ETHICAL CONSIDERATIONS IN USING AI

Prof. Dr. Sigrid Sterckx

Bioethics Institute Ghent, Ugent

Centre for Health, Law and Emerging Technologies, University of Oxford

Koninklijke Academie voor Geneeskunde van België

Belgian Society Neonatology – AI within the NICU – 8 Oct 2026



https://www.academiapress.be/nl/geneeskunde-tijden-van-ai

JAMA – 17 AUGUSTUS 2026

AI IN MEDICINE

Will Autonomous AI Exceed AI-Aided Physicians as the Best Medical Care?

Ezekiel J. Emanuel, MD, PhD; Abe Baker-Butler, BA; Neal Khosla, MS; Vinod Khosla, MS, MBA

The prevailing view of artificial intelligence (AI) in medicine is that it will support physician-led care. The American Medical Association regularly calls AI augmented intelligence to focus on AI's assistive role. Similarly, the American College of Physicians argues that AI "should be limited to a supportive role in clinical decision-making" and "should not replace physician decision-making." In *A Giant Leap: How AI Is Transforming Healthcare and What That Means for Our Future*, Wachter[1] argues that the highest tier of care will be AI-aided physicians, whereas AI-only care will be medicine's "economy class."

Supplemental content

We disagree. In cognitive medical functions, AI-alone medical care is likely to be better than physician-only or physician-AI hybrid care. Large language models (LLMs) were only publicly introduced in November 2022, and already generative AI rivals or outperforms licensed physicians at 5 fundamental cognitive medical tasks: (1) eliciting medically relevant information; (2) establishing a differential diagnosis; (3) specifying diagnostic testing; (4) prescribing guideline-concordant treatment; and (5) managing chronic diseases. The gap between physicians' and LLMs' performance will likely widen because AI is rapidly improving, whereas physicians' skills are threatened by AI-induced deskilling.[2,3]

Data from medicine and other fields suggest that when AI-alone performance is consistently superior to human-alone performance, AI alone surpasses human-AI hybrids. Paradoxically, hybrid care in which humans are in (or on) the loop to correct AI errors is likely to worsen rather than improve AI performance. Review of all published articles on AI in medicine since January 1, 2024, shows that medicine is rapidly approaching the transition point at which AI alone will exceed physicians and physician-AI hybrids in providing the best care at 5 fundamental cognitive medical tasks.

EVERY YEAR 12 MILLION DIAGNOSTIC ERRORS RE ADULT PATIENTS IN THE US ALONE (HALF OF WHICH: POTENTIALLY SEVERE DAMAGE)

BMJ Open Access Visit this article BMJ

▸ BMJ Qual Saf. 2014 Apr 17;23(9):727–731. doi: 10.1136/bmjqs-2013-002627

The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations

Hardeep Singh [1], Ashley N D Meyer [1], Eric J Thomas [2]

JAMA – 17 AUGUSTUS 2026

AI IN MEDICINE

Will Autonomous AI Exceed AI-Aided Physicians as the Best Medical Care?

Ezekiel J. Emanuel, MD, PhD; Abe Baker-Butler, BA; Neal Khosla, MS; Vinod Khosla, MS, MBA

"Rather than improving on AI, physicians-in-the-loop consistently degrade AI's performance. …

We need to stop denying that and confront it, and rethink medical practice and education in light of the new reality."

CONFLICTING MESSAGES …



Editorial

https://doi.org/10.1038/s41591-026-04389-4

Show us the evidence for the value of medical AI

nature medicine

Volume 32 | April 2026 | 1163 | 1163

cal information and advice[2]. Yet evidence that AI tools create value for patients, providers or health systems remains scarce.

Nonetheless, in publications, and in product materials, claims about clinical impact are increasingly more common, even though there is no clear agreement on what level of evidence should be required before such claims are considered credible. The result is not only scientific uncertainty but also often premature implementation and adoption. If

CONFLICTING MESSAGES …

The Illusion of Readiness in Health AI

Yu Gu[1*], Jingjing Fu[1], Xiaodong Liu[1], Jeya Maria Jose Valanarasu[1], Noel C. F. Codella[1], Reuben Tan[1], Qianchu Liu[1], Ying Jin[1], Sheng Zhang[1], Jinyu Wang[1], Rui Wang[1], Lei Song[1], Guanghui Qin[1], Naoto Usuyama[1], Cliff Wong[1], Hao Cheng[1], HoHin Lee[1], Praneeth Sanapathi[1], Sarah Hilado[1], Tristan Naumann[1], Javier Alvarez-Valle[1], Jiang Bian[1], Mu Wei[1], Khalil Malik[1], Lidong Zhou[1], Jianfeng Gao[1], Eric Horvitz[1], Matthew P. Lungren[1], Doug Burger[1], Eric Topol[2*], Hoifung Poon[1*], Paul Vozila[1]

[1]Microsoft Research, Health & Life Sciences, Redmond, WA, USA.
[2]Scripps Research Translational Institute, Scripps Research, La Jolla, CA, USA.

*Corresponding author(s). E-mail(s): aiden.gu@microsoft.com; etopol@scripps.edu; hoifung@microsoft.com;

Abstract

Large language models have demonstrated remarkable performance in a wide range of medical benchmarks. Yet underneath the seemingly promising results lie salient growth areas, especially in cutting-edge frontiers such as multimodal reasoning. In this paper, we introduce a series of adversarial stress tests to systematically assess the robustness of flagship models and medical benchmarks. Our study reveals prevalent brittleness in the presence of simple adversarial transformations: leading systems can guess the right answer even with key inputs removed, yet may get confused by the slightest prompt alterations, while fabricating convincing yet flawed reasoning traces. Using clinician-guided rubrics, we demonstrate that popular medical benchmarks vary widely in what they truly measure. Our study reveals significant competency gaps of frontier AI in attaining real-world readiness for health applications. If we want AI to earn trust in healthcare, we must demand more than leaderboard wins and must hold AI systems accountable to ensure robustness, sound reasoning, and alignment with real medical demands.

