

AI for Clinician Experience and Beyond

Jacqueline G. You, MD, MBI

Digital Clinical Lead for Clinician Experience, Digital, Mass General
Brigham

Associate Physician, Division of General Internal Medicine and
Primary Care, Department of Medicine
Brigham and Women's Hospital

Instructor

Harvard Medical School

Jacqueline G. You, MD, MBI



Boston University Chobanian & Avedisian School of Medicine

Internal Medicine Residency @ Boston Medical Center

Clinical Informatics Fellowship @ MGB

Masters of Biomedical Informatics @ HMS

Instructor @ HMS

Associate Physician, Division of General Internal Medicine and Primary Care, Department of Medicine

- Clinical focus: Primary care
- Operational focus: Clinical informatics

DISCLOSURES

I have no relevant financial relationships with ineligible companies.



OBJECTIVES

- Explain basic terminology for AI in healthcare.
- Describe major use cases of generative AI for clinician experience.



Caveats



AI sampler, by no means comprehensive



Ongoing, rapidly evolving space

Outline

Introduction to AI in Healthcare

AI for Clinician Experience

Other Specialty Applications



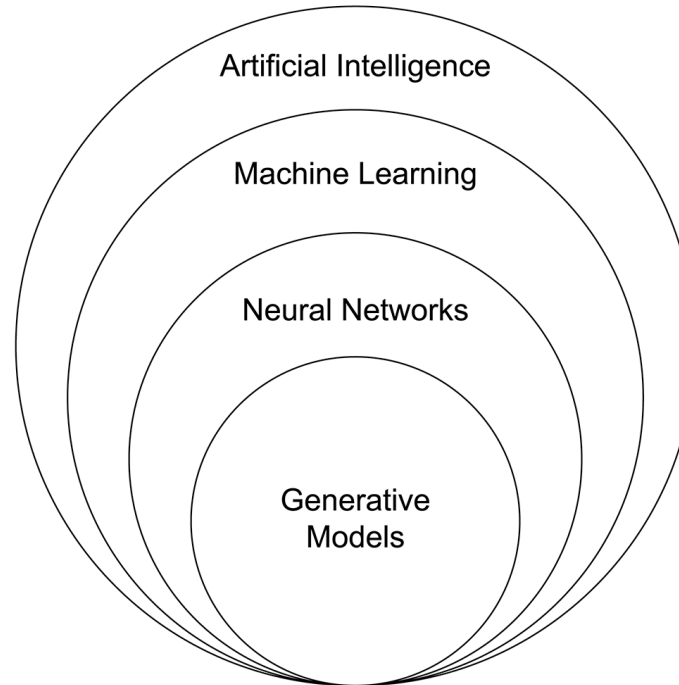
Introduction to *AI* in Healthcare



AI Overview



Impact in healthcare
– broad across
patient care,
research, education.
Estimated ~5% of US
companies using AI
for healthcare based
on 2023-2025 US
Census Bureau data.



Source:
https://commons.wikimedia.org/wiki/File:AI_relation_to_Generative_Models_subset_venn_diagram.png

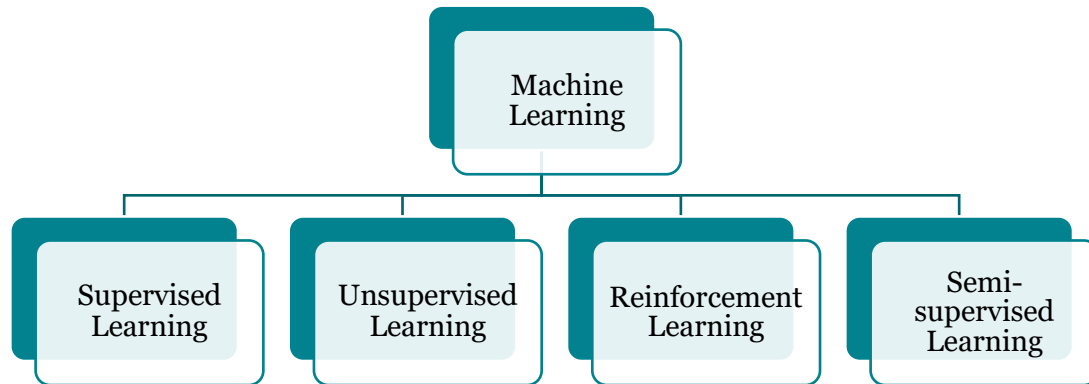
Opportunity for clinicians:

- Challenging time in healthcare with high levels of burnout
- Healthcare systems using AI as a tool in toolkit to address challenges in healthcare
- As clinicians - not the time to take a backseat on this technology
- Ask questions, get involved
- Learn more about gen AI:
 - Introduction to AI post 2022, Isaac Kohane (HMS, Department of Biomedical Informatics):
<https://www.zaklab.org/blog/resources-for-introduction-to-ai-post-2022/>
 - Generative AI for Healthcare, Yao and Videk (Stanford):
https://youtu.be/eLAq8yzvu8Q?si=i_yFoF-C04tZ-4u



