Institute *for* Healthcare Improvement

Artificial Intelligence in Health Care

Implications for Patient and Workforce Safety

Innovation Report
ihi.org

Authors

Kate Feske-Kirby, MA, Research Associate, IHI

Kaveh Shojania, MD, Professor and Vice Chair (Quality and Innovation), Department of Medicine, University of Toronto

Patricia McGaffigan, RN, MS, CPPS, Vice President, IHI

This IHI innovation project was conducted from July to October 2023.

IHI's innovation process seeks to research innovative ideas, assess their potential for advancing quality improvement, and bring them to action. The process includes time-bound learning cycles (30, 60, or 90 days) to scan for innovative practices, test theories and new models, and synthesize the findings (in the form of the summary Innovation Report).

Acknowledgments

The authors are thankful to the interviewees who shared their insights, experiences, and expertise to inform this report. IHI is grateful to the Gordon and Betty Moore Foundation for their generous funding that supported this innovation project.

How to Cite This Document: Feske-Kirby K, Shojania K, McGaffigan P. *Artificial Intelligence in Health Care: Implications for Patient and Workforce Safety*. Boston: Institute for Healthcare Improvement; 2024. (Available at ihi.org)

Institute for Healthcare Improvement

For more than 30 years, the Institute for Healthcare Improvement (IHI) has used improvement science to advance and sustain better outcomes in health and health systems across the world. We bring awareness of safety and quality to millions, accelerate learning and the systematic improvement of care, develop solutions to previously intractable challenges, and mobilize health systems, communities, regions, and nations to reduce harm and deaths. We work in collaboration with the growing IHI community to spark bold, inventive ways to improve the health of individuals and populations. We generate optimism, harvest fresh ideas, and support anyone, anywhere who wants to profoundly change health and health care for the better. Learn more at ihi.org.

© 2024 Institute for Healthcare Improvement. All rights reserved. Individuals may photocopy these materials for educational, not-for-profit uses, provided that the contents are not altered in any way and that proper attribution is given to IHI as the source of the content. These materials may not be reproduced for commercial, for-profit use in any form or by any means, or republished under any circumstances, without the written permission of the Institute for Healthcare Improvement.

Contents

Executive Summary

The excitement surrounding use of artificial intelligence (AI) in health care is palpable, especially for generative AI (genAI) tools such as ChatGPT. Yet, the specific potential benefits of such tools in health care, as well as the risks and unintended challenges potentially associated with their implementation, are not well understood.

To better understand the current landscape, the Institute for Healthcare Improvement (IHI) conducted a 90-day innovation project to identify high-level applications of genAI along with their advantages and disadvantages, including unexpected consequences or new safety hazards for patients and the health care workforce. The innovation project reviewed three primary types of clinical applications in health care: documentation support, clinical decision support, and chatbots that provide patient support.

While genAI applications in health care hold much promise, they also present serious potential risks that must be considered and addressed prior to more widespread implementation.

This report describes:

- A detailed review of three primary applications of genAI in health care: documentation support, clinical decision support, and chatbots that provide patient support;
- Analysis and review of the three genAI applications — including functions, benefits, risks, and examples — relevant to inpatient, primary care, and ambulatory care settings;
- Potential patient safety risks and unintended consequences resulting from genAI tools; and
- Considerations and recommendations for ongoing work to finetune AI tools and mitigate potential risks.

Intent and Aim

In May 2023, the Institute for Healthcare Improvement (IHI) Lucian Leape Institute identified artificial intelligence (AI), specifically generative AI (genAI), and safety considerations for applications in health care as an area of interest to further investigate. The intent of this IHI 90-day innovation project (conducted from July to October 2023) was to identify the high-level applications of generative AI in health care and assess implications for safety, including unexpected consequences or new safety hazards for patients and the health care workforce. Generative AI includes machine learning systems capable of generating text, images, code, or other types of content, often in response to a prompt or question entered by a user through a chat interface.

With the proliferation of AI and generative AI applications in health care, it is necessary to examine how to appropriately integrate these technologies into clinical workflows and health care operations with careful attention to the impact on patient care and safety, workforce safety and well-being, and the health care sector as a whole. There are also considerations for the rapid adoption of this technology, which may outpace the education and agility of the workforce as well as existing technological infrastructure.

