CLRPD Report 2-I-23, Generative AI in Medicine and Health Care

EXECUTIVE SUMMARY

This report provides information on the fundamentals of generative AI in medicine and health care: terminologies and components of artificial intelligence (AI) and augmented intelligence, definitions, prominent models (Open AI ChatGPT, Google Bard and Med-PaLM, and Microsoft Bing), promises, challenges, and pitfalls, AMA partnerships and resources, and potential ethical and regulatory frameworks. The report concludes with insight from CLRPD members on the trend.

Generative AI models are commercial natural language processing tools known as large language models (LLMs). At their core, all AI innovations utilize sophisticated statistical techniques to discern patterns within extensive datasets using increasingly powerful computational technologies. Three components—big data, advanced statistical methods, and computing resources—have not only become available recently but are also being democratized and made accessible to at a pace unprecedented in previous technological innovations.

While LLMs show promise to make a significant contribution to health care in the future, physicians currently considering using generative AI models in a clinical setting or direct patient care should exercise caution and be aware of the real challenges that remain to ensure reliability: confident responses that are not justified by the model's training data, the "black box" nature of AI, biased and discriminatory tendencies in outputs, lack of knowledge-based reasoning, lack of current ethical and regulatory frameworks, patient privacy and security concerns, and potential liability.

Generative AI systems are not sentient, they simply use massive amounts of text to predict one word after another, and their outputs may mix truth with patently false statements. As such, physicians will need to learn how to integrate these tools into clinical practice, defining clear boundaries between full, supervised, and proscribed autonomy. Physicians should be clear-eyed about the risks inherent to any new technology, especially ones that carry existential implications, while cautiously optimistic about a future of improved health care system efficiency, better patient outcomes, and reduced burnout. Extant AI-assistant programs and rapidly developing systems are incredibly sophisticated, and as physicians have already begun to demonstrate on social media, they might soon be able to reliably perform test result notifications, work letters, prior authorizations, and the like—the mundane necessities that not only cumulatively consume valuable time but are substantial contributors to physician burnout.

Projecting further into an AI-enhanced future, imagine that instead of writing follow-up care instructions, physicians could ask a generative AI system to create a synopsis of the patient's treatment course. With the time saved, physicians could step away from the computer, face the patient, and explain the most salient follow-up items, prepped with materials that are compatible with best practices in health literacy. Likewise, these programs might help actualize the admirable intentions behind the provisions in the 21st Century Cures Act that have given patients access, but not accessibility, to their jargon-laden electronic medical records.

Given opportunities to offer clinical insight into the development and deployment of these systems, Generative AI may provide physicians with technological tools that reduce administrative burden and enable them to get back to the reason why they decided to pursue medicine in the first place—to improve patients' lives—meanwhile, improving physicians' wellbeing.

REPORT OF THE COUNCIL ON LONG RANGE PLANNING AND DEVELOPMENT

CLRPD Report 2-I-23

Subject: Generative AI in Medicine and Health Care

Presented by: Gary Thal, MD, Chair

BACKGROUND

The functions of the Council on Long Range Planning and Development (CLRPD) include to study and make recommendations concerning the long-range objectives of the American Medical Association (AMA), and to serve in an advisory role to the Board of Trustees concerning strategies by which the AMA attempts to reach its long-range objectives. To accomplish its role, the Council studies anticipated changes in the environment in which medicine and the AMA must function and develops memos to the Board, which include CLRPD deliberations and insight on emerging issues, such as generative artificial intelligence (AI).

This informational report presents material on the fundamentals of generative AI in medicine and health care including terminologies and components, definition, prominent models, promises and pitfalls, AMA partnerships and resources, potential ethical and regulatory frameworks, and CLRPD insight.

TERMINOLIGIES AND COMPONENTS OF AI

CLRPD Report 1-A-18, A Primer on Artificial and Augmented Intelligence[1] defines the relative terminologies of artificial intelligence (AI), which are not well understood:

- *Algorithms* are a sequence of instructions used to solve a problem. Developed by programmers to instruct computers in new tasks, algorithms are the building blocks of the advanced digital world. Computer algorithms organize enormous amounts of data into information and services, based on certain instructions and rules.

- *Artificial Intelligence* is the ability of a computer to complete tasks in a manner typically associated with a rational human being—a quality that enables an entity to function appropriately and with foresight in its environment. True AI is widely regarded as a program or algorithm that can beat the Turing Test, which states that an artificial intelligence must be able to exhibit intelligent behavior that is indistinguishable from that of a human.

- *Augmented Intelligence* is an alternative conceptualization that focuses on AI's assistive role, emphasizing the fact that its design enhances human intelligence rather than replaces it.

- *Machine Learning* is a part of the discipline of artificial intelligence and refers to constructing algorithms that can make accurate predictions about future outcomes. Machine learning can be supervised or unsupervised.

- In supervised learning, algorithms are presented with "training data" that contain examples with their desired conclusions, such as pathology slides that contain cancerous cells as well as slides that do not.
- Unsupervised learning does not typically leverage labeled training data. Instead, algorithms are tasked with identifying patterns in data sets on their own by defining signals and potential abnormalities based on the frequency or clustering of certain data.

- *Deep Learning* is a subset of machine learning that employs artificial neural networks (ANNs) and algorithms structured to mimic biological brains with neurons and synapses. ANNs are often constructed in layers, each of which performs a slightly different function that contributes to the result. Deep learning is the study of how these layers interact and the practice of applying these principles to data.

- *Cognitive Computing*, a term coined by IBM, is often used interchangeably with machine learning and artificial intelligence. However, cognitive computing systems do not necessarily aspire to imitate intelligent human behavior, but instead to supplement human decision-making power by identifying potentially useful insights with a high degree of certainty. Clinical decision support and augmented intelligence come to mind when considering this definition.

- *Natural Language Processing* (NLP) forms the foundation for many cognitive computing exercises. The ingestion of source materials, such as medical literature, clinical notes, or audio dictation records requires a computer to understand what is written, spoken, or otherwise being communicated. One commonly used application of NLP is optical character recognition (OCR) technology that can turn static text, such as a PDF of a lab report or a scan of a handwritten clinical note, into machine readable data. Once data is in a workable format, the algorithm parses the meaning of each element to complete a task such as translating into a different language, querying a database, summarizing information, or supplying a response to a conversation partner. In the health care field, where acronyms and abbreviations are common, accurately parsing through this "incomplete" data can be challenging.