VERY HARD TO 'ISOLATE' THE **'IMPACT OF AI'** ON **HUMANITY IN MEDICINE** FROM THE IMPACT OF OTHER TRENDS IN MEDICINE (AND SOCIETY) !!

>> What are new factors due to digital automation and AI ? and what is a continuation or strengthening of earlier factors?

>> Perhaps many of the issues we face today are not new questions but new formulations of centuries old questions in medicine ? >> historical (as opposed to futuristic) perspective also matters!

Main hypothesis

(you will be the judge of its plausibility >>

react, refute, corroborate, … as we go along) :

AI transforms what physicians (don't) 'see' …

and hence also what they (don't) 'do'

‘SEEING’ = DIAGNOSIS

DIAGNOSIS = TECHNIQUE + SELECTION CRITERIA

Stanley Reiser (physician & founder ‘Program in Medical History’, Harvard Medical School; Clinical Professor George Washington Univ & Univ of Texas):

“the change from one order of diagnostic procedure to another involves far more than a change of technique: it involves, basically, a transformation of the criteria we use to decide what to direct our attention to from the universe of possible observations that can be made on the person who is sick. It is the alteration of these fundamental norms of selection …”

[**Reiser**, S.J. (1982), ‘Decision Making and the Evolution of Modern Medicine’, in Connelly, Benson & Burke (eds), *Clinical Decisions and Laboratory Use*. University of Minnesota Press, pp. 11-15, 11.]

DIAGNOSTIC METHODS

Since start 19th C: increasing **use of technology** for diagnosis:

- to **reduce physician uncertainty**
- based on the idea that technology (whether a stethoscope or an AI-based Clinical Decision Support System) yields '**objective**' **and standardised information** regarding disease

DIAGNOSTIC METHODS

- 17th and 18th C: diagnosis based on patient testimony (listening), looking at the patient, and (rarely) manual assessment (touching) of the body
- Early 19th C: instruments appear (e.g. stethoscope, other 'scopes' to look inside via orifices)
- Late 19th C: X-rays rendering the human skin 'transparent'
- Measuring & quantification of various bodily phenomena (e.g. blood pressure, lung capacity, temperature)
- Chemical testing of substances removed from the body

--- patient testimony increasingly seen as unnecessary /superfluous/ unreliable ! (ousting of 'subjective' elements)

1816: stethoscoop

The name of this instrument derives from the Greek words for “chest” and “I see”

???? but is *not* about “seeing”, surely ?

>> what is the idea ??

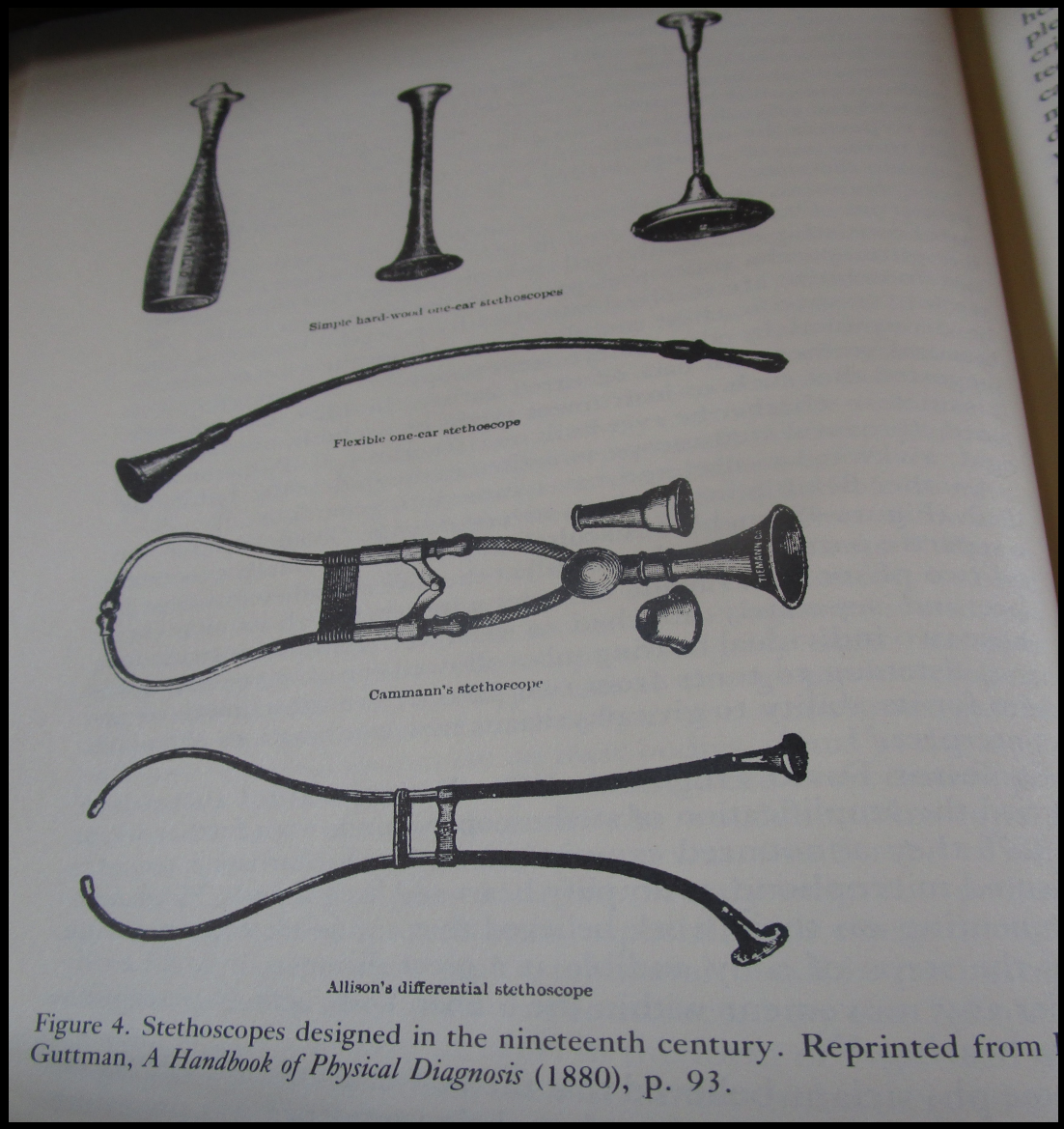


Figure 4. Stethoscopes designed in the nineteenth century. Reprinted from Guttman, *A Handbook of Physical Diagnosis* (1880), p. 93.

WHAT DO YOU SEE?

WHAT STRIKES YOU?



Figure 7. A patient being X-rayed in 1896. Reprinted from Sydney Rowland, "Report on the Application of the New Photography to Medicine and Surgery," *British Medical Journal*, vol. 1 (1896), p. 557

The “technological imperative” in medicine

Co-driven by fear of uncertainty and ambiguity

>> becomes a **problem if machines pre-empt the clinician’s judgment**:

*“I know that the minute I see the x-rays of the patient, before looking at the patient or before working on him, I will **fit** the patient’s story into whatever the x-rays or other images are showing me.”*

-- Howard M. Spiro, Yale School of Medicine

BMJ Overdiagnosis in primary care – Framing the problem and finding solutions

Table 1 Drivers of overdiagnosis

Category	Factor	Example
Broadening disease definitions	Lowering of diagnostic thresholds	Changes that defined CKD at a higher creatinine clearance led to diagnosis of CKD in 25-35% of people over age 65, few of whom will progress to end stage renal disease[15] [16]
	Recognition of risk factors as pre-diseases	Pre-diabetes is highly prevalent (eg. prevalence nearly 36% in China); many patients are treated with drugs.[17] Only about a third will progress to true diabetes over 10 years[9] [18]
Technology	Use of advanced technology for diagnosis	Increasing use of CT, ultrasound, and MRI over time lead to a dramatic rise in the incidence of incidentally detected thyroid cancer, with no concurrent change in mortality[19]
	Use of more sensitive screening tests	Digital mammography is more sensitive than film mammography in some groups but tumors detected have better prognosis, suggesting overdiagnosis[20] [21]
Public health interventions	Widespread screening	Population based breast cancer screening results in 1-10% of cancers; this represents overdiagnosis (in European countries)[22]
Culture of medical care	Value of diagnosis for its own sake	Both patients and physicians feel anxious when problems are not labeled with a diagnosis[23]

DRIVER : TECHNOLOGY : "technological imperative"

There is an **imperative of <u>possibility</u>** : cf. statements such as:

- *"we have to try everything possible"*
- *"more choices [of things to try] are better than few"*

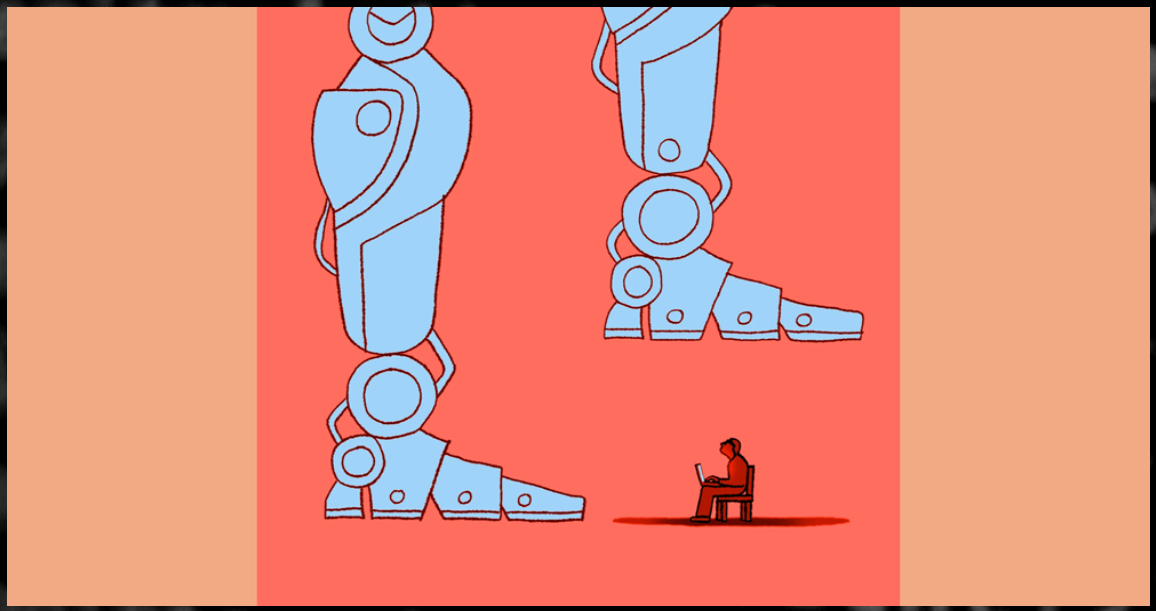
There is an **imperative of <u>action</u>** : cf. statements such as:

- *"do something!"*
- *"we do not want to be accused of not having tried"*

>> **action** is preferable over no action & '**effective**' action is seen as action that involves **technology**

Hofmann 2002 Internat. J. of Technol. Assess. in Healthcare 18(3): 675

WHAT ABOUT UN-INTENDED EFFECTS OF AI ?



Hypotheses:

>> the illusion of certainty and deskilling of physicians will increase ;

>> burnout may decrease, but this comes with a novel major risk

THE ILLUSION OF CERTAINTY

Even though AI hype creates the impression that medical decision-making will become ever more robust, **uncertainty** will always remain a characteristic element of medicine.

However: the implementation of AI systems can make it easier to simply forget or deny this, as this is much more psychologically comfortable:

> *"[T]he digitization of medicine has contributed to shifting the idea of uncertainty, from being a natural and irreducible element of medical practice in the **interpretation** of subtle and sometimes contradictory clues in the existential and complex context of idiosyncratic patients, into the domain of those rational problems that can be modeled to pursue … a **computational** [solution]."*
>
> (Cabitza et al. 2018)

DESKILLING EFFECT

Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task△

Nataliya Kosmyna [1]
MIT Media Lab
Cambridge, MA

Eugene Hauptmann
MIT
Cambridge, MA

Ye Tong Yuan
Wellesley College
Wellesley, MA

Jessica Situ
MIT
Cambridge, MA

Xian-Hao Liao
Mass. College of Art and Design (MassArt)
Boston, MA

Ashly Vivian Beresnitzky
MIT
Cambridge, MA

Iris Braunstein
MIT
Cambridge, MA

Pattie Maes
MIT Media Lab
Cambridge, MA

GHENT UNIVERSITY

DESKILLING EFFECT: MIT STUDY

We assigned participants to three groups: LLM group, Search Engine group, Brain-only group, where each participant used a designated tool (or no tool in the latter) to write an essay. We conducted 3 sessions with the same group assignment for each participant. In the 4th session we asked LLM group participants to use no tools (we refer to them as LLM-to-Brain), and the Brain-only group participants were asked to use LLM (Brain-to-LLM). We recruited a total of 54 participants for Sessions 1, 2, 3, and 18 participants among them completed session 4.

We used electroencephalography (EEG) to record participants' brain activity in order to assess their cognitive engagement and cognitive load, and to gain a deeper understanding of neural activations during the essay writing task. We performed NLP analysis, and we interviewed each participant after each session. We performed scoring with the help from the human teachers and an AI judge (a specially built AI agent).

DESKILLING EFFECT: MIT STUDY

n-grams, ontology of topics within each group. EEG analysis presented robust evidence that LLM, Search Engine and Brain-only groups had significantly different neural connectivity patterns, reflecting divergent cognitive strategies. Brain connectivity systematically scaled down with the amount of external support: the Brain-only group exhibited the strongest, widest-ranging networks, Search Engine group showed intermediate engagement, and LLM assistance elicited the weakest overall coupling. In session 4, LLM-to-Brain participants showed weaker neural connectivity and under-engagement of alpha and beta networks; and the Brain-to-LLM participants demonstrated higher memory recall, and re-engagement of widespread occipito-parietal and prefrontal nodes, likely supporting the visual processing, similar to the one frequently perceived in the Search Engine group. The reported ownership of LLM group's essays in the interviews was low. The Search Engine group had strong ownership, but lesser than the Brain-only group. The LLM group also fell behind in their ability to quote from the essays they wrote just minutes prior.

DESKILLING EFFECT: MIT STUDY

As the educational impact of LLM use only begins to settle with the general population, in this study we demonstrate the pressing matter of a likely decrease in learning skills based on the results of our study. The use of LLM had a measurable impact on participants, and while the benefits were initially apparent, as we demonstrated over the course of 4 months, the LLM group's participants performed worse than their counterparts in the Brain-only group at all levels: neural, linguistic, scoring.

DESKILLING EFFECT: MIT STUDY

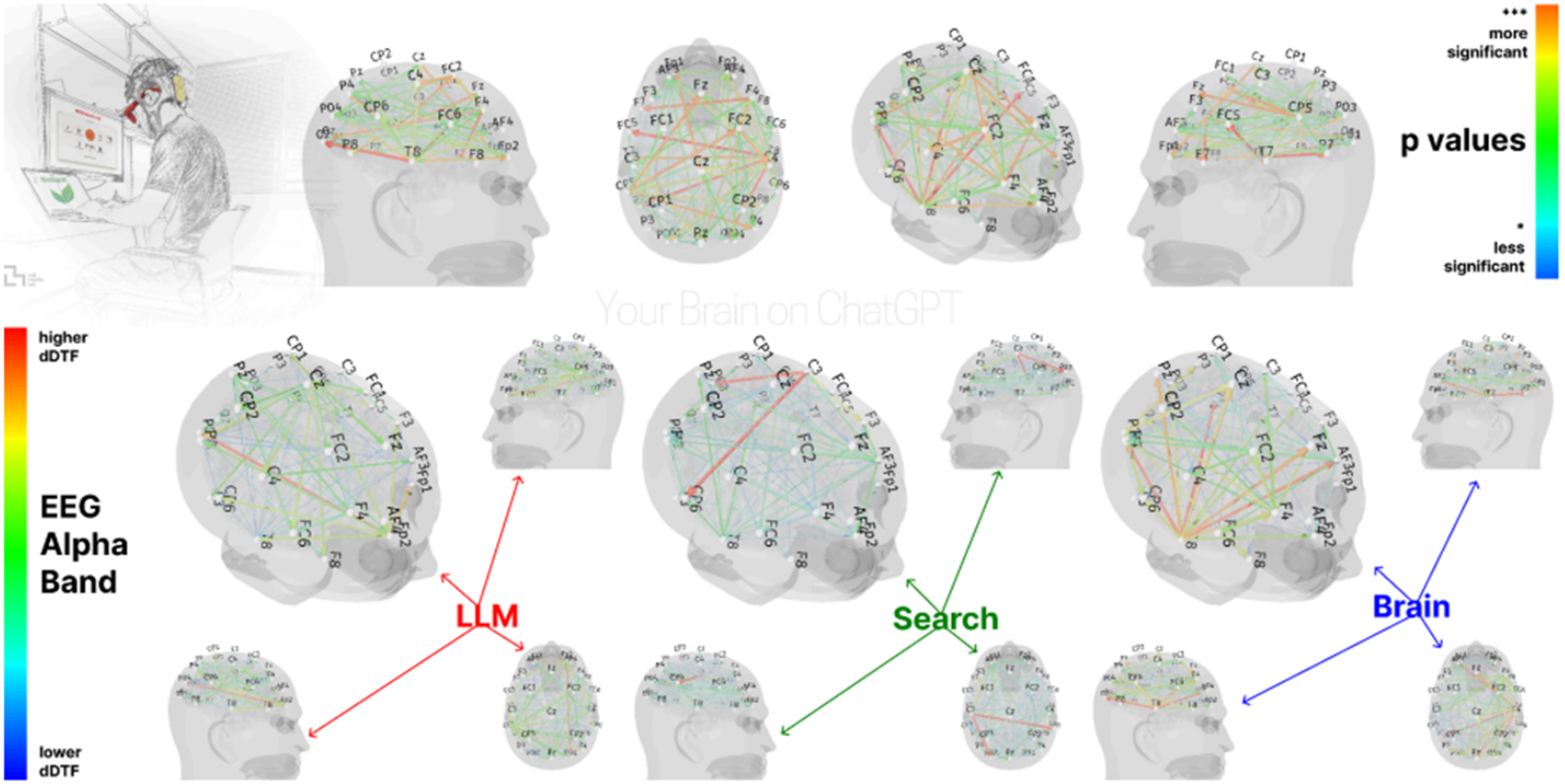


Figure 1. The dynamic Direct Transfer Function (dDTF) EEG analysis of Alpha Band for groups: LLM, Search Engine, Brain-only, including p-values to show significance from moderately significant (*) to highly significant (***).

GHENT UNIVERSITY

DESKILLING

“AI supplies structure [and] vocabulary […] that **inflate the feeling of mastery** during study without strengthening memory traces through retrieval or learner-generated elaboration. […] **Students confuse the AI’s fluency with their own understanding**, a **metacognitive blindness** that prevents them from recognising when assistance is undermining their learning. […] The AI handles the cognitive labor, but the student experiences the outcome as their own achievement. **When the tool is removed, the illusion collapses**. **But by then, the opportunity for genuine learning has passed**. […]

The **point** of a **degree**, on the older view, was never the certificate. It was the slow cultivation of mind that the certificate happened to represent. […] what is valuable about higher education is the knowledge acquired in the doing of it, and […] **the credential is merely the residue of an interior transformation**. But that defence **collapses** when the **learning never happens**.”

Carl Hendrick 2026 – The Death of the University Degree

DESKILLING EFFECT: SPECIFICALLY PHYSICIANS

THE LANCET

This journal Journals Publish Clinical Global health Multimedia Events About

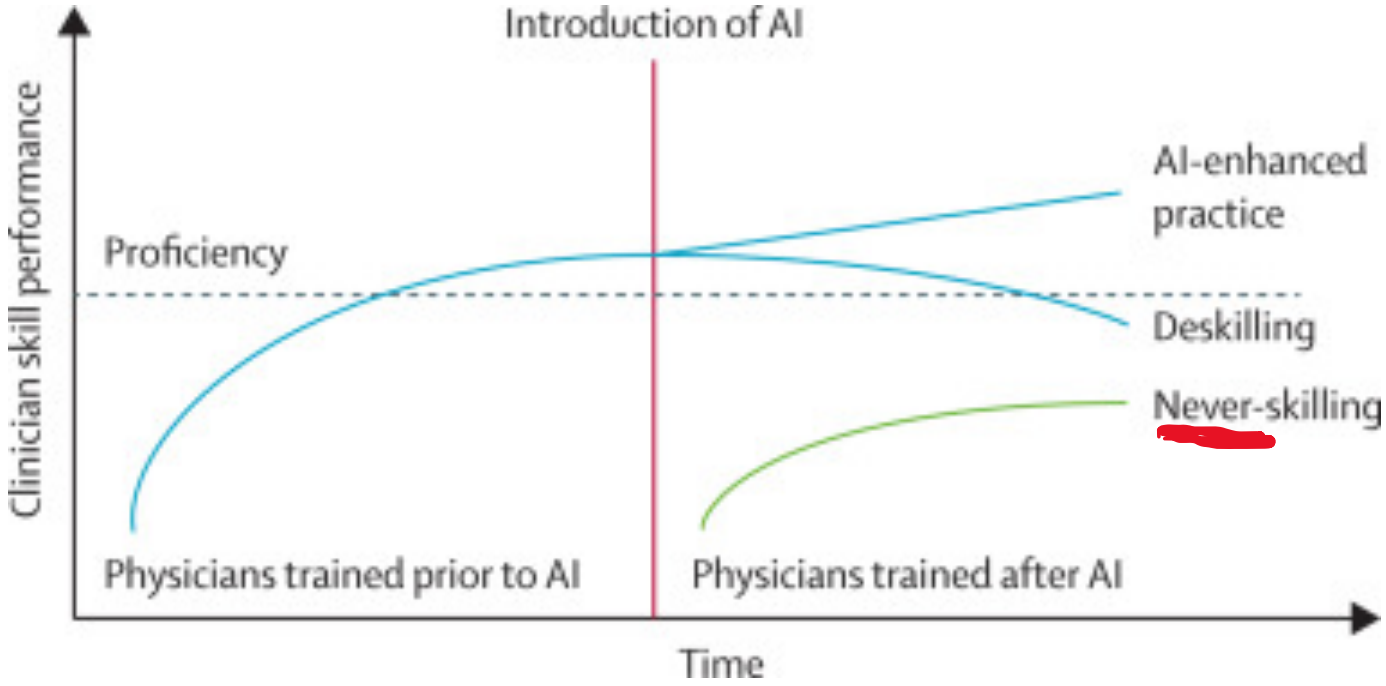
PERSPECTIVES | Digital medicine · Volume 406, Issue 10513, P1719 / October 18, 2025 Download Full Issue

Preserving clinical skills in the age of AI assistance

Tyler M Berzin [a] · Eric J Topol [b]

Affiliations & Notes Article Info

a Center for Advanced Endoscopy, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA 02215, USA
b Scripps Research Translational Institute, La Jolla, CA, USA



GHENT UNIVERSITY

Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: a multicentre, observational study

Krzysztof Budzyń, Marcin Romańczyk, Diana Kitala, Paweł Kołodziej, Marek Bugajski, Hans O Adami, Johannes Blom, Marek Buszkiewicz, Natalie Halvorsen, Cesare Hassan, Tomasz Romańczyk, Øyvind Holme, Krzysztof Jarus, Shona Fielding, Malina Kunar, Maria Pellise, Nastazja Pilonis, Michał Filip Kamiński, Mette Kalager, Michael Bretthauer, Yuichi Mori

Summary

Background It is not known if continuous exposure to artificial intelligence (AI) changes endoscopists' behaviour when conducting colonoscopy. We assessed how endoscopists who regularly used AI performed colonoscopy when

Lancet Gastroenterol Hepatol 2025

Comment

Endoscopist deskilling: an unintended consequence of AI-assisted colonoscopy?

Computer-aided polyp detection (CADe) in colonoscopy represents one of the most extensively evaluated uses of AI in medicine, demonstrating clinical efficacy in multiple randomised controlled trials (RCTs). A recently published meta-analysis of 44 RCTs suggested an absolute increase of 8% in the adenoma detection rate (ADR) with CADe-assisted colonoscopy.[1] However, professional

There are of course limitations to the current study that the authors acknowledge. These include the observational study design, absence of procedural metrics such as withdrawal time, small per-operator sample sizes limiting endoscopist-level analyses, use of only experienced endoscopists, and evaluation of just one AI manufacturer's device. Nevertheless,

Lancet Gastroenterol Hepatol 2025
Published Online August 12, 2025
https://doi.org/10.1016/S2468-1253(25)00164-5
See Online/Articles
https://doi.org/10.1016/S2468-1253(25)00133-5

19 gastroenterologists and surgeons who used an AI polyp detection system for 3 months showed a decline in adenoma detection rates when subsequently performing colonoscopies without AI.
As a group, their post-AI polyp detection performance fell below their own baseline pre-AI performance.

The colonoscopy AI deskilling study is not an indictment of AI. Evidence from over 40 randomised trials for AI polyp detection shows it helps physicians detect more precancerous colon polyps. We should welcome such progress, but remain vigilant. ... The choices we make now about how we design, integrate, and train around AI will determine whether these systems elevate our profession or quietly erode the skills that define it.

DESKILLING BY AUTOMATION

Automation : the exercise by a machine (usually a computer) of a function previously exercised by a human being.

Parasuraman, R., & Riley, V. (1997). Humans and automation: use, misuse, disuse, abuse. *Human factors*, *39*(2), 230–253.

automation changes the work AND the worker

"A work-saving device does not just provide a substitute for a specific part of a job; it changes the character of the entire task + the roles, attitudes and skills of the people involved" (Parasuraman et al., 2000)

>> redistribution of cognitive functions between human (e.g. physician) and machine

>>> QUESTION: how does increasing (digital) automation **reconfigure** the conditions for *agency* and *moral responsibility* of physicians ?

automation changes the work AND the worker

Crucial question : which functions should /may be automated and to what extent ?

(Parasuraman et al. 2000: 286)

>>> Automation as an ethical question



Trends - Katleen De Stobbeleir, Professor Leadership & Partner Vlerick Business School

OR HAS MEDICINE ITSELF BECOME DEHUMANISED ?
ANNALS - SEPTEMBER 2026

IDEAS AND OPINIONS

Annals of Internal Medicine

Why Patients Ask the Chatbot First

Robert B. Shpiner, MD

ten, that patient is responding to the lived reality of modern care: no timely appointment, no established clinician, no confidence that a portal message will be answered, and no wish to generate an urgent care bill for a question that may be minor. The attraction is not that the tool is wise but that it is available.

Trust is the other half of the story. In a 50-state survey, the share of U.S. adults reporting a lot of trust in physicians and hospitals decreased from 71.5% in April 2020 to 40.1% in January 2024 (6). The decline has been steepest across educational and political lines, with a 12-point drop between 2019 and 2024 concentrated among adults without a college degree

Even excellent governance will not solve the problem if patients use chatbots to compensate for failures in access, affordability, continuity, and respect. In that sense, access reform is part of AI policy.

The New York Times https://www.nytimes.com/2025/11/16/well/care-advice.html

Empathetic, Available, Cheap: When A.I. Offers What Doctors Don't

UNINTENDED EFFECTS OF ALIENATION AND BURNOUT

Toll E. (Brown University) The cost of technology. JAMA 2012:



GHENT UNIVERSITY

UNINTENDED EFFECTS OF ALIENATION AND BURNOUT

Toll E. (Brown University) The cost of technology. JAMA 2012:

"No one was more surprised than the physician himself. The drawing was unmistakable. It showed the artist—a 7-year-old girl—on the examining table. Her older sister was seated nearby in a chair, as was her mother, cradling her baby sister. **The doctor sat staring at the computer, his back to the patient—and everyone else**. … you couldn't miss the message.

When he saw the drawing, **the physician wrote** a caption for it: "[A new law] has directed $20 billion to healthcare information technology, largely funding electronic medical record incentives. **I wonder how much this technology will really cost?**"

GHENT UNIVERSITY

UNINTENDED EFFECTS OF ALIENATION AND BURNOUT

"[The **Electronic Health Record**] has become a key instrument for **measuring the duration** and **standardizing the content** of patient-doctor interactions in pursuit of "the one best way" [Taylorism]. … **[Patient-physician] encounters have been restructured** around the demands of the EHR. … **Open-ended interviews**, vital for obtaining accurate clinical information and understanding patients' mindsets, have become **almost impossible**."

Hartzband, Pamela & Groopman, Jerome (2016), "Medical Taylorism", ***New England Journal of Medicine*** 374(2), pp. 106-108, p. 107

AMBIENT AI AS THE SOLUTION ?

- Ambient AI tools/scribes: *streamline documentation by listening in the background and drafting notes during a clinical encounter*
- June 2025: already in use in two thirds of US hospitals which use an Epic EHR

(***Am J Manag Care. 2026;32(1):e25-e30. https://doi.org/10.37765/ajmc.2026.89876***)

- Explicitly marketed as a remedy for burnout in physicians

GHENT UNIVERSITY

AMBIENT AI AS THE SOLUTION ?

Ambience

Product Specialties EHR Integrations Resources Careers Book a demo ›

Specialty-specific AI scribe with compliant CDI

Reduce clinician burnout, improve system efficiency, and enable high quality care.

Book a demo ›



$4.50

Average increase in revenue per E/M visit

AMBIENT AI AS THE SOLUTION ?

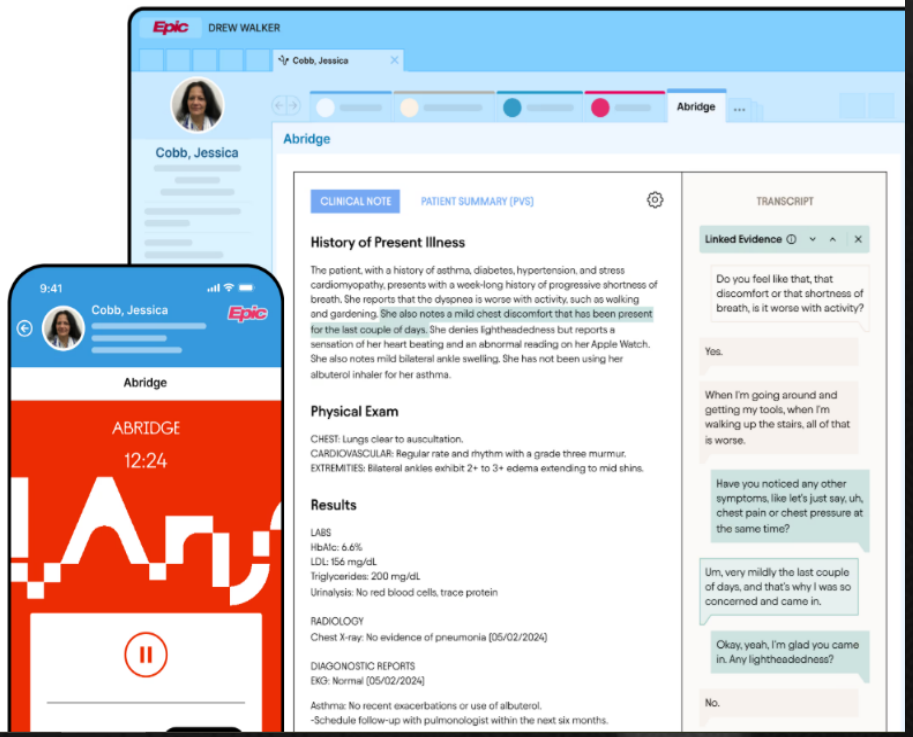
ABRIDGE

PRODUCT AI RESOLUTIONS ABOUT LOGIN CONTACT US

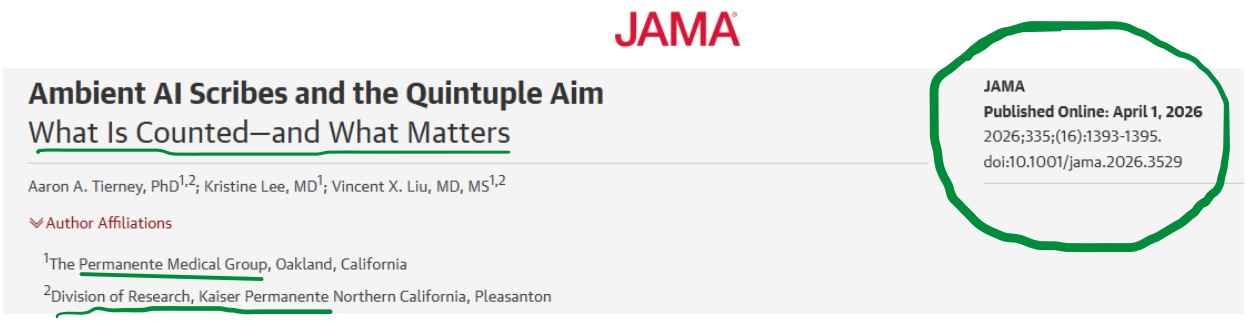
ENTERPRISE-GRADE AI

GENERATIVE AI PLATFORM FOR CLINICAL CONVERSATIONS

Clinically useful and billable AI notes at the point of care, integrated directly into EHR workflows. Powered by the Abridge Contextual Reasoning Engine.



EVIDENCE REGARDING AMBIENT AI TOOLS?



JAMA

Ambient AI Scribes and the Quintuple Aim

What Is Counted—and What Matters

Aaron A. Tierney, PhD[1,2]; Kristine Lee, MD[1]; Vincent X. Liu, MD, MS[1,2]

Author Affiliations

[1]The Permanente Medical Group, Oakland, California

[2]Division of Research, Kaiser Permanente Northern California, Pleasanton

JAMA
Published Online: April 1, 2026
2026;335;(16):1393-1395.
doi:10.1001/jama.2026.3529



JAMA | Original Investigation | AI IN MEDICINE

Changes in Clinician Time Expenditure and Visit Quantity With Adoption of Artificial Intelligence–Powered Scribes

A Multisite Study

- "AI scribe adoption was associated with 13.4 … fewer minutes of EHR time, 16.0 … fewer minutes of documentation time [per 8h planned patient care]"
- "the question is no longer whether ambient AI can save documentation time for clinicians but whether that time is reinvested in ways that measurably improve outcomes and equity for patients"
- "If health systems rely on increasing visit volume to justify the cost of ambient AI adoption, they risk squandering the benefits of time savings if that time is simply converted into more visits per clinician, rather than investments in higher-quality care."
- What about effects on : experience of patients; health outcomes; equity ??

GHENT UNIVERSITY

JAMA HEALTH FORUM 9 JAN 2026

Viewpoint | AI in Health Policy

Unintended Consequences of Using Ambient Artificial Intelligence Scribes for Billing

Paige Nong, PhD; Hannah T. Neprash, PhD

Among artificial intelligence (AI) applications, ambient scribes have been the focus of particularly high enthusiasm across the US health care system.[1] These tools, driven by generative AI, offer a powerful solution to a persistent challenge related to the use of electronic health records: clinician documentation burden. Itself an unforeseen consequence of enthusiastic information technology deployment, documentation burden has chronically reduced clinician satisfaction and fueled burnout.[2] The hope for ambient scribe technologies, which use generative AI to automatically transcribe and summarize clinical encounters to produce draft clinical documentation, is to ease documentation burden. Pilots and single-organization analyses find ambient scribes are associated with some reduced documentation time per visit and after-hours time, for the clinicians who have opted into use.[2-4]

Ambient scribes can be expensive, with monthly costs ranging from $300 to $500 per clinician. Health care delivery organizations may be more eager to adopt the technology if they can demonstrate financial return on investment (ROI), beyond alleviating clinician documentation burden, through automating more intensive billing. We argue that the use of ambient scribes for billing may have multiple negative unintended consequences, including higher health care spending, lower clinician satisfaction in the long term, and eroded patient trust.

+ Editorial

Author affiliations and article information are listed at the end of this article.

CLINICIANS AS QUANTIFIED WORKERS (“veredelde magazijniers”) ?

“medical AI tools … also create a glide path for turning clinicians into “quantified workers” — workers whose daily tasks are monitored and controlled by AI technologies, which denies them autonomy and the benefit of discretion based on their expertise”

“By analyzing speech patterns, sentiment, and content, health care systems could use AI scribes to assess how often clinicians’ recommendations deviate from institutional guidelines. In addition, these systems could detect “efficiency outliers” — clinicians who spend more time conversing with patients than employers consider ideal, at the expense of conducting new-patient visits or more total visits.”

GHENT UNIVERSITY

Medical AI and Clinician Surveillance — The Risk of Becoming Quantified Workers

Author: I. Glenn Cohen, J.D. , Ifeoma Ajunwa, J.D., L.L.M., Ph.D., and Ravi B. Parikh, M.D., M.P.P. Author Info & Affiliations

Published June 14, 2025 | N Engl J Med 2025;392:2289-2291 | DOI: 10.1056/NEJMp2502448 | VOL. 392 NO. 23

Copyright © 2025

SIGNATURE READY NOTES = PARADISE ?

Compare Augmedix products

	Go Self-service	Assist Hybrid	Live Full-service
Daily EHR time savings	1 hour	1.5 – 2 hours	2 – 3 hours
Personalization	Some	More	Most
Produces AI-drafted notes within seconds	✓	✓	✓
Web and mobile apps	✓	✓	✓
HITRUST certified & HIPAA compliant	✓	✓	✓
Bi-directional EHR integrations (Epic, Cerner, Meditech, athenahealth, and more)	✓	✓	✓
Support for 32 languages	✓	✓	✓
CareCues™ - Real-time care gap and billing reminders	✓	✓	✓
Clinical data summary for easy reference during visits	✓	✓	✓
Suggested ICD-10 codes	✓	✓	✓
CPT, E&M, LOS, HCC coding		✓	✓
Produces signature-ready notes	–	Within 1 hour	Moments after every visit

"Total game changer. I'm taking lunch breaks for the first time in years, and I was even able to attend my kid's school play for the first time."

Dr. Robert Grigg, DMSc, PA-C
Paradise Medical Group



SOME FUNDAMENTAL QUESTIONS TO ANIMATE LUNCH

What makes a physician a physician ?

What makes a good physician a good physician ?

Why should the profession of medicine continue to deserve awe (beyond respect) ?

How should the use of AI impact physicians' status and remuneration ?

Do you think the deployment of AI will increase or decrease your professional autonomy ? (cf. ambient AI and physicians becoming 'quantified workers')

THANKS FOR YOUR ATTENTION

sigrid.sterckx@ugent.be