Machine Learning

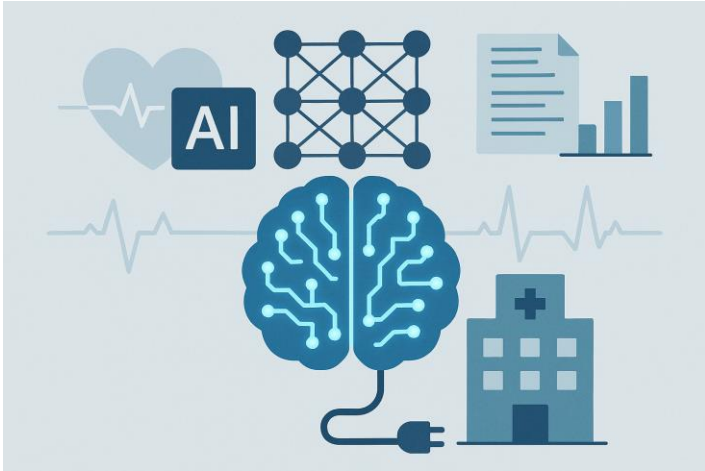
Paradigms



- Supervised
 - Labeled dataset
 - E.g. set of cat vs dog photos, training a model to distinguish between the two
- Unsupervised
 - Unlabeled dataset to make inferences
 - E.g. unlabeled food photos -> might cluster 'sweet' vs 'savory' foods, might cluster based on color
- Reinforcement
 - 'Reward' or 'penalty' to try to nudge system to a metric
 - E.g. training a robot to not hit objects in a room.
- Semi-supervised
 - Hybrid of supervised and unsupervised – you get some labeled data and a larger pool of unlabeled data
 - Use unsupervised learning to cluster data, then supervised to label after



Deep Learning and Generative AI



Deep Learning: subset of machine learning that uses 'layers' of nodes to learn from data

Large Language Models: statistical language model, often with deep learning architecture.

Generative AI: deep learning models with input of large amounts of data, outputs text/images based on statistical likelihood (think very sophisticated auto-complete)

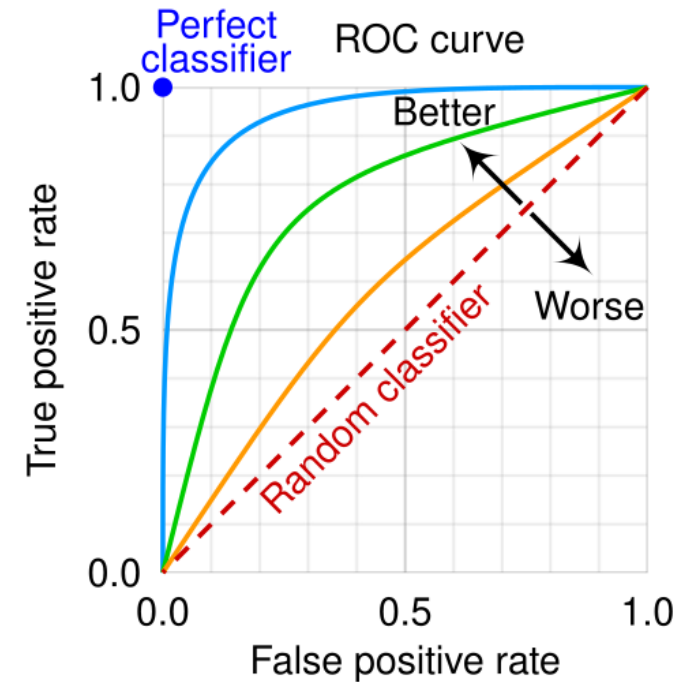
Machine Learning Model Performance

Common metrics:

- Accuracy (correct predictions / total predictions)
- Recall aka sensitivity (true positives / true positives + false negatives)
- Precision (true positives / positive predictions)
- F1 score – blend of precision and recall

For classification models

- ROC curve – receiver operating characteristic curve – plots true positive rate to true negative rate, area under the curve



https://en.wikipedia.org/wiki/File:Roc_curve.svg

Why is Generative AI a Different Beast in Healthcare?