DEFINTION OF GENERATIVE AI

Generative AI is a broad term used to describe any type of artificial intelligence that can be used to create new text, images, video, audio, code, or synthetic data. Progress with generative AI was relatively slow until around 2012, when a single idea shifted the entire field. It was called a neural network—inspired by the inner workings of the human brain—a mathematical system that learns skills by finding statistical patterns in enormous amounts of data. By analyzing thousands of cat photos, for instance, it can learn to recognize a cat. Neural networks enable Siri and Alexa to understand what you are saying, identify people and objects in Google Photos and instantly translate dozens of languages.[2]

The next big change was large language models (LLMs), which consist of a neural network. Around 2018, companies like Google, Microsoft, and OpenAI began building neural networks that were trained on vast amounts of text from the internet, including Wikipedia articles, digital books, and academic papers. Somewhat to the experts' surprise, these systems learned to write unique prose and computer code and carry-on sophisticated conversations, which is termed generative AI.[3]

LLMs are a class of technologies that drive generative AI systems. The first LLMs appeared about five years ago, but were not very sophisticated; however, today they can draft emails, presentations, and memos. Every AI system needs a goal. Researchers call this an objective function. It can be simple, such as “win as many chess games as possible” or complicated, such as “predict the three-dimensional shapes of proteins, using only their amino acid sequences.”[4] Most LLMs have the same basic objective function, which is, given a sequence of text, to guess what comes next. Though trained on simple tasks along the lines of predicting the next word in a sentence, neural language models with sufficient training and parameter counts are found to capture much of the syntax and semantics of human language. In addition, LLMs demonstrate considerable general knowledge about the world and can memorize a great quantity of facts during training.

Training the model involves feeding algorithms large amounts of data, which serves as the foundation for the AI model to learn from. This can consist of text, code, graphics, or any other types of content relevant to the task at hand. Once the training data has been collected, the AI model analyzes the patterns and relationships within the data to understand the underlying rules governing the content. Continuously, the AI model fine-tunes its parameters as it learns, improving its ability to simulate human-generated content. The more content the AI model generates, the more sophisticated and convincing its outputs become.[5]

Typing in the precise words and framing to generate the most helpful answers is an art. Beginning a prompt with “act as if” will instruct the model to emulate an expert. For example, typing “Act as if you are a tutor for the SATs” or “Act as if you are a personal trainer” will guide the systems to model themselves around people in those professions. These prompts provide additional context for the generative AI model to produce its response by helping the tool to draw on specific statistical patterns in its training data.[6]

Generative AI outputs are calibrated combinations of the data used to train the algorithms. Because the amount of data used to train these algorithms is so incredibly massive—multiple terabytes of text data—the models can appear to be “creative” when producing outputs. Moreover, the models usually have random elements, which means they can produce a variety of outputs from one input request—making them seem even more lifelike. The unmanageably huge volume and complexity of data (unmanageable by humans, anyway) that is now being generated has increased the potential of the technologies.[7]

Tech companies are confronting a challenge: how to balance asking users for more data to deliver new AI features without scaring away privacy-conscious businesses and consumers that consistently tell pollsters they want transparency about when AI is used and trained. But when companies provide such detail, it is often written in legalese and buried in fine print that is often being rewritten to give tech companies more rights. Video conferencing company Zoom encountered a massive backlash over concerns the contents of video chat might be used to train AI systems. The move prompted an apologetic post from Zoom’s CEO, but the company is far from alone in seeking more consumer data to train AI models. Companies are deploying different approaches to ensure they have access to user data. At the same time, many are also adding in language to prevent anyone else from scraping their websites to train AI systems.[8]

According to the JAMA Forum article, “ChatGPT and Physicians’ Malpractice Risk,”[9] most LLMs are trained on indiscriminate assemblages of web text with little regard to how sources vary in reliability. They treat articles published in the *New England Journal of Medicine* and Reddit discussions as equally authoritative. In contrast, Google searches let physicians distinguish expert from inexpert summaries of knowledge and selectively rely on the best. Other decision-support tools provide digests based on the best available evidence. Although efforts are underway[10] to train LLMs on exclusively authoritative, medically relevant texts, they are still nascent and prior efforts have faltered.[11]

Generative AI models have been observed to experience-confabulations or delusions— confident responses by an AI model that does not seem to be justified by its training data. Such phenomena are termed by the tech industry as "hallucinations," in loose analogy with the phenomenon of hallucination in human psychology; however, one key difference is that human hallucinations are usually associated with false percepts, while an AI hallucination is associated with the category of unjustified responses or beliefs.[12]

GENERATIVE AI MODELS

There are several types of generative AI models, each designed to address specific challenges and applications. These generative AI models can be broadly categorized into the following types:[13]

- Transformer-based models: These models, such as OpenAI's ChatGPT and GPT-3.5, are neural networks designed for natural language processing. They are trained on large amounts of data to learn the relationships between sequential data — like words and sentences — making them useful for text-generation tasks.

- Generative adversarial networks (GANs): GANs are made up of two neural networks, a generator, and a discriminator that work in a competitive or adversarial capacity. The generator creates data, while the discriminator evaluates the quality and authenticity of said data. Over time, both networks get better at their roles, leading to more realistic outputs.

- Variational autoencoders (VAEs): VAEs use an encoder and a decoder to generate content. The encoder takes the input data, such as images or text, and simplifies it into a more compact form. The decoder takes this encoded data and restructures it into something new that resembles the original input.

- Multimodal models: Multimodal models can process multiple types of input data, including text, audio, and images. They combine different modalities to create more sophisticated outputs, such as DALL-E 2[14] and OpenAI's GPT-4[15], which is also capable of accepting image and text inputs.

OpenAI ChatGPT

Researchers have been working on generative AI for a long time. OpenAI, developer of ChatGPT (Generative Pretrained Transformer), is over seven years old. Launched in November 2022, ChatGPT is a LLM that leverages huge amounts of data to mimic human conversation and assess language patterns. Currently, the basic system is free via a simple web interface that lets users pose questions and give directions to a bot that can answer with conversation, term papers, sonnets, recipes—almost anything.[16]

GPT-4 is the newest version of OpenAI's language model systems, and it is much more advanced than its predecessor GPT-3.5, which ChatGPT runs on. GPT-4 is a multimodal model that accepts both text and images as input and output text. This can be useful for uploading worksheets, graphs, and charts to be analyzed. GPT-4 has advanced intellectual capabilities that allow it to outperform

GPT-3.5 in a series of simulated benchmark exams. It has also reduced the number of "hallucinations" produced by the chatbot.[17]

ChatGPT has passed a series of benchmark exams. Christian Terwiesch, a professor at Wharton, the University of Pennsylvania's business school, used ChatGPT to take an MBA exam. ChatGPT not only passed the exam but also scored a B to B-. The professor was impressed at its basic operations management, process analysis questions, and explanations. Although ChatGPT could pass many of these benchmark exams, its scores were usually in the lower percentile. However, with GPT-4, scores were much higher. For example, ChatGPT in the 3.5 series scored in the lower 10th percentile of a simulated Bar Exam, while GPT-4 scored in the top 10th percentile.[18]

Google Bard and Med-PaLM

Bard is Google's AI chat service, a rival to ChatGPT.[19] On February 6, 2023, Google introduced its experimental AI chat service. Over a month after the announcement, Google began rolling out access to Bard via a waitlist. Bard uses a lightweight version of Google's Language Model for Dialogue Applications (LaMDA)[20] and draws on all the information from the web to respond -- a stark contrast from ChatGPT, which does not have internet access. Google's chat service had a rough launch, with a demo of Bard delivering inaccurate information about the James Webb Space Telescope.[21] ChatGPT's advanced capabilities exceed those of Google Bard. Even though Google Bard has access to the internet and ChatGPT does not, it fails to produce answers much more often than ChatGPT.