The 90-day innovation project included these activities:

- Literature scan to analyze current and future uses of AI, particularly generative AI, in health care: This included literature on generalized perspectives of the use and risk of AI in health care and publications specific to various health care settings and specialties.
- Interviews with 27 individuals from selected health care and other organizations to gain further insights: This included academic experts, practitioners, and technical experts with relevant expertise in the topic area.

The innovation project reviewed three primary types of clinical applications in health care (see Table 1) and focused on generative AI tools already available for implementation and those likely to find application in clinical care within the next year.

Table 1. Generative AI Applications in Health Care

GenAI Application	Description
Documentation support	Functions include producing formatted visit or encounter notes from transcripts of patient-clinician interactions. Other examples include generating discharge summaries from the notes and results in the electronic health record (EHR), summarizing patient information (e.g., for handovers), and resolving inaccuracies and redundancies in the EHR.
Clinical decision support	Functions include providing diagnostic support by providing potential diagnosis based on symptoms, test and lab results, and patient history; providing early detection or warning on patient condition; and developing potential treatment plans.

Chatbots	Functions include personalized patient communication and navigation services, particularly supporting clinicians in responding to messages from patients in EHR inboxes.
Additional non-clinical uses (see Appendix B)	Includes administrative assistance (e.g., medical billing and coding), as well as support for medical research and education, and quality improvement (QI) and safety efforts (e.g., preparing case summaries for root cause analyses).

The aim of this innovation project is to inform health care personnel about the specific functions genAI tools in health care can accomplish — now or in the immediate future — while also identifying potential risks and mitigation strategies.

Background

With the emergence of publicly available AI chatbots such as OpenAI's ChatGPT (GPT-3.5, GPT-4) and Google's Bard, AI applications in health care are rapidly growing. While not yet widely applied in practice, published reports have already shown that generative AI tools can pass medical licensing examinations, generate clinical notes, and respond empathetically and accurately to medical questions from patients.[1,2,3,4] These and other applications, such as point-of-care decision support for diagnosis and treatment, have understandably generated widespread excitement.

Yet, this growing interest in implementing genAI tools in health care coincides with calls from prominent AI experts to pause the development and implementation of AI applications for varied concerns, including accuracy, the perpetuation of systemic biases, spreading disinformation, privacy and security threats, and fears about the negative implications for the human workforce and the existential threat of AI developing autonomously.

One concrete example of concerns over accuracy consists of so-called "hallucinations" or "confabulations" — that is, false or made-up statements presented as factual (see Table 2 and Appendix A for terminology and definitions). Non-medical examples of such hallucinations have appeared in the news, such as the example of a legal firm that submitted a brief containing non-existent legal precedents.[5] Examples of chatbots making up medical citations also exist.[6] In one article about using GPT-4, an AI-generated medical note presents the patient's Body Medicine Index (BMI) as a specific value despite the lack of supporting information from the original clinical encounter note.[7] As the article authors noted, "such errors can be particularly dangerous in medical scenarios because the errors or falsehoods can be subtle and are often stated by the chatbot in such a convincing manner that the person making the query may be convinced of its veracity."[8]

In addition to the issue of hallucinations and confabulations, there remain many problems involving the perpetuation of social biases and inequities that are often baked into the datasets used by genAI applications.[9,10,11,12,13] For instance, AI tools used to diagnose skin conditions may have been trained on datasets without sufficient representation of the variation in common

skin diseases (never mind rare ones) across different skin types, especially people with darker skin tones.[14] Insufficient representation in training datasets can also lead to AI tools missing the presence of disease in chest radiographs of patients from racialized or underserved groups.[15]

Artificial intelligence is not new to health care. Early expert systems for medical diagnosis emerged in the 1960s followed by the development of early applications for image recognition and diagnosis in the 1970s and the use of natural language processing for documentation in the 1990s.[16,17] More recent years (circa 2016) saw machine learning programs akin to those used to beat world champions in chess and demonstrate expert-level performance in interpreting medical images in radiology, pathology, and dermatology.[18,19,20,21]

These algorithms and models differ, though, from contemporary genAI in that they were typically trained on large datasets containing "gold standard judgments" by human experts (e.g., tens of thousands of moles clearly identified by majority consensus among dermatologists as either benign or malignant). These models also tend to have known performance in terms of the accuracy of their specific judgments (e.g., if a skin lesion not from the training dataset is benign or malignant).

GenAI tools have the advantage of not being so focused — they can make judgments or provide answers in response to questions or prompts for which they were not specifically trained. On the other hand, they were trained on datasets with no "gold standard" answers. For instance, there may be numerous images and text materials related to moles, but no human experts have verified the veracity of the statements contained in these materials. The performance of the AI tools in terms of generating accurate answers has thus far been assessed mostly at a general level (e.g., not in terms of arriving at the correct diagnosis or treatment recommendation for specific clinical situations).

Table 2. Terminology and Definitions

Term	Definition
Artificial intelligence (AI)	The branch of computer science that aims to create systems capable of performing tasks that typically require human intelligence.[22] These tasks can include problem-solving, understanding natural language, recognizing patterns, making decisions, and adapting to new information. AI systems can be based on rules and algorithms or can be trained using data-driven approaches like machine learning, where they improve over time by learning from data.
Big data	An indetermined phrase that often refers to a large, diverse, and complex set of data from various sources. Authors of one article note that there is no agreed on standard definition, although there are important overlaps recognized in the definition above.[23]
Generative artificial intelligence (genAI)	Broadly speaking, generative AI refers to machine learning systems capable of generating text, images, code, or other types of content, often in response to a prompt or question entered by a user through a chat interface. This differs from discriminative AI models that perform

	functions such as data classification or data grouping, or action oriented.[24] GenAI models are initially trained on broad datasets (e.g., massive amounts of text from the Internet), then the model is finetuned, often with a more specific or specialized dataset (e.g., text relevant to health care), in which human reviewers refine the model's behavior and ensure that it aligns with intended use cases and ethical guidelines.[25] Generative artificial intelligence systems include large language models as well as image generators, code generator tools, and auto generation.[26] Prominent commercial examples of gen AI models include OpenAI ChatGPT, Google Bard, and Llama 2 created by Meta.
Generative pretrained transformer (GPT)	AI that can produce human-like writing such as ChatGPT.[27]
Hallucinations or confabulations	False or made-up statements presented as factual.
Large language models (LLMs)	A type of AI that engages with language. These models are referred to as "large" due to the increase in parameters for training language models, including increased data and computational power.[28] Large language models may also be called foundational models.
Machine learning (ML)	Machine learning is a subfield of artificial intelligence that uses algorithms trained on datasets to create models that enable machines to perform tasks that would otherwise only be possible for humans.[29]
Natural language processing (NLP)	A type of speech recognition AI focused on producing human-like language processes and outputs using text and spoken word through a computer.[30,31] Functions range from simple analytics tasks like classifying documents to advanced tasks like responding to questions and summarizing literature.
Neural network	Computing systems with interconnected nodes that can identify hidden patterns and correlations in raw data. These patterns and correlations are clustered and classified and, over time, the system can continue to learn and improve.[32] Multilayered neural networks constitute the "deep learning" programs used for medical image interpretation, among other applications.
Transformer model architecture	A neural network algorithm that learns contextual information from sequential data (e.g., wearables, videos, textual works in written or spoken language, or other audio signals).[33] Transformer model architecture is the basis for large learning models.[34]

Generative AI in Health Care

The excitement over AI in health care reflects its potential to achieve several distinct goals. AI-powered decision support tools will almost certainly increase the proportion of patients receiving accurate diagnoses and evidence-based treatments. These tools may also foster earlier recognition of deteriorating patients and play a role in reducing common, preventable adverse events as well as address other patient safety and quality of care issues. Existing genAI tools can facilitate clinical documentation and administrative tasks that contribute to clinician burnout. There also are genAI tools that can produce patient-facing versions of complex medical notes, answer medical questions for patients, and provide substantial assistance with care coordination and patient navigation.

With many potential genAI applications in health care — which vary based on technologic simplicity or complexity and readiness for adoption[35] — the field must proceed with caution since most AI tools have not undergone evaluation or validation to assess their accuracy or the extent of adverse consequences. Notable concerns include the problem of authoritative-looking pronouncements containing factually incorrect material (i.e., hallucinations, confabulations), misleading or inappropriate outputs due to biased training sets, and various ethical, legal, and regulatory concerns (e.g., the need for FDA/regulatory oversight of AI tools that amount to medical devices), among others.[36,37,38]

The combination of massive potential benefits accompanied by numerous serious unintended consequences place health care workers, executives, and consumers in a bind. Numerous genAI tools offer promise for making clinicians' work much easier (e.g., instantly generating medical notes from patient-provider interactions) and value for patients (e.g., instantly generating patient-friendly versions of medical notes and enabling clinicians to provide timely and better care).