High stakes – patient safety, data, privacy, ethics, medicolegal

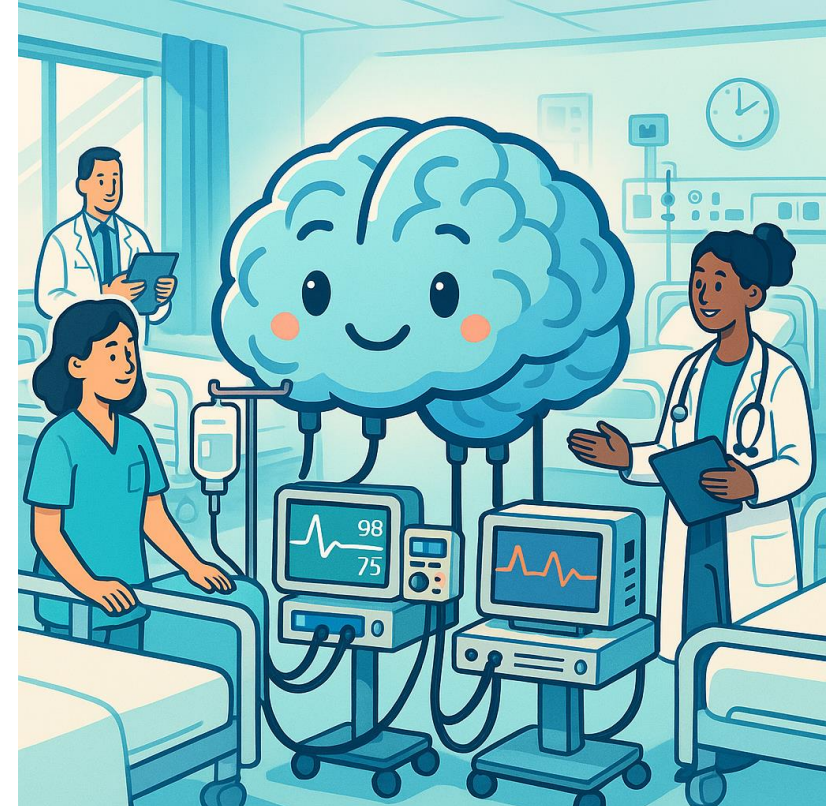
Complexity, data messiness

Less deterministic

Data volume/computing power

Shift to rapid tech evolution

Ongoing multidisciplinary evaluation



?Q : What is the difference between supervised and unsupervised learning?

- A. Supervised learning involves clustering, while unsupervised learning uses a reward-penalty system.
- B. Supervised learning uses a labeled dataset, while unsupervised learning uses unlabeled data.
- C. Supervised learning only learns from large numbers of images, while unsupervised learning uses multiple types of data inputs including text and images.
- D. Supervised learning uses HIPAA-compliant data, whereas unsupervised learning uses publicly available data.



?Q : What is the difference between supervised and unsupervised learning?

A. Supervised learning involves clustering, while unsupervised learning uses a reward-penalty system. → *K-means clustering is an unsupervised machine learning algorithm; reinforcement learning uses a reward-penalty system.*

B. Supervised learning uses a labeled dataset, while unsupervised learning uses unlabeled data.

C. Supervised learning only learns from large numbers of images, while unsupervised learning uses multiple types of data inputs including text and images. → Multimodal models use multiple data types.

D. Supervised learning uses HIPAA-compliant data, whereas unsupervised learning uses publicly available data. → *Supervised versus unsupervised learning is unrelated to the privacy/security of the data.*



AI for Clinician Experience

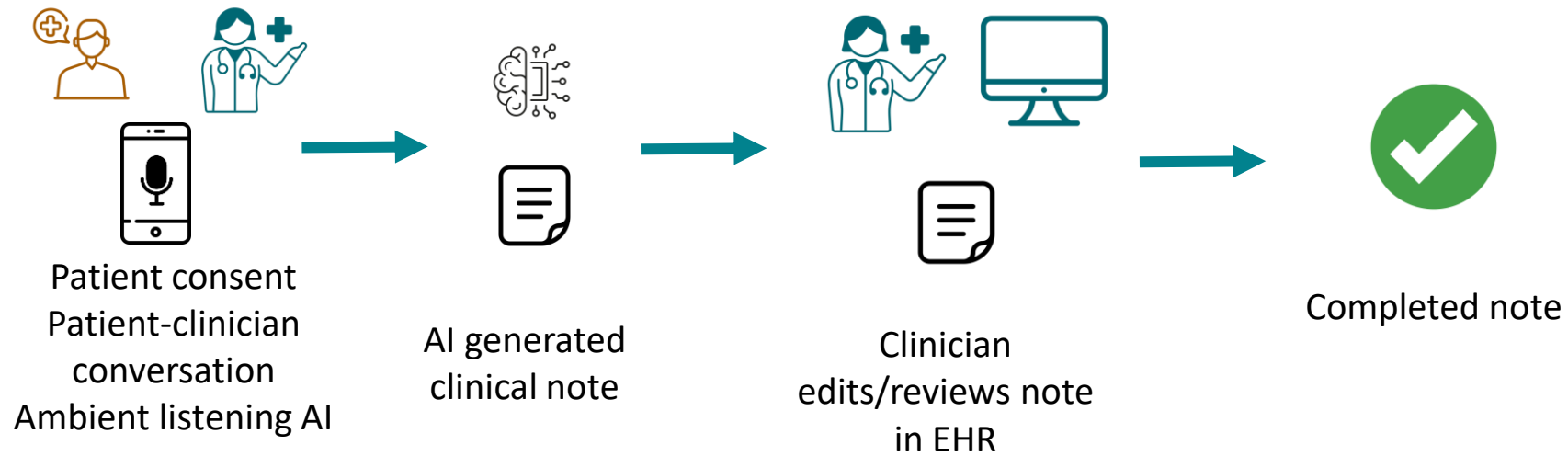


AI for Clinician Experience – Core Internal Medicine Use Cases

- Ambient Documentation
- Chart Summarization
 - Includes various specific use case scenarios such as pre-charting, results summarization
- AI Assistants
- Medical Point-of-Care Information Tools



Ambient Documentation (AI Scribing)



Reduction in documentation after hours, cognitive load, burnout

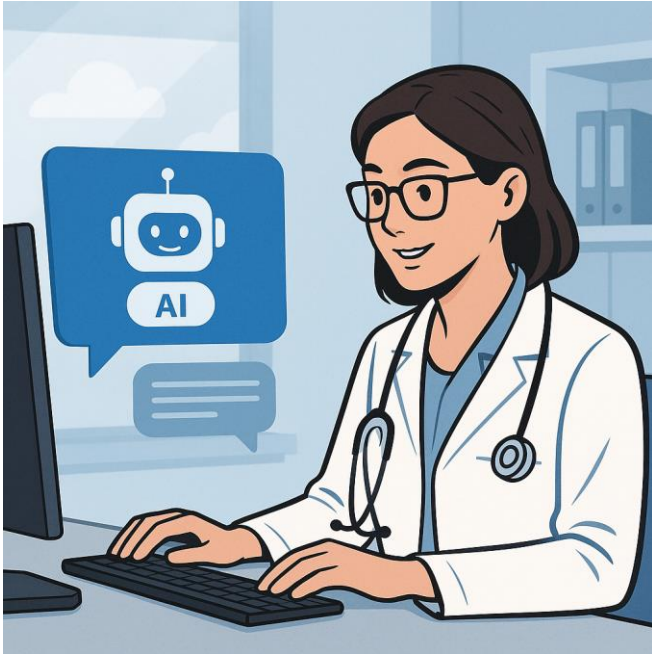
Study	Setting	# of Clinicians	Duration	Improvements
Albrecht et al <i>JAMIA Open</i> 2025	University of Kansas Abridge	191	3 months (avg use)	- Ease of documentation - Job satisfaction/burnout risk
Guo et al <i>JAMIA</i> 2025	UC Irvine Nuance DAX Copilot	167 (quantitative) 65 (qualitative)	3 months	- Time/note, note time/day, note time/appt - Cognitive demand, documentation effort
Liu et al <i>NEJM AI</i> 2024	Atrium Health Nuance DAX Copilot	112 Primary Care	6 months	- Daily documentation time with high usage only - No change in RVUs
Olson et al <i>JAMA Network Open</i> , 2025	6 academic and community health systems (Yale, MemorialCare, Christus Health, Sutter Health, University of Chicago) Abridge	263	30 days	- Burnout - Self-reported cognitive load - Documentation after hours - Focused attention on patients
Ma et al <i>JAMIA</i> 2024 Shah et al <i>JAMIA</i> 2024	Stanford Nuance DAX Copilot	45 (quantitative) 38 (qualitative)	3 months	- Burnout - Taskload - Usability scores - Time per note (0.57 minutes), daily documentation time (6.89 min) - After hours EHR time (5.17 min), Total EHR time (19.95 min)
You et al <i>JAMA Network Open</i> , 2025	2 academic and community health systems (MGB, Emory) Abridge and Nuance DAX Copilot	327	3 months	- Burnout (MGB: sustained at 1 year) - Well-being

More recent publications also show (modest) time savings and financial productivity

Study	Setting	# of Clinicians	Duration	Improvements
Holmgren et al <i>JAMA Network Open</i> 2025	UCSF 2 vendors, unspecified	1565 attending physicians (698 adopters vs 867 non-adopters)	1 year 8 months	• Higher RVUs in adopters without evidence of increased denials • More encounters per week
Rotenstein et al <i>JAMA</i> 2026	5 AMC systems (UCSF, MGB, Emory, Yale, NYU) Multiple vendors	8581 clinicians (1809 adopters)	2 years 2 months, June 2023 – August 2025	• Modest decrease in total EMR time/documentation time • Modest increase in weekly visit volume



AI Chart Summarization



Summary of prior visits for outpatient chart review

Draft discharge summary

Chat format to query questions from the EHR

AI Assistants

General Purpose:

- Large language model vendors (e.g. OpenAI's ChatGPT, Google's Gemini, Meta's Meta AI, Claude AI)

Specific Purpose:

- Work productivity AI tools (e.g. Microsoft Copilot)
- Research AI tools (e.g. ResearchRabbit)
- Medical point-of-care information tools (e.g. OpenEvidence, UpToDate Expert AI, Dyna AI, Elsevier ClinicalKey AI, Doximity Ask)

**** Only use your healthcare system's approved AI tools for business/patient care, even if an external system is labeled as HIPAA-compliant.**



AI in Healthcare: Other Real World Applications



Patient Education



Billing / Coding and
Prior Authorization



Medical Education
(Teaching Diagnostic
Reasoning, Practicing
Patient Education)

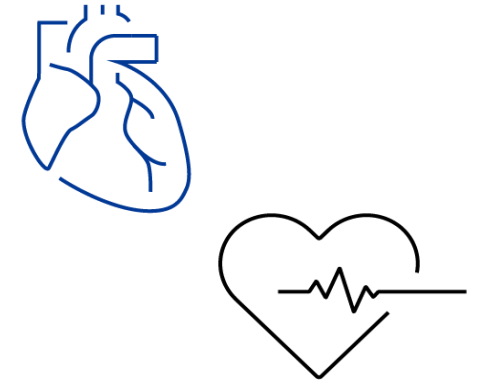


Patient Portal
Messages (AI-
generated Draft
Replies, Draft Results)



Emergency Room
Triage (AI-Based
Triage Support)

IM Subspecialty: Cardiology



- Review of 277 FDA-cleared AI/machine learning medical devices for cardiology:
 - Imaging – 65%, diagnostics – 64%, predictive 6.4%, prescriptive 0.7%
 - Limited support with few high quality trials
 - Example applications
 - ECG tools for atrial fibrillation detection
 - Remote monitoring of heart failure for prediction of exacerbations
 - Quantifying calcification of coronary artery plaques
 - AI-guidance of echocardiography image acquisition

IM Subspecialty: Gastroenterology



- ACG task force Delphi consensus on AI for GI practice (2026):
 - Endoscopy
 - Adenoma detection
 - Diagnostic promise in risk stratification, diagnostic accuracy, treatment selection
 - Clinician administrative support:
 - Ambient documentation, NLP-coding, workflow, authorization support
 - Training: Balance structured AI curriculum with procedural competencies



?Q : Ambient documentation, or AI scribing, has not been shown to:

- A. Improve burnout
- B. Decrease taskload
- C. Improve time in notes
- D. Decrease RVU output



?Q : Ambient documentation, or AI scribing, has not been shown to:

- A. Improve burnout
- B. Decrease taskload
- C. Improve time in notes
- D. Decrease RVU output** → net neutral to positive impact on RVUs/financial efficiency



Other Specialty Applications





- In US, majority of FDA approved AI-enabled medical devices in radiology (2025 review: 77%)
- Between 1995-2024, 950 AI/ML devices authorized by the FDA
 - Often not validated against clinical end points or tested with human operators even though by design are often “human in the loop” → challenging to determine clinical generalizability
- In European Union, higher bar for clinical outcomes
 - Of 173 commercially available devices:
 - Subspecialty: Neuro (34%) > Chest (31%) > Breast (10%)
 - Modality: CT (35%) > MR (27%) > XR (21%)
 - Function: Quantification, Detection, Diagnosis
 - Minority of papers reflecting proof of advanced levels of efficacy (diagnostic thinking, therapeutic efficacy, patient outcome efficacy, or societal efficacy)

Pathology



- Multimodal AI models – incorporating image-based, text-based data
- Clinical adoption lagging behind state-of-the-art research/industry
 - Limitations in transition from physical slides/manual interpretation to digital pathology to fully harness AI potential
- FDA-approved models often more for cancer-related applications (e.g. quantification, identifying areas on slides, metastases detection)

Psychiatry/Behavioral Health



- Potential to Transform Traditional Mental Healthcare Paradigm
 - Current state: questionnaires, psychiatric evaluation, symptoms
 - Future state: precision psychiatry for customized treatments, better diagnosis and prognostics, tying in neuroimaging, psychomotor, genetics data
- Chatbots
 - Commercial LLMs
 - National survey of US adolescents/young adults 12-21 (Nov 2025): 19.2% of US adults using AI chatbots for mental health advice, majority (63.3%) did not disclose chatbot use for this purpose to anyone
 - Healthcare system-supported chatbots for referral management

Conclusion



MOC REFLECTIVE STATEMENT (BRIEF TAKE HOME NOTES FOR REFERENCE)

- Different levels of AI:
 - Machine Learning (Unsupervised, Supervised, Reward, Semi-supervised)
 - Deep Learning
 - Large Language Models
- Generative AI applications for clinician experience include:
 - Ambient documentation
 - AI chart summarization
 - AI assistants
 - Other promising applications such as patient messaging/education
- In pathology and radiology: quantification and detection as most clinically promising areas



References

- Guo Y, Wang J, Hu D, et al. Evaluating ambient artificial intelligence documentation: effects on work efficiency, documentation burden, and patient-centered care. *J Am Med Inform Assoc*. 2025;(ocaf180):ocaf180. doi:10.1093/jamia/ocaf180
- Rodenhouse AJ, Davis E, Burger P, Nicandri G, DiGiovanni B. Utilization of an artificial intelligence-based documentation system improves provider efficiency in outpatient orthopaedic clinics: Reducing the afterhours burden of the electronic health record-A pilot study. *J Am Acad Orthop Surg*. Published online September 16, 2025. doi:10.5435/JAAOS-D-24-01321
- Afshar M, Resnik F, Baumann MR, et al. A novel playbook for pragmatic trial operations to monitor and evaluate ambient artificial intelligence in clinical practice. *NEJM AI*. 2025;2(9). doi:10.1056/aidbp2401267
- Guo Y, Hu D, Wang J, et al. Ambient listening in clinical practice: Evaluating EPIC Signal data before and after implementation and its impact on physician workload. *Stud Health Technol Inform*. 2025;329:653-657. doi:10.3233/SHTI250921
- Olson KD, Meeker D, Troup M, et al. Use of ambient AI scribes to reduce administrative burden and professional burnout. *JAMA Netw Open*. 2025;8(10):e2534976. doi:10.1001/jamanetworkopen.2025.34976
- Lawrence K, Kuram VS, Levine DL, et al. Informed consent for ambient documentation using generative AI in ambulatory care. *JAMA Netw Open*. 2025;8(7):e2522400. doi:10.1001/jamanetworkopen.2025.22400
- Wendt SJ, Dinh CT, Sutcliffe M, Jones K, Scanlan JM, Smitherman JS. Deploying ambient clinical intelligence to improve care: A research article assessing the impact of nuance DAX on documentation burden and burnout. *Future Healthcare Journal*. 2025;(100450):100450. doi:10.1016/j.fhj.2025.100450
- R. Wojda T, Vino Dominic Savio F, Hochheiser H, Amaravathi A, Bharat K, Fischer G. Understanding the human-computer interaction of ambient AI in healthcare: A qualitative study of early provider experiences. In: Stawicki DSP, Wojda MDTR, eds. *Artificial Intelligence in Medicine and Surgery - An Exploration of Current Trends, Potential Opportunities, and Evolving Threats, Volume 3 [Working Title]*. IntechOpen; 2025. doi:10.5772/intechopen.1011340
- Wright DS, Kanaparthi N, Melnick ER, et al. The effect of ambient artificial intelligence scribes on trainee documentation burden. *Appl Clin Inform*. Published online July 2, 2025. doi:10.1055/a-2647-1142
- Shin HS, Braykov NP, Jahan A, Meller J, Orenstein E. The influence of artificial intelligence scribes on clinician experience and efficiency among pediatric subspecialists: A rapid randomized quality improvement trial. *Appl Clin Inform*. Published online July 17, 2025. doi:10.1055/a-2657-8087
- Pelletier JH, Watson K, Michel J, McGregor R, Rush SZ. Effect of a generative artificial intelligence digital scribe on pediatric provider documentation time, cognitive burden, and burnout. *JAMIA Open*. 2025;8(4):oaf068. doi:10.1093/jamiaopen/oaf068
- Owens LM, Wilda JJ, Grifka R, Westendorp J, Fletcher JJ. Effect of ambient voice technology, natural language processing, and artificial intelligence on the patient-physician relationship. *Appl Clin Inform*. 2024;15(4):660-667. doi:10.1055/a-2337-4739
- Stults CD, Deng S, Martinez MC, et al. Evaluation of an ambient artificial intelligence documentation platform for clinicians. *JAMA Netw Open*. 2025;8(5):e258614. doi:10.1001/jamanetworkopen.2025.8614
- Shah SJ, Crowell T, Jeong Y, et al. Physician perspectives on ambient AI scribes. *JAMA Netw Open*. 2025;8(3):e251904. doi:10.1001/jamanetworkopen.2025.1904
- Hassan H, Zipursky AR, Rabbani N, et al. Special topic on burnout: Clinical implementation of artificial intelligence scribes in healthcare: A systematic review. *Appl Clin Inform*. 2025;0(AAM). doi:10.1055/a-2597-2017
- Tierney AA, Gayre G, Hoberman B, et al. Ambient artificial intelligence scribes: Learnings after 1 year and over 2.5 million uses. *NEJM Catalyst*. 2025;6(2):CAT.25.0040. doi:10.1056/CAT.25.0040
- Liu TL, Hetherington TC, Stephens C, et al. AI-powered clinical documentation and clinicians' electronic health record experience: A nonrandomized clinical trial: A nonrandomized clinical trial. *JAMA Netw Open*. 2024;7(9):e2432460. doi:10.1001/jamanetworkopen.2024.32460
- Liu TL, Hetherington TC, Dharod A, et al. Does AI-powered clinical documentation enhance clinician efficiency? A longitudinal study. *NEJM AI*. 2024;1(12). doi:10.1056/aioa2400659
- Cao MS DY, Jamie R. Silkey PA-C MPAS MHA MBA, Decker MC, Wanat KA. Artificial intelligence-driven digital scribes in clinical documentation: Pilot study assessing the impact on dermatologist workflow and patient encounters. *JAAD International*. 2024;15:149-151. doi:10.1016/j.jdin.2024.02.009
- Albrecht M, Shanks D, Shah T, et al. Enhancing clinical documentation with ambient artificial intelligence: a quality improvement survey assessing clinician perspectives on work burden, burnout, and job satisfaction. *JAMIA Open*. 2025;8(1). Accessed March 5, 2025. <https://academic.oup.com/jamiaopen/article/8/1/oaf013/8029407>
- Duggan MJ, Gervase J, Schoenbaum A, et al. Clinician experiences with ambient scribe technology to assist with documentation burden and efficiency. *JAMA Netw Open*. 2025;8(2):e2460637. doi:10.1001/jamanetworkopen.2024.60637
- Galloway JL, Munroe D, Vohra-Khullar PD, et al. Impact of an Artificial Intelligence-Based Solution on Clinicians' Clinical Documentation Experience: Initial Findings Using Ambient Listening Technology. *J Gen Intern Med*. 2024;39(13):2625-2627. doi:10.1007/s11606-024-08924-2
- Owens LM, Wilda JJ, Hahn PY, Koehler T, Fletcher JJ. The association between use of ambient voice technology documentation during primary care patient encounters, documentation burden, and provider burnout. *Fam Pract*. 2024;41(2):86-91. doi:10.1093/fampra/cmadv092
- Shah SJ, Devon-Sand A, Ma SP, et al. Ambient artificial intelligence scribes: physician burnout and perspectives on usability and documentation burden. *J Am Med Inform Assoc*. Published online December 5, 2024:ocae295. doi:10.1093/jamia/ocae295
- Haberle T, Cleveland C, Snow GL, et al. The impact of nuance DAX ambient listening AI documentation: a cohort study. *J Am Med Inform Assoc*. 2024;31(4):975-979. doi:10.1093/jamia/ocae022
- Misurac J, Knake LA, Blum JM. The Effect of Ambient Artificial Intelligence Notes on Provider Burnout. *Appl Clin Inform*. (AAM). doi:10.1055/a-2461-4576
- Ma SP, Liang AS, Shah SJ, et al. Ambient artificial intelligence scribes: utilization and impact on documentation time. *J Am Med Inform Assoc*. Published online December 17, 2024:ocae304. doi:10.1093/jamia/ocae304
- Tierney AA, Gayre G, Hoberman B, et al. Ambient Artificial Intelligence Scribes to Alleviate the Burden of Clinical Documentation. *NEJM Catal*. 2024;5(3):CAT.23.0404. doi:10.1056/CAT.23.0404



References

- Caballar R, Stryker C. What is model performance? November 17, 2025. Accessed November 19, 2025. <https://www.ibm.com/think/topics/model-performance>
- What is Deep Learning. Google Cloud. 2025. Accessed November 19, 2025. <https://cloud.google.com/discover/what-is-deep-learning>
- Chen S, Gao M, Sasse K, et al. When helpfulness backfires: LLMs and the risk of false medical information due to sycophantic behavior. *npj Digit Med*. 2025;8(1). doi:10.1038/s41746-025-02008-z
- Omar M, Sorin V, Collins JD, et al. Multi-model assurance analysis showing large language models are highly vulnerable to adversarial hallucination attacks during clinical decision support. *Commun Med (Lond)*. 2025;5(1):330. doi:10.1038/s43856-025-01021-3
- China CR. Five machine learning types to know. IBM. 2025. Accessed November 18, 2025. <https://www.ibm.com/think/topics/machine-learning-types>
- The Office of the National Coordinator for Health Information Technology. EHR Contracts Untangled: Selecting Wisely, Negotiating Terms, and Understanding the Fine Print. September 2016. Accessed November 18, 2025. https://www.healthit.gov/sites/default/files/EHR_Contracts_Untangled.pdf
- Business Associate Contracts. US Department of Health and Human Services. January 25, 2013. Accessed November 18, 2025. <https://www.hhs.gov/hipaa/for-professionals/covered-entities/sample-business-associate-agreement-provisions/index.html>
- Currey D, Tarabichi Y. External validation of a widely available, localizable sepsis early detection model in the emergency department setting. *Am J Emerg Med*. 2025;97:147-151. doi:10.1016/j.ajem.2025.07.056
- Wong A, Otles E, Donnelly JP, et al. External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. *JAMA Intern Med*. 2021;181(8):1065-1070. doi:10.1001/jamainternmed.2021.2626
- Hassan H, Zipursky AR, Rabbani N, et al. Special topic on burnout: Clinical implementation of artificial intelligence scribes in healthcare: A systematic review. *Appl Clin Inform*. 2025;0(AAM). doi:10.1055/a-2597-2017
- Saenz AD, Mass General Brigham AI Governance Committee, Centi A, et al. Establishing responsible use of AI guidelines: a comprehensive case study for healthcare institutions. *NPJ Digit Med*. 2024;7(1):348. doi:10.1038/s41746-024-01300-8
- You JG, Hernandez-Boussard T, Pfeffer MA, Landman A, Mishuris RG. Clinical trials informed framework for real world clinical implementation and deployment of artificial intelligence applications. *NPJ Digit Med*. 2025;8(1):107. doi:10.1038/s41746-025-01506-4



References

- Informatics: Research and practice. AMIA. 2023. Accessed July 11, 2023. <https://amia.org/about-amia/why-informatics/informatics-research-and-practice>
- Hersh, WR. 2022. Health Informatics: Practical Guide, 8th Edition. Lulu.com.
- Martineau K. What is Generative AI? IBM Research Blog. April 20, 2023. Accessed April 11, 2024. <https://research.ibm.com/blog/what-is-generative-AI>.
- Meskó B, Topol EJ. The imperative for regulatory oversight of large language models (or generative AI) in healthcare. *npj Digit. Med.* **6**, 120 (2023). <https://doi.org/10.1038/s41746-023-00873-0>
- Raza, M.M., Venkatesh, K.P. & Kvedar, J.C. Generative AI and large language models in health care: pathways to implementation. *npj Digit. Med.* **7**, 62 (2024). <https://doi.org/10.1038/s41746-023-00988-4>
- Peng, C., Yang, X., Chen, A. *et al.* A study of generative large language model for medical research and healthcare. *npj Digit. Med.* **6**, 210 (2023). <https://doi.org/10.1038/s41746-023-00958-w>
- Shasha Han, Tait D. Shanafelt, Christine A. Sinsky, et al. [Estimating the Attributable Cost of Physician Burnout in the United States](#). *Ann Intern Med.* 2019;170:784-790. [Epub 28 May 2019]. doi:[10.7326/M18-1422](https://doi.org/10.7326/M18-1422)
- Menon NK, Shanafelt TD, Sinsky CA, et al. Association of Physician Burnout With Suicidal Ideation and Medical Errors. *JAMA Netw Open.* 2020;3(12):e2028780. doi:10.1001/jamanetworkopen.2020.28780
- Budd J. Burnout Related to Electronic Health Record Use in Primary Care. *Journal of Primary Care & Community Health.* 2023;14. doi:[10.1177/21501319231166921](https://doi.org/10.1177/21501319231166921)
- Gaffney A, Woolhandler S, Cai C, et al. Medical Documentation Burden Among US Office-Based Physicians in 2019: A National Study. *JAMA Intern Med.* 2022;182(5):564-566. doi:10.1001/jamainternmed.2022.0372
- Lance M Owens, Joshua J Wilda, Peter Y Hahn, Tracy Koehler, Jeffrey J Fletcher, The association between use of ambient voice technology documentation during primary care patient encounters, documentation burden, and provider burnout, *Family Practice*, 2023;, cmad092, <https://doi-org.ezp-prod1.hul.harvard.edu/10.1093/fampra/cmad092>
- Gliadkovskaya A. Medical ai scribe startup nabla rolling out tool to the Permanente Medical Group Docs in Northern California. Fierce Healthcare. October 9, 2023. Accessed February 27, 2024. <https://www.fiercehealthcare.com/health-tech/medical-scribe-startup-nabla-rollout-tool-kaizer-permanente-docs>.
- Tierney Aaron A., Gayre Gregg, Hoberman Brian, et al. Ambient Artificial Intelligence Scribes to Alleviate the Burden of Clinical Documentation. *NEJM Catalyst.* 5(3):CAT.23.0404. doi:[10.1056/CAT.23.0404](https://doi.org/10.1056/CAT.23.0404)
- Liu TL, Hetherington TC, Dharod A, et al. Does AI-powered clinical documentation enhance clinician efficiency? A longitudinal study. *NEJM AI.* 2024;1(12). doi:10.1056/aioa2400659
- Ma SP, Liang AS, Shah SJ, et al. Ambient artificial intelligence scribes: utilization and impact on documentation time. *J Am Med Inform Assoc.* Published online December 17, 2024:ocae304. doi:10.1093/jamia/ocae304
- Shah SJ, Devon-Sand A, Ma SP, et al. Ambient artificial intelligence scribes: physician burnout and perspectives on usability and documentation burden. *J Am Med Inform Assoc.* Published online December 5, 2024:ocae295. doi:10.1093/jamia/ocae295
- Holmgren AJ, Fenton CL, Thombley R, et al. Ambient artificial intelligence scribes and physician financial productivity. *JAMA Netw Open.* 2026;9(1):e2553233. doi:10.1001/jamanetworkopen.2025.53233



References

McBain RK, Cantor JH, Breslau J, et al. AI chatbot use and disclosure for mental health among US adolescents and young adults. *JAMA Pediatr*. Published online June 1, 2026. doi:10.1001/jamapediatrics.2026.2015

Brodeur PG, Buckley TA, Kanjee Z, et al. Performance of a large language model on the reasoning tasks of a physician. *Science*. 2026;392(6797):524-527. doi:10.1126/science.adz4433

Rao AS, Esmail KP, Lee RS, et al. Large language model performance and clinical reasoning tasks. *JAMA Netw Open*. 2026;9(4):e264003. doi:10.1001/jamanetworkopen.2026.4003

Gross SA, Shaukat A, Afzali A, et al. Artificial intelligence for gastroenterology practice: A modified Delphi consensus. *Am J Gastroenterol*. 2026;121(4):1017-1024. doi:10.14309/ajg.0000000000003944

Zhou L, Pacchiardi L, Martínez-Plumed F, et al. General scales unlock AI evaluation with explanatory and predictive power. *Nature*. 2026;652(8108):58-67. doi:10.1038/s41586-026-10303-2

Everett SS, Bunning BJ, Jain P, et al. From tool to teammate in a randomized controlled trial of clinician-AI collaborative workflows for diagnosis. *NPJ Digit Med*. Published online March 18, 2026:1-11. doi:10.1038/s41746-026-02545-1

Huang T, Tse G, Pageler NM, Bannett Y. Large language models using clinical text in pediatrics: A scoping review: A scoping review. *JAMA Netw Open*. 2026;9(3):e262443. doi:10.1001/jamanetworkopen.2026.2443

Hussain A, Guni A, Gandhewar R, et al. Is the FDA regulation of cardiology AI devices supporting cardiovascular innovation: a scoping review. *Heart*. 2026;112(6):300-306. doi:10.1136/heartjnl-2025-326307

Antonissen N, Tryfonos O, Houben IB, Jacobs C, de Rooij M, van Leeuwen KG. Artificial intelligence in radiology: 173 commercially available products and their scientific evidence. *Eur Radiol*. 2026;36(1):526-536. doi:10.1007/s00330-025-11830-8

Angus DC, Khera R, Lieu T, et al. AI, health, and health care today and tomorrow: The JAMA Summit report on artificial intelligence: The JAMA Summit report on Artificial intelligence. *JAMA*. 2025;334(18):1650-1664. doi:10.1001/jama.2025.18490



Nguyen TD, Whaley CM, Simon K, et al. Adoption of artificial intelligence in the health care sector. *JAMA Health Forum*. 2025;6(11):e255029. doi:10.1001/jamahealthforum.2025.5029

Thank you!

Jacqueline You
jyou5@mgb.org





Mass General Brigham