In April 2023, Google announced a new version of its medical LLM, called Med-PaLM 2.[22] An AI platform for analyzing medical data, it aims to assist physicians with routine tasks and provide more reliable answers to patient questions than "Dr. Google." *PaLM 2, the Pathways Language Model*, is more critical than Bard for medicine. With 540 billion parameters, it draws knowledge from scientific papers and websites, can reason logically, and perform complex mathematical calculations.[23] Google is actively developing its large language model (LLM), Med-PaLM 2, which they anticipate will excel at healthcare discussions over general-purpose algorithms, given its training on questions and answers from medical licensing exams. They are collaborating with Mayo Clinic and other health systems and partnering with the healthcare technology vendor, CareCloud.[24]

Microsoft Bing AI

In early February 2023, Microsoft unveiled[25] a new version of Bing[26] -- and its standout feature is its integration with GPT-4. When it was announced, Microsoft shared that Bing Chat was powered by a next-generation version of OpenAI's large language model, making it "more powerful than ChatGPT."[27]

Five weeks after launch, Microsoft revealed that, since its launch, Bing Chat had been running on GPT-4, the most advanced Open AI model, before the model even launched. Because Bing's ChatGPT is linked to the internet, the biggest difference from ChatGPT is that Bing's version has information on current events, while ChatGPT is limited to knowledge before 2021. Another major advantage of the new Bing is that it links to the sites it sourced its information from using footnotes, whereas ChatGPT does not.

Building a generative AI model has for the most part been a major undertaking, to the extent that only a few well-resourced tech heavyweights have tried. OpenAI, the company behind ChatGPT, former GPT models, and DALL-E (a tool for AI-generated art), has billions in funding from highprofile donors. DeepMind is a subsidiary of Alphabet, the parent company of Google, and Meta has released its Make-A-Video product based on generative AI. These companies employ some of the world's best computer scientists and engineers. However, when you are asking a model to train using nearly the entire internet, it is going to be costly. OpenAI has not released exact costs, but estimates indicate that GPT-3 was trained on a vast amount of text data that was equivalent to one million feet of bookshelf space, or a quarter of the entire Library of Congress at an estimated cost of several million dollars. These are not resources that your garden-variety start-up can access.[28]

PROMISES AND PITFALLS

The latest McKinsey Global Survey breaks down how corporate leaders worldwide are using generative AI. By interviewing thousands of managers and executives across the globe, McKinsey gained a high-level view on where AI is being deployed already (especially in marketing, product development, and service operations), as well as the biggest perceived risks of implementing AI (including inaccurate outputs, cybersecurity threats, and intellectual property infringement).[29] In June, McKinsey projected that generative AI could add $4.4 trillion to global GDP, 75% of which would emerge from use cases in customer operations, marketing and sales, software engineering, and R&D.[30]

In the medical device industry, product developers are integrating AI capabilities into a wide variety of health care technologies, from imaging and surgical systems to vital sign monitors, endoscopes, and diagnostic devices. New players range from Big Tech behemoths to entrepreneurial startups to the individual visionaries who, in the digital age, create algorithms that could lead to the next breakthrough technology.

AMA surveys of physicians conducted in 2016, 2019, and 2022 show growing use of and plans to use AI in the short term. In the latest survey, nearly one in five physicians say their practice incorporates AI for practice efficiencies and clinical applications, while just over one in 10 use biometrics, precision and personalized medicine, or digital therapeutics. More than twice as many expect to adopt such advanced technologies within one year. However, unlike other health care technologies, AI-enabled medical devices can perform in mysterious and unexpected ways—introducing a whole new set of uncertainties. This so-called "black box conundrum"—knowing what goes in and what comes out of the system, but not what happens in between—can be disconcerting.[31]

In 2021, two experts explained the fundamentals of machine learning, what it means in the clinical setting and the possible risks of using the technology, "Machine Learning: An Introduction and Discussion of Medical Applications" that took place during the June 2021 AMA Sections Meetings and was hosted by AMA Medical Student Section:[32]

- A key aspect of machine learning is that it continuously improves the model by weighing the data with minimal human interaction, explained Herbert Chase, MD, MA, professor of clinical medicine in biomedical informatics at Vagelos College of Physicians and Surgeons at Columbia University. It may be able to pick up factors leading to disease that a physician does not. For example, people who all worked in a factory that had heavy metals in the atmosphere or people in the same zip code are experiencing the same thing. People with a certain disease are taking the same vitamins or they all had a previous surgery. "The EHR has hundreds of different attributes, thousands of different values that can be mined. This is classic data mining in an unsupervised way to make the prediction model better and there are many examples in the literature now of how this approach has dramatically improved the prediction for coronary artery disease, heart failure and many other chronic conditions," Dr. Chase said.

- While machine learning can help medicine in tremendous ways, physicians must also be mindful that bias in machine learning is a problem, Ravi Parikh, MD, MPP, assistant professor of medical ethics and health policy and medicine at the University of Pennsylvania, explained during the educational session. There are three distinct things you need to specify for a supervised machine-learning algorithm. You start with a population. A series of variables is derived from the population. Those variables are then used for a predictive algorithm to predict an outcome.

- "Any amount of those three steps could be biased and could generate bias in the context of the algorithm," Dr. Parikh said. So, how can bias be addressed? Dr. Parikh said physicians can identify bias and potentially flawed decision making in real time, use unbiased data sources and track algorithm outputs continuously to monitor bias.

- Drs. Parikh and Chase said physicians do not need to worry about machine learning eliminating physicians' jobs. "The workforce will just be the same as it always has been … but you will be operating at a higher level and I think that will make the profession to some extent more interesting," Dr. Chase said.

Augmented intelligence promises to be a transformational force in health care, especially within primary care. Experts outline ways that innovations driven by this technology can aid rather than subvert the patient-physician relationship. Steven Y. Lin, MD, and Megan R. Mahoney, MD, associate clinical professor of medicine and clinical professor of medicine, respectively, in the Division of Primary Care and Population Health at Stanford University School of Medicine, and AMA vice president of professional satisfaction Christine A. Sinsky, MD—reviewed promising inventions in 10 distinct problem areas:[33]

- Risk prediction and intervention: Drawing on EHR data, AI-driven predictive modeling can outperform traditional predictive models in forecasting in-hospital mortality, 30-day unplanned readmission, prolonged length of stay and final discharge diagnoses.

- Population health management: With the move from fee-for-service to value-based payments, AI could help identify and close care gaps and optimize performance with Medicare quality payment programs.

- Medical advice and triage: Some companies have developed "AI doctors" to provide health advice to patients with common symptoms, freeing up primary care appointments for patients requiring more complex care. "Rather than replacing physicians for some conditions, AI support can be integrated into team-based care models that make it easier for primary care physicians to manage a patient panel," the authors wrote. Risk-adjusted paneling and resourcing EHR data on utilization can be used to create algorithms for weighing panel sizes in primary care. This can be used to determine the level of staffing support needed for primary care practices based on the complexity and intensity of care provided.

- Device integration: Wearable devices can track vital signs and other health measures, but their data's volume and its incompatibility with EHRs make it unwieldy without the help of AI. Apple's Health Kit is a tool that integrates data from multiple wearable devices into the EHR, enabling care teams to map trends and spot deviations that suggest illness.

- Digital health coaching: Companies are now offering digital health coaching for diabetes, hypertension and obesity, and similar programs integrated in health systems have shown reductions in cost per patient through reduced office and hospital visits.

- Chart review and documentation: Technology companies with expertise in automatic speech recognition are teaming up with health systems to develop AI-driven digital scribes that can listen in on patient-physician conversations and automatically generate clinical notes in the EHR.

- Diagnostics: AI-powered algorithms for diagnosing disease "are now outperforming physicians in detecting skin cancer, breast cancer, colorectal cancer, brain cancer and cardiac arrhythmias," the authors wrote, citing numerous tools, such as IDx-DR, Aysa, and Tencent. "This could reduce the need for unnecessary referrals, increase continuity with patients and enhance mastery for primary care physicians."[34]
- Clinical decision-making: Next generation platforms do much more than provide alerts and best practice advisories. eClinicalWorks, for example, is developing a new version of its EHR that will feature an AI assistant that provides evidence-based clinical suggestions in real time.

- Practice management: AI can also automate repetitive clerical tasks. Eligibility checks, insurance claims, prior authorizations, appointment reminders, billing, data reporting and analytics can all now be automated using AI, and some companies have developed AI-powered category auditors to help optimize coding for quality payment programs.

AMA partners with technology and health care leaders to bring physicians critical insights on AI's potential applications and ensure that physicians have a voice in shaping AI's role in medicine.

- Health2047, the innovation subsidiary of the American Medical Association (AMA), has launched a startup that develops augmented intelligence technologies to support clinical decision making.[35] Called RecoverX, the startup creates technologies that leverage research, medical charts, patient conversations, and test results to provide evidence-based clinical insights and suggested actions for clinicians in real time. For example, one of the technologies on the core RecoverX platform, called Diagnostic Glass, provides decision-making support to clinicians in more than 30 specialties.[36]

- To develop actionable guidance for trustworthy AI in health care, the AMA reviewed literature on the challenges health care AI poses and reflected on existing guidance. These findings are published in a paper in *Journal of Medical Systems*: Trustworthy Augmented Intelligence in Health Care.[37]

- The AMA Intelligent Platform's CPT® Developer Program allows developers to access the latest content and resources, Access the Developer Portal on the AMA Intelligent Platform.[38]

- Kimberly Lomis, MD, AMA vice president of undergraduate medical innovations, co-authored a discussion paper, Artificial Intelligence for Health Professions Educators in *NAM Perspectives.*[39]

The technological capacity exists to use AI algorithms and tools to transform health care, but real challenges remain in ensuring that tools are developed, implemented and maintained responsibly, according to a *JAMA* Viewpoint column, "Artificial Intelligence in Health Care: A Report From the National Academy of Medicine."[40] The NAM report recommends that people developing, using, implementing, and regulating health care AI do seven key things:[41]

- Promotion of population-representative data with accessibility, standardization and quality is imperative: This is the way to ensure accuracy for all populations. While there is a lot of data now, there are issues with data quality, appropriate consent, interoperability, and scale of data transfers.

- Prioritize ethical, equitable and inclusive medical AI while addressing explicit and implicit bias: Underlying biases need to be scrutinized to understand their potential to worsen or address existing inequity and whether and how it should be deployed.

- Contextualize the dialogue of transparency and trust, which means accepting differential needs: AI developers, implementers, users, and regulators should collaboratively define guidelines for clarifying the level of transparency needed across a spectrum and there should be a clear separation of data, performance, and algorithmic transparency.

- Focus in the near term on augmented intelligence rather than AI autonomous agents: Fully autonomous AI concerns the public and faces technical and regulatory challenges. Augmented intelligence—supporting data synthesis, interpretation and decision-making by clinicians and patients—is where opportunities are now.

- Develop and deploy appropriate training and educational programs: Curricula must be multidisciplinary and engage AI developers, implementers, health care system leadership, frontline clinical teams, ethicists, humanists, patients, and caregivers.

- Leverage frameworks and best practices for learning health care systems, human factors, and implementation science: Health care delivery systems should have a robust and mature information technology governance strategy before embarking on a substantial AI deployment and integration.

- Balance innovation with safety through regulation and legislation to promote trust: AI developers, health system leaders, clinical users, and informatics and health IT experts should evaluate deployed clinical AI for effectiveness and safety based on clinical data.

The AMA recently developed a ChatGPT primer for physicians with questions regarding the technology and use in medical practice. The primer outlines considerations for physicians and patients when considering utilizing the tool and is available on the AMA website.[42]

Researchers from the University of Arizona Health Sciences found that patients are almost evenly split about whether they would prefer a human clinician or an AI-driven diagnostic tool, with preferences varying based on patient demographics and clinician support of the technology.[43] The results of the study, demonstrated that many patients do not believe that the diagnoses provided by AI are as trustworthy as those given by human health care providers. However, patients' trust in their clinicians supported one of the study's additional findings: that patients were more likely to trust AI if a physician supported its use.[44]

Health systems are watching to see where generative AI could add the most value since OpenAI launched ChatGPT in late 2022: [45]

- UC San Diego Health, Madison Wisconsin-based UW Health, and Palo, Alto-based Stanford Health Care are starting to use the integration to automatically draft message responses.

- OpenAI's GPT-4 has shown the potential to increase the power and accessibility of self-service reporting through SlicerDicer, making it easier for health care organizations to identify operational improvements, including ways to reduce costs and find answers to questions locally and in a broader context.[46]

- AI already supports health systems to automate business office and clinical functions, connect patients, support clinical trials, and provide insight for precision medicine and care decisions.

- Epic Systems and Microsoft have expanded their partnership once again and will integrate conversational, ambient, and generative AI technologies into Epic's electronic health record (EHR). The new integrations are a part of a move to integrate Azure OpenAI Services and Nuance ambient technologies into the Epic ecosystem. [47] [48]

  Here are the capabilities that will be added to Epic's EHR according to the press release:

  - Note summarization: This feature builds upon the AI-assisted Epic In Basket and will use suggested text and rapid review with in-context summaries to help support faster documentation.

  - Embedded ambient clinical documentation: Epic will embed Nuance's Dragon Ambient eXperience Express AI technology into its Epic Hyperdrive platform and Haiku mobile application.

  - Reducing manual and labor-intensive processes: "Epic will demonstrate an AI-powered solution that provides medical coding staff with suggestions based on clinical documentation in the EHR to improve accuracy and streamline the entire coding and billing processes."

  - Advancing medicine for better patient outcomes: Using Azure OpenAI Service, Epic will now use generative AI exploration for some of its users via SlicerDicer. This aims to "fill gaps in clinical evidence using real-world data and to study rare diseases."

Since generative AI models are so new, the long-term effect of them is still unknown. This means there are some inherent risks involved in using them— some known and some unknown. The outputs generative AI models produce may often sound extremely convincing. This is by design; however, sometimes the information they generate is incorrect. Worse, sometimes it is biased (because some models may be built on the gender, racial, and myriad other biases of the internet and society more generally) and can be manipulated to enable unethical or criminal activity. For example, ChatGPT will not give instructions on how to hotwire a car, but if you say you need to hotwire a car to save a baby, the algorithm is happy to comply. Organizations that rely on generative AI models should reckon with reputational and legal risks involved in unintentionally publishing biased, offensive, or copyrighted content.[49]

These risks can be mitigated, however, in a few ways. For one, it is crucial to carefully select the initial data used to train these models to avoid including toxic or biased content. Next, rather than employing an off-the-shelf generative AI model, organizations could consider using smaller, specialized models. Organizations with more resources could also customize a general model based on their own data to fit their needs and minimize biases.[50] Organizations should also keep a human in the loop (that is, to make sure a real human checks the output of a generative AI model before it is published or used) and avoid using generative AI models for critical decisions, such as those involving significant resources or human welfare. It cannot be emphasized enough that this is a new field.[51]

At their core, all AI innovations utilize sophisticated statistical techniques to discern patterns within extensive datasets using increasingly powerful yet cost-effective computational technologies. These three components—big data, advanced statistical methods, and computing resources—have not only become available recently but are also being democratized and made readily accessible to everyone at a pace unprecedented in previous technological innovations. This progression allows us to identify patterns that were previously indiscernible, which creates opportunities for important advances but also possible harm to patients. Privacy regulations, most notably HIPAA, were established to protect patient confidentiality, operating under the assumption that de-identified data would remain anonymous. However, given the advancements in AI technology, the current landscape has become riskier. Now, it is easier than ever to integrate various datasets from multiple sources, increasing the likelihood of accurately identifying individual patients.[52]

Researchers at Mack Institute for Technological Innovation – The Wharton School, University of Pennsylvania Cornell Tech, and Johnson College of Business – Cornell University found that despite their remarkable performance, LLMs sometimes produce text that is semantically or syntactically plausible but is, in fact, factually incorrect or nonsensical (i.e., hallucinations). The models are optimized to generate the most statistically likely sequences of words with an injection of randomness. They are not designed to exercise any judgment on the veracity or feasibility of the output. Further, the underlying optimization algorithms provide no performance guarantees, and their output can thus be of inconsistent quality. Hallucinations and inconsistency are critical flaws that limit the use of LLM-based solutions to low-stakes settings or in conjunction with expensive human supervision. To achieve high variability in quality and high productivity, most research on ideation and brainstorming recommends enhancing performance by generating many ideas while postponing evaluation or judgment of ideas (Girotra et al., 2010). This is hard for human ideators to do, but LLMs are designed to do exactly this— quickly generate many somewhat plausible solutions without exercising much judgment. Further, the hallucinations and inconsistent behavior of LLMs increase the variability in quality, which, on average, improves the quality of the best ideas. For ideation, an LLM's lack of judgment and inconsistency could be prized features, not bugs. Thus, the researchers hypothesize that LLMs will be excellent ideators.[53]

The landscape of risks and opportunities is likely to change rapidly in the coming weeks, months, and years. New use cases are being tested monthly, and new models are likely to be developed in the coming years. As generative AI becomes increasingly, and seamlessly, incorporated into business, society, and our personal lives, we can also expect a new regulatory climate to take shape. As organizations begin experimenting—and creating value—with these tools, physicians will do well to keep a finger on the pulse of benefits and drawbacks with the use of generative AI in medicine and health care.[54]

ETHICS FRAMEWORK FOR USE OF GENERATIVE AI IN HEALTH CARE

A new paper published by leading Australian AI ethicist Stefan Harrer PhD proposes for the first time a comprehensive ethical framework for the responsible use, design, and governance of Generative AI applications in health care and medicine. The study highlights and explains many key applications for health care:[55]

- assisting clinicians with the generation of medical reports or preauthorization letters,
- helping medical students to study more efficiently,
- simplifying medical jargon in clinician-patient communication,
- increasing the efficiency of clinical trial design,
- helping to overcome interoperability and standardization hurdles in EHR mining,
- making drug discovery and design processes more efficient.

However, the paper also highlights that the inherent danger of LLM-driven generative AI arising from the ability of LLMs to produce and disseminate false, inappropriate, and dangerous content at unprecedented scale is increasingly being marginalized in an ongoing hype around the recently released latest generation of powerful LLM systems authoritatively and convincingly.

Dr. Harrer proposes a regulatory framework with 10 principles for mitigating the risks of generative AI in health care:

1. Design AI as an assistive tool for augmenting the capabilities of human decision makers, not for replacing them.
2. Design AI to produce performance, usage and impact metrics explaining when and how AI is used to assist decision making and scan for potential bias.
3. Study the value systems of target user groups and design AI to adhere to them.
4. Declare the purpose of designing and using AI at the outset of any conceptual or development work.
5. Disclose all training data sources and data features.
6. Design AI systems to label any AI-generated content clearly and transparently as such.
7. Ongoingly audit AI against data privacy, safety, and performance standards.
8. Maintain databases for documenting and sharing the results of AI audits, educate users about model capabilities, limitations, and risks, and improve performance and trustworthiness of AI systems by retraining and redeploying updated algorithms.
9. Apply fair-work and safe-work standards when employing human developers.
10. Establish legal precedence to define under which circumstances data may be used for training AI, and establish copyright, liability, and accountability frameworks for governing the legal dependencies of training data, AI-generated content, and the impact of decisions humans make using such data.

Dr. Harrer said, “Without human oversight, guidance and responsible design and operation, LLM-powered generative AI applications will remain a party trick with substantial potential for creating and spreading misinformation or harmful and inaccurate content at unprecedented scale.” He predicts that the field will move from the current competitive LLM arms race to a phase of more nuanced and risk-conscious experimentation with research-grade generative AI applications in health, medicine, and biotech, which will deliver first commercial product offerings for niche applications in digital health data management within the next 2 years. “I am inspired by thinking about the transformative role generative AI and LLMs could one day play in health care and medicine, but I am also acutely aware that we are by no means there yet and that despite the prevailing hype, LLM-powered generative AI may only gain the trust and endorsement of clinicians and patients if the research and development community aims for equal levels of ethical and technical integrity as it progresses this transformative technology to market maturity."

"Ethical AI requires a lifecycle approach from data curation to model testing, to ongoing monitoring. Only with the right guidelines and guardrails can we ensure our patients benefit from emerging technologies while minimizing bias and unintended consequences," said John Halamka, MD, MS, President of Mayo Clinic Platform, and a co-founder of the Coalition for Health AI (CHAI).[56]

"This study provides important ethical and technical guidance to users, developers, providers, and regulators of generative AI and incentivizes them to responsibly and collectively prepare for the transformational role this technology could play in health and medicine," said Brian Anderson, MD, Chief Digital Health Physician at MITRE.[57]

REGULATORY FRAMEWORK FOR USE OF GENERATIVE AI IN MEDICINE

AMA's President Jesse Ehrenfeld, MD, MPH co-chairs the AI committee of the Association for the Advancement of Medical Instrumentation (AAMI)[58] and co-authored an article, "Artificial Intelligence in Medicine & ChatGPT: De-Tether the Physician," published in the *Journal of Medical Systems.* He says, "A competitive marketplace requires regulatory flexibility from the Federal Drug Administration (FDA). Regulation of AI systems is still in its infancy but AI that improves physician workflow should require less regulatory oversight than algorithms that make diagnoses, recommend treatments, or otherwise impact clinical decision making. While AI algorithms may one day independently learn to read CT scans, identify skin lesions, and provide medical diagnoses, the low-hanging fruit is in improving physician efficiency, e.g., de-tethering clinicians from the computer. This should be embraced by the health care industry now." Physicians have a critical role to play in this endeavor. Without physician knowledge, expertise and guidance on design and deployment, most of these digital innovations will fail, he predicted. They will not be able to achieve their most basic task of streamlining workflows and improving patient outcomes.

Dr. Ehrenfeld said, the AMA is working closely with the FDA to support efforts that create new pathways and approaches to regulate AI tools:

- Any regulatory framework should ensure that only safe, clinically validated, high-quality tools enter the marketplace. "We can't allow AI to introduce additional bias" into clinical care, cautioning that this could erode public confidence in the tools that come to the marketplace.[59]

- There also needs to be a balance between strong oversight and ensuring the regulatory system is not overly burdensome to developers, entrepreneurs, and manufacturers, "while also thinking about how we limit liability in appropriate ways for physicians," added Dr. Ehrenfeld.

- The FDA has a medical device action plan on AI and machine-learning software that would enable the agency to track and evaluate a software product from premarket development to post market performance.[60] The AMA has weighed in on the plan, saying the agency must guard against bias in AI and focus on patient outcomes.[61]

In April 2023, the European Union (EU) proposed new copyright rules for generative AI.[62] In its most recent AI Act, the EU requires that AI-generated content be disclosed to consumers to prevent copyright infringement, illegal content, and other malfeasance related to end-user lack of understanding about these systems.[63] As more chatbots mine, analyze, and present content in accessible ways for users, findings are often not attributable to any one or multiple sources, and despite some permissions of content use granted under the fair use doctrine in the United States that protects copyright-protected work, consumers are often left in the dark around the generation and explanation of the process and results.[64]

In the United States, the U.S. Food and Drug Administration (FDA) published a regulatory framework for AI applications in medicine in April 2019 and an action plan in January 2021. The FDA's leadership role in formulating regulatory guidance is a manifestation of the broader U.S. national approach to the regulation of AI. In contrast to the EU, the U.S. policy sustains from broad and comprehensive regulation of AI and instead delegates responsibilities to specific federal agencies, with an overarching mandate to avoid overregulation and promote innovation.[65]

CLRPD DISCUSSION

Generative AI systems are not sentient, they simply use massive amounts of text to predict one word after another, and their outputs may mix truth with patently false statements. As such, physicians will need to learn how to integrate these tools into clinical practice, defining clear boundaries between full, supervised, and proscribed autonomy. Physicians should be clear-eyed about the risks inherent to any new technology, especially ones that carry existential implications, while cautiously optimistic about a future of improved health care system efficiency, better patient outcomes, and reduced burnout.

Extant AI-assistant programs and rapidly developing systems are incredibly sophisticated, and as physicians have already begun to demonstrate on social media, they might soon be able to reliably perform test result notifications, work letters, prior authorizations, and the like—the mundane necessities that not only cumulatively consume valuable time but are a substantial contributor to physician burnout.

Projecting further into an AI-enhanced future, imagine that instead of writing discharge instructions, physicians could ask a generative AI system to create a synopsis of the patient's hospital course. With the time saved, physicians could step away from the computer, go to the patient's room, and explain the most salient follow-up items face-to-face, prepped with materials that are compatible with best practices in health literacy. Integrating AI into routine clinical practice will require careful validation, training, and ongoing monitoring to ensure its accuracy, safety, and effectiveness in supporting physicians to deliver care. While AI can be an asset in the medical field, it cannot replace the human element. However, AI can and should be used to enhance the practice of medicine, empowering physicians with the latest technological tools to serve our patients better. Moreover, Generative AI may provide physicians with a future that enables them to fully experience the reason why they decided to pursue medicine in the first place—to interact with their patients.

The AMA has addressed the importance of AI, has advocated for the use of the expression augmented intelligence, and has assumed thought leadership with its reports and guidelines for physicians. AMA policy states, "as a leader in American medicine, our AMA has a unique opportunity to ensure that the evolution of AI in medicine benefits patients, physicians, and the health care community."

Relevant AMA Policy

Augmented Intelligence in Health Care H-480.939[66]
Augmented Intelligence in Health Care H-480.940[67]
Augmented Intelligence in Medical Education H-295.857[68]
Professionalism in Health Care Systems E-11.2.1 [69]
Assessing the Potentially Dangerous Intersection Between AI and Misinformation H-480.935[70]

Three AI-related resolutions were introduced for consideration by the House of Delegates at the 2023 AMA Annual Meeting. They were combined into one measure, RES 609-A-23 Encouraging Collaboration Between Physicians and Industry in AI (Augmented Intelligence) Development, urging physicians to educate patients on benefits and risks and directing the AMA to work with the federal government to protect patients from false or misleading AI-generated medical advice. The HOD action was referral. A BOT report is scheduled for consideration by the HOD at the 2024 AMA Annual Meeting.

Specifically, the AMA was directed to:

- Study and develop recommendations on the benefits of and unforeseen consequences to the medical profession of large-language models (LLMs) such as generative pretrained transformers (GPTs) and other augmented intelligence-generated medical advice or content.

- Propose appropriate state and federal regulations with a report back at the 2024 AMA Annual Meeting.

- Work with the federal government and other appropriate organizations to protect patients from false or misleading AI-generated medical advice.

- Encourage physicians to educate patients about the benefits and risks of LLMs including GPTs.

- Support publishing groups and scientific journals to establish guidelines to regulate the use of augmented intelligence in scientific publications that include detailing the use of augmented intelligence in the methods and exclusion of augmented intelligence systems as authors and the responsibility of authors to validate veracity of any text generated by augmented intelligence.

REFERENCES

[1] AMA. "CLRPD Report 1-A-18: A Primer on Artificial and Augmented Intelligence." https://www.ama-assn.org/sites/ama-assn.org/files/corp/media-browser/public/hod/a18-clrpd01.pdf. Accessed August 22, 2023.
[2] McKinsey. "What is ChatGPT, DALL-E, and generative AI?" https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai. Accessed August 22, 2023.
[3] Friedland A. "What are Generative AI, Large Language Models, and Foundation Models?" Center for Security and Emerging Technology. https://cset.georgetown.edu/article/what-are-generative-ai-large-language-models-and-foundation-models/. Accessed August 22, 2023.
[4] Callaway E. "DeepMind's AI Makes Gigantic Leap in Solving Protein Structures." Scientific American. https://www.scientificamerican.com/article/deepminds-ai-makes-gigantic-leap-in-solving-protein-structures/. Accessed August 22, 2023.

[5] Gordon R. "3 Questions: Jacob Andreas on Large Language Models." MIT News | Massachusetts Institute of Technology. https://news.mit.edu/2023/3-questions-jacob-andreas-large-language-models-0511. Published May 11, 2023. Accessed August 22, 2023.
[6] Chen BX. "Get the Best from ChatGPT with These Golden Prompts." The New York Times. https://www.nytimes.com/2023/05/25/technology/ai-chatbot-chatgpt-prompts.html. Published May 25, 2023. Accessed August 22, 2023.
[7] Pifer R. "Hurtling into the Future: The Potential and Thorny Ethics of Generative AI in Healthcare." Healthcare Dive. https://www.healthcaredive.com/news/generative-AI-healthcare-gpt-potential/648104/. Accessed August 22, 2023.
[8] AXIOS. " Terms-of-service land grab: Tech firms seek private data to train AI." Accessed August 24, 2023.
[9] Mello M, Guha N. "ChatGPT and Physicians' Malpractice Risk." JAMA. https://jamanetwork.com/journals/jama-health-forum/fullarticle/2805334? Accessed August 22, 2023.
[10] Bolton E, Hall D, Yasunaga M, et. All. "Stanford CRFM Introduces PubMed GPT 2.7B." Stanford HAI. https://hai.stanford.edu/news/stanford-crfm-introduces-pubmedgpt-27b. Accessed August 22, 2023.
[11] Heaven WD. "Why Meta's latest Large Language Model Survived Only Three Days Online." MIT Technology Review. https://www.technologyreview.com/2022/11/18/1063487/meta-large-language-model-ai-only-survived-three-days-gpt-3-science/. Accessed August 22, 2023.
[12] Field H. "OpenAI is Pursuing a New Way to Fight A.I. "Hallucinations." CNBC. https://www.cnbc.com/2023/05/31/openai-is-pursuing-a-new-way-to-fight-ai-hallucinations.html. Published May 31, 2023. Accessed August 22, 2023.
[13] Hu L. "Generative AI and Future. Medium." https://pub.towardsai.net/generative-ai-and-future-c3b1695876f2. Published November 15, 2022. Accessed August 22, 2023.
[14] "OpenAI. DALL·E 2. OpenAI." https://openai.com/dall-e-2. Published 2023. Accessed August 22, 2023.
[15] Crouse M. "GPT-4 Cheat Sheet: What is GPT-4, and What is it Capable Of?" TechRepublic. https://www.techrepublic.com/article/gpt-4-cheat-sheet/. Published July 7, 2023. Accessed August 22, 2023.
[16] Agomuoh F. "ChatGPT: How to Use the AI Chatbot Faking Over the World." https://www.digitaltrends.com/computing/how-to-use-openai-chatgpt-text-generation-chatbot/. Published January 19, 2023. Accessed August 22, 2023.
[17] "OpenAI GPT-4." https://openai.com/research/gpt-4. Accessed August 22, 2023.
[18] NBC News. "ChatGPT passes MBA exam given by a Wharton Professor". https://www.nbcnews.com/tech/tech-news/chatgpt-passes-mba-exam-wharton-professor-rcna67036. Accessed August 22, 2023.
[19] Digital Trends. "How to Use Google Bard". https://www.digitaltrends.com/computing/how-to-use-google-bard/. Accessed August 22, 2023.
[20] Google Technology. "AI LaMDA". https://blog.google/technology/ai/lamda/. Accessed August 22, 2023.
[21] ZdNet. "Google's ChatGPT Competitor Spaces Out at the Start". https://www.zdnet.com/article/googles-chatgpt-competitor-spaces-out-at-the-start/. Accessed August 22, 2023.
[22] Google Technology. "AI MedPalm Research". https://blog.google/technology/health/ai-llm-medpalm-research-thecheckup/. Accessed August 22, 2023.
[23] Cloud Google. "MedPalm2 Large Language Model". https://cloud.google.com/blog/topics/healthcare-life-sciences/sharing-google-med-palm-2-medical-large-language-model. Accessed August 22, 2023.
[24] Forbes. "Big Techs New Weapon Against Physician Burnout. https://www.forbes.com/sites/davidchou/2023/08/02/big-techs-new-weapon-against-physician-burnout-generative-ai/?sh=79c0241b1f7b. Accessed August 22, 2023.
[25] Zdnet. "Microsoft Taps ChatGPT to Improve Bing and Edge Browser." https://www.zdnet.com/article/microsoft-taps-chatgpt-to-improve-bing-and-the-edge-browser/. Accessed August 22, 2023.
[26] ZdNet. "What is the New Bing. Here's everything You Need to Know." https://www.zdnet.com/article/what-is-the-new-bing-heres-everything-you-need-to-know/. Accessed August 22, 2023.
[27] Digital Trends. "How to Use Microsoft ChatGPT Bing Edge." https://www.digitaltrends.com/computing/how-to-use-microsoft-chatgpt-bing-edge/ Accessed August 22, 2023.
[28] McKinsey. "What is Generative AI?" https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai. Accessed August 22, 2023.

[29] McKinsey. The state of AI in 2023: Generative AI's breakout year. McKinsey. "The state of AI in 2023: Generative AI's breakout year | McKinsey." Accessed August 24, 2023.
[30] McKinsey. The economic potential of generative AI. McKinsey. "The-economic-potential-of-generative-ai-the-next-productivity-frontier-vf.pdf." (mckinsey.com). Accessed August 24, 2023.
[31]AMA. AMA digital health care 2022 study findings. https://www.ama-assn.org/about/research/ama-digital-health-care-2022-study-findings. Accessed August 22, 2023.
[32] AMA. "Digital Machine Learning 101, Promise and Pitfalls." https://www.ama-assn.org/practice-management/digital/machine-learning-101-promise-pitfalls-and-medicine-s-future. Accessed August 22, 2023.
[33] AMA. "10 Ways AI Could Transform Primary Care." https://www.ama-assn.org/practice-management/digital/10-ways-health-care-ai-could-transform-primary-care. Accessed August 22, 2023.
[34] AMA. "Ophthalmologist Doing Health Care AI the Right Way." https://www.ama-assn.org/practice-management/digital/ophthalmologist-doing-health-care-ai-right-way. Accessed August 22, 2023.
[35] AMA. "Health2047's New Company Plans to Use AI Research to Improve Clinical Decision Making." https://health2047.com/media_coverage/health2047s-new-company-plans-to-use-ai-research-access-to-improve-clinical-decision-making/. Accessed August 22, 2023.
[36] Recoverx. "Scaling equitable, science-based healthcare." https://www.recoverx.com/. Accessed August 22, 2023.
[37] Crigger E. Rinebold K. Hanson C. Kao A. Blake D. Irons M. "Trustworthy Augmented Intelligence in Health Care." https://rdcu.be/de3cZ. Accessed August 22, 2023.
[38] AMA. CPT Developer Program. https://platform.ama-assn.org/ama/#/dev-program. Accessed August 22, 2023.
[39] National Academy of Medicine. "Artificial Intelligence for Health Professions Educators." https://nam.edu/artificial-intelligence-for-health-professions-educators/. Accessed August 22, 2023.
[40] JAMA. "Artificial Intelligence in Health Care. A Report From the National Academy of Medicine." https://jamanetwork.com/journals/jama/fullarticle/2757958. Accessed August 22, 2023.
[41] AMA. "7 Tips on Responsible Use of Health Care AI." https://www.ama-assn.org/practice-management/digital/7-tips-responsible-use-health-care-ai. Accessed August 22, 2023.
[42] AMA. "ChatGPT, What Physicians Should Consider." https://www.ama-assn.org/system/files/chatgpt-what-physicians-should-consider.pdf. Accessed August 22, 2023.
[43] Health Sciences, Arizona Edu. "Would You Trust AI Doctor, New Research Shows Patients Are Split." https://healthsciences.arizona.edu/newsroom/news-releases/0523/would-you-trust-ai-doctor-new-research-shows-patients-are-split. Accessed August 22, 2023.
[44] Robertson, C. Woods A. Bergstrand K. Findley J. Balser C. Slepian M. "Diverse patients' attitudes towards Artificial Intelligence (AI) in diagnosis." https://doi.org/10.1371/journal.pdig.0000237. Accessed August 22, 2023.
[45] Fierce Healthcare. "Epic Moves Forward Bringing Generative AI Healthcare. Here's Why a Handful of Health Systems Are." https://www.fiercehealthcare.com/health-tech/epic-moves-forward-bring-generative-ai-healthcare-heres-why-handful-health-systems-are. Accessed August 22, 2023.
[46] UCDavis Edu. "Epic Slicer Dicer." https://health.ucdavis.edu/data/epic-slicer-dicer.html. Accessed August 22, 2023.
[47] Microsoft. "Azure AI Overview." https://learn.microsoft.com/en-us/azure/ai-services/openai/overview. Accessed August 22, 2023.
[48] Microsoft Blog. "Microsoft and Epic expand AI collaboration to accelerate generative AI's impact in healthcare, addressing the industry's most pressing needs." Accessed August 22, 2023.
[49] Science Direct. "Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges, and implications of generative conversational AI for research, practice, and policy." https://www.sciencedirect.com/science/article/pii/S0268401223000233. Accessed August 22, 2023.
[50] Frontier. "AI chatbots not yet ready for clinical use." https://www.frontiersin.org/articles/10.3389/fdgth.2023.1161098/full?utm_source=S. Accessed August 22, 2023.
[51] JAMA. "Garbage in, Garbage out—Words of Caution on Big Data and Machine Learning in Medical Practice." https://jamanetwork.com/journals/jama-health-

forum/fullarticle/2801776?widget=personalizedcontent&previousarticle=2806091. Accessed August 22, 2023.
[52] The Brookings Institute. "Around the Halls, What Should the Regulation of Generative AI Look Like?" https://www.brookings.edu/blog/techtank/2023/06/02/around-the-halls-what-should-the-regulation-of-generative-ai-look-ike/?utm_campaign=Brookings%20Brief&utm_medium=email&utm_content=260949068&utm_source=hs_email. Accessed August 22, 2023.
[53] Girotra K. Meincke L. Terwiesch C. Ulrich K. "Working Paper, Ideas are Dimes a Dozen: Large Language Models for Idea Generation in Innovation." https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4526071. Accessed August 22, 2023.
[54] Al-Medfa MK. Al-Ansari A. Darwish AH. Qreeballa TA. Jahrami H. "Physicians' attitudes and knowledge toward artificial intelligence in medicine: Benefits and drawbacks." https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10073828/. Accessed August 22, 2023.
[55] Harner S. "Attention is not all you need: the complicated case of ethically using large language models in healthcare and medicine." https://www.thelancet.com/journals/ebiom/article/PIIS2352-3964(23)00077-4/fulltext. Accessed August 22, 2023.
[56] Coalition For Health AI. "Coalition for Health AI (CHAI™) Updates Progress and Plans to Issue Guidelines for the Responsible Use of AI in Healthcare." https://www.coalitionforhealthai.org/updates/october-6th-2022. Accessed August 22, 2023.
[57] "MITRE's Brian Anderson, M.D., Explains Coalition's AI Blueprint" https://www.hcinnovationgroup.com/analytics-ai/artifical-intelligence-machine-learning/article/53063115/mitres-brian-anderson-md-explains-coalitions-ai-blueprint. Accessed August 22, 2023.
[58] Digiorgio A. Ehrenfeld J. "Artificial Intelligence in Medicine & ChatGPT: De-Tether the Physician." https://link.springer.com/article/10.1007/s10916-023-01926-3. Accessed August 23, 2023.
[59] AMA. "Feds Warn Algorithms Can Induce Bias in Clinical Decisions." https://www.ama-assn.org/delivering-care/health-equity/feds-warned-algorithms-can-introduce-bias-clinical-decisions. Accessed August 22, 2023.
[60] FDA. "Artificial Intelligence and Machine Learning Software Regulation." https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device#regulation. Accessed August 22, 2023.
[61] AMA. "FDA Must Guard Against Bias AI Focus in Patient Outcomes." https://www.ama-assn.org/practice-management/digital/fda-must-guard-against-bias-ai-focus-patient-outcomes. Accessed August 22, 2023.
[62] Reuters. "Lawmakers Committee Reaches Deal on AI Act." https://www.reuters.com/technology/eu-lawmakers-committee-reaches-deal-artificial-intelligence-act-2023-04-27/. Accessed August 22, 2023.
[63] The AI Act. "What is the AI Act?" https://artificialintelligenceact.eu/. Accessed August 22, 2023.
[64] U.S. Copyright Office Fair Use Index. "2013 Joint Strategic Plan on Intellectual Property Enforcement." https://www.copyright.gov/fair-use/. Accessed August 22, 2023.
[65] Vokinger K. Gasser U. "Regulating AI in medicine in the United States and Europe." "https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7611759/. Accessed August 22, 2023.
[66] AMA PolicyFinder. Augmented Intelligence in Health Care H-480.939. Accessed August 22, 2023.
[67] AMA PolicyFinder. Augmented Intelligence in Health Care H-480.940. Accessed August 22, 2023.
[68] AMA PolicyFinder. Augmented Intelligence in Medical Education H-295.857. Accessed August 22, 2023.
[69] AMA PolicyFinder. Professionalism in Health Care Systems: Augmented Intelligence E-11.2.1. Accessed August 22, 2023.
[70] AMA PolicyFinder. Assessing the Potentially Dangerous Intersection Between AI and Misinformation H-480.935. Accessed August 22, 2023.