On the other hand, implementing these tools could lead to massive, irreversible new problems. As a corollary, electronic health records (EHRs) were also supposed to make clinicians' lives easier, when many cite the EHR "solution" makes the work more complex and inefficient.[39] Yet, one of the main arguments in favor of AI tools is that they will help address the problem of clinician burnout widely attributed to EHRs.

Results of the 90-Day Innovation Project

The landscape of AI in health care is highly dynamic, rapidly expanding and changing, including the state of technology developments, the implications for our knowledge of how well AI tools perform, and related safety implications. Consequently, we caution readers that this report may be dated by the time of publication or shortly thereafter. Additionally, the findings from the literature review and interviews during this innovation project reveal a lack of expert consensus on application readiness, notable risks or challenges, and conflicting viewpoints.

Even with the enthusiasm surrounding AI in health care, the promise must first be examined. What applications are ready to go? How, where, and with whom should they be designed and implemented? What are the benefits, risks, and implications of using specific AI applications?

The findings in this report primarily focus on applications and uses of AI, particularly generative AI, for clinical care delivery systems. This report addresses three primary genAI clinical applications: documentation support, clinical decision support, and chatbots that provide patient support. Analysis and review of each AI application includes functions, benefits, risks, and real-world examples and are relevant to multiple care settings, including inpatient, primary care, and ambulatory care.

The report provides real-world examples of AI applications in development or implementation. Any references to information and technology companies and health care system partnerships are not an endorsement of said companies or partnerships; it was beyond the scope of the innovation project to validate or test related applications.

Please see Appendix A for a description of risks, challenges, and unintended consequences of clinical applications of genAI, and Appendix B for details on non-clinical uses of genAI technology in health care.

Review of Three Clinical Applications of GenAI

1. Documentation Support

- Produce Patient Interaction Notes
- Produce Patient Medical History Summaries
- Resolve Patient Record Inaccuracies and Redundancies
- Provide Accessible Patient-Facing Documentation

2. Clinical Decision Support

- Suggest Potential Diagnoses
- Provide Early Detection or Warning
- Propose Treatment Plans

3. Chatbots That Provide Patient Support

- Patient Communication
- Triage Patient Check-In and Information

Review of Three Clinical Applications of GenAI

This section discusses three clinical applications of genAI: documentation support, clinical decision support, and chatbots that provide patient support. Table 3 at the end of this section provides a summary review of these applications and potential risks.

1. Documentation Support

A key application of genAI is documentation support. Currently, clinical documentation is typically entered into electronic health records (EHRs) by clinicians or clinical support staff such as scribes. Yet, this manual effort can be burdensome, time-intensive, and is frequently cited as a major contributor to clinician burnout.[40,41,42] The use of genAI for documentation support has multiple functions and can be used to support both patients and clinicians.

Clinicians can utilize genAI applications for a broad array of documentation support purposes, including to:

- Record visit or encounter notes during patient-clinician interactions;
- Summarize patient information including demographics, medical history, and medical records; and
- Resolve inaccuracies and redundancies in a patient's electronic health record.

Produce Patient Interaction Notes

Using ambient listening, genAI can produce detailed and accurate notes of patient interactions, including scheduled care visits and bedside visits.[43] Early adopters have begun using AI technology in multiple settings, including primary care, ambulatory care, and the emergency department.[44,45,46] While genAI can produce the note, clinical oversight is needed to both review and edit the note prior to completion. In fact, one expert from a large health system shared that their organization had deliberately designed their AI-powered scribe tool to generate notes with features indicating that notes are only approximately 90 percent complete, so that clinicians must review and provide final input to publish a completed note.

Using AI-powered scribe tools during clinical encounters involving multiple speakers (e.g., patient, family members, and clinician or a patient and several clinicians) is still a work in progress. Ambient noise associated with inpatient encounters (e.g., in the emergency department or an intensive care unit) may also present problems for AI-driven scribe tools, although improvements to mitigate such issues are actively being addressed by some tools.

Produce Patient Medical History Summaries

GenAI can provide clinicians with succinct summaries of a patient's medical history, including reasons for patient visits, previous clinical encounter notes (including primary care or specialist visits), charting for inpatient patients, and previous labs or imaging. Patient summaries have multiple uses such as reducing clinical time spent creating and reviewing patients notes or reviewing patient history during a visit, improving communication for clinical handoffs or referrals, and ensuring continuity in clinical care delivery.

Resolve Patient Record Inaccuracies and Redundancies

GenAI can be used to resolve inaccuracies and redundancies in a patient's electronic health record (EHR). While not discussed as thoroughly in the literature, interviewed experts suggested that genAI technology could be applied to patient EHRs to improve the quality of notes by removing inaccuracies and reducing redundant notes. While less studied than other genAI applications, ensuring that electronic clinical documentation is accurate and accessible not only improves record keeping, but also provides higher quality data for clinical encounters and for use in de-identified datasets.

Provide Accessible Patient-Facing Documentation

GenAI can produce patient-facing documentation — accessible material provided to the patient that details specific health care-related information. GenAI applications could efficiently create visit notes, discharge summaries, and treatment plans in simple language (or at the reading proficiency level of the patient and/or approved caregiver), removing clinical or medical jargon without decreasing the accuracy of information. Additionally, it's possible for genAI tools to provide language translations of patient notes, records, and resources to advance equitable care for patients who speak a different primary language.

Potential Benefits and Risks

The use of genAI to support clinical documentation has notable benefits:

- Reduce documentation burden for clinicians;
- Increase clinical time available for direct patient care; and
- Improve accessibility for patients, given that AI tools can produce patient-friendly versions of medical notes in multiple languages.

What is not yet known is how often genAI tools produce inaccurate statements within such notes. Some experts indicated that, apart from input errors (e.g., incorrectly transcribed words due to ambient noise or unfamiliar accents), AI-generated content that is factually incorrect (i.e., hallucinations or confabulations) is rare. Other interviewed experts highlighted that hallucinations can happen whenever a genAI tool analyzes or synthesizes information, which can also occur with producing formatted clinical notes and not just transcriptions of the verbal interaction.[47] If AI programs produce inaccurate outputs, that documentation's accuracy could be further diluted due to recall bias by the provider(s).

Additional risks that may be encountered with genAI applications for documentation support are described below.

- **Deskilling:** Clinicians, especially those early in their career, may lose the ability to synthesize and concisely summarize the key information obtained during exchanges with patients.
- **Patient privacy and data rights:** Chatbots have already produced privacy scandals.[48] In March 2023, a software glitch allowed ChatGPT users to see queries posed by others and in some cases could see their credit card information.[49] The authors of one article assert, with the limited privacy and data protection laws in the United States and

technological advancements such as the EHR, "HIPAA's Privacy Rule is based on a fallacy: the belief that data can be successfully stripped of personal information (de-identified) and that doing so renders it safe. In the age of LLMs, data can easily be re-identified, and even de-identified data can cause significant harm."[50]

- **Patient preference:** Some patients may perceive use of AI technology as an intrusion and threat to privacy, though others may welcome clinicians not having to focus so much on computer screens during clinical encounters.
- **Application errors:** Hallucinations and confabulations in AI-produced content have already been mentioned, and there may be other AI-generated errors related to the quality of the ambient audio in different and difficult clinical settings with multiple participants and/or noise.

Examples

Despite these concerns, medical documentation facilitated by generative AI tools is already proceeding, as the following examples illustrate.

- As of 2023 Houston Methodist is applying Pieces technology to provide a "working summary" for clinical handovers to improve patient care quality, reduce documentation burden, and improve clinical communication.[51] The working summary in the acute setting includes patient demographics, patient condition (current and recent), relevant patient history, recent procedures, predictive discharge status, and contextual risk assessment and communication.
- In 2023, HCA Healthcare partnered with Augmedix to accelerate the development of AI-enabled ambient documentation solutions for acute care providers.[52] Clinicians can use the Augmedix app to create hands-free, precise, and timely medical notes from clinician-patient interactions. The proprietary platform leverages NLP (in addition to Google Cloud's genAI technology, which HCA Healthcare is also partnered with, and multi-party medical speech-to-text processing) to promptly convert the collected data into medical notes, which clinicians review and finalize prior to transfer into the EHR.[53]
- As part of their Partners and Pals program, Epic, the developer and vendor of electronic health records, has named their first Pal, Abridge, a leading company in genAI for clinical documentation.[54] Through this collaboration, Epic and Abridge plan to help providers focus on patients and reduce time spent on documentation, help patients better understand their conversations with providers, and help health care systems rapidly adopt genAI-based solutions. Emory Health will be leveraging this partnership, announcing an enterprise-wide agreement to make Abridge's Epic-integrated genAI solution available to clinicians over the next three years.[55] Emory will join the University of Pittsburgh Medical Center (UPMC) and the University of Kansas Health System in using Abridge through Epic EHR.[56]

2. Clinical Decision Support

The use of clinical decisions support (CDS) in health care is not novel, but genAI could advance CDS tools. Traditional CDS consisted of software designed for a specific purpose to directly aid clinical decision-making. This traditional rule-based application has changed over time with the advancement of technology, with CDS increasingly developing the capacity to leverage data and observations independently (e.g., AI and genAI solutions).[57]

As CDS aims to inform clinicians' decisions about patient care to improve patient outcomes, thereby leading to higher-quality health care, AI applications can support clinical decision-making through the following functions:

- Provide diagnostic support by suggesting potential diagnoses to explain the constellation of symptoms and test results for a given patient, as well as diagnostic support for interpreting medical images;
- Provide early detection or warning on changes to patient condition, including risk assessment and suggested interventions; and
- Develop potential treatment plans, including suggested medications and tests.

Suggest Potential Diagnoses

A genAI application could suggest diagnoses to consider based on presenting symptoms and clinical assessment. To have a sense of how well genAI technology performs in this regard, one research group evaluated ChatGPT's diagnostic performance on 36 clinical vignettes. The researchers characterized its performance as on par with the level of an intern or resident, with 77 percent accuracy on final diagnosis, 68 percent accuracy on clinical management decisions, and 60 percent accuracy on differential diagnoses.[58]

In another example, researchers assessed the performance of Chat GPT-4 in suggesting the correct diagnosis for cases presented as clinicopathologic conferences in the *New England Journal of Medicine.* The genAI model's leading diagnosis agreed with the final diagnosis in 39 percent of cases; in 64 percent of cases, the model listed the final diagnosis in its so-called differential diagnosis.[59]

These results attest to the impressive ability of genAI tools to digest disparate elements of a complex medical case and suggest the correct diagnosis or differential diagnoses for the clinician to consider. Yet, they also highlight the potential problem of automation bias, which raises the risk of diagnostic error and clinician complacency. Clinicians must not defer, consciously or subconsciously, to the AI tool's suggested diagnosis without considering alternatives — at least not with the current approximately 25 to 35 percent inaccuracy of these suggestions, based on available evidence. While many experts expressed enthusiasm for AI-powered decision support, most acknowledged that accuracy needs to improve — as well as establishing reliable performance in a robust manner — to justify widespread implementation.

Provide Early Detection or Warning

While more simple rules-based decision support has been used for clinical decision support for many years, genAI could further support the refinement and expansion of early detection and

warning systems. GenAI can support predictive analytics for earlier detection of deteriorating patient conditions and support predicting patient risk for contracting certain conditions or predicting patient likelihood of responding to treatment.[60,61] As noted in the "Documentation Support" section above, early detection and warnings can also be included in clinical documentation (see Houston Methodist example).

Propose Treatment Plans

GenAI could act as a CDS tool by proposing a suggested treatment plan that includes potential medication, appointment follow-ups, referrals, a care plan, and patient education. These recommendations would be based on a patient's medical history, genetic information, and lifestyle and behavioral factors.[62] Clinicians would be required to review and alter the treatment plan as needed to meet the patient's needs and preferences. As with diagnostic decision support, though, the performance characteristics of such AI-generated recommendations have yet to be established as suitably accurate for clinical application.

Potential Benefits and Risks

To be clear, genAI-powered CDS tools aimed at specific conditions will undoubtedly undergo evaluation and be shown to perform well (e.g., see Intermountain Health example below). Yet, decision support focused on specific conditions differs from the more general decision support envisioned for the years to come, in which AI-driven programs will make diagnostic and therapeutic suggestions related to whatever condition seems suggested by the content of the clinical note the user has opened. For example, "It looks like your patient has 'X' illness. You should consider initiating the following medications…." — where "X" might include common conditions such as diabetes or pneumonia, but also much less common diagnoses.

One key risk associated with AI clinical decision support relates to biased or incomplete training datasets such that AI model outputs are less accurate for patients from different racial/ethnic backgrounds. For instance, one of the experts we interviewed mