressed

### Uphold denial

**Recommendation:** Not supported — gap resolved

## Audit Trail

**Original claim denied (MN-401, MN-403)**

- Date: 12/02/2024

**Appeal received**

### Prompt

Analyze this insurance claim appeal and help me understand why the claim was originally rejected. Cross-reference the patient's medical records and coverage policy to identify whether any supporting justification has been provided. Draft a comprehensive review summary that addresses each denial reason with specific clinical evidence, cites relevant medical necessity criteria, and references any additional applicable information.

### Attachments

Appeal submission  
8kb

PDF

Member medical  
records  
14kb

pdf

# Burnout is Complex

- EHR time alone is a modest predictor of burnout when controlling for autonomy, meaning, and workload
- Other causes:
  - Over 70% of physicians are employed by hospitals or corporate entities
  - With benefits, come reduced control of schedules, staffing and clinical decisions
  - Physicians consistently rank loss of autonomy and inability to provide the care they believe patients deserve as a key issue

HEALTHTECH

## **Doctor Slave Launches as a Bold Contender for the New York Times Best Seller List; Exposing the Hidden Truth Behind Medicine's Corporate Takeover**



By Zeeshan Yousaf

Posted on April 25, 2025



# The Productivity Paradox: Caution Before Celebration

- The paradox: AI may enable more patients/day but if quality suffers, moral injury may rise even as charting time falls (Goodson, Learning Health Systems 2025)
- Volume vs. quality: lower-quality care to more patients can worsen burnout even when documentation time is reduced
- Structural drivers remain: Malpractice stress, loss of autonomy, staffing — ambient scribes alone do not fully address these
- The imperative: Physician co-design with rigorous validation, and post-deployment monitoring to ensure digital empowerment and not a potential new burden

# Governance Frameworks for AI in the Patient Encounter

- Rigorous clinical validation to confirm real-world safety and effectiveness
- Independent review to audit bias to detect and correct demographic or socioeconomic disparities in outputs
- Institutional committees to monitor fairness, accountability, and unintended harms
- Continuous life cycle monitoring post-deployment to detect drift, errors, or safety issues
- Stakeholder Engagement: Include patients, clinicians, ethicists, and technologists in AI deployment decisions

“Believe that a farther shore is reachable from here.”

Seamus Heaney (*The Cure at Troy*)

# Summary: AI/Digital Empowerment and the Future of the Patient Encounter

- From digitization to transformation: Inbox AI and ambient scribes now have RCT-level evidence showing meaningful reductions in burnout
- The evidence base has matured: multicenter studies show burnout significantly dropping with deployment of AI tools, but also reveal tool-specific differences
- The next frontier is agentic AI: workflow automation of prior authorizations, inbox triage, refills, and scheduling — moving beyond documentation to reshape clinical workflows
- BUT the productivity paradox is real: AI tools can increase volume without improving quality; structural drivers of burnout are not solved by scribes and AI alone. **They are a necessary but insufficient piece of a larger systemic response.**
- Governance and physician co-design are non-negotiable: rigorous validation, bias auditing, and continuous post-deployment monitoring are prerequisites, not afterthoughts

Thanks  
sidsinha@stanford.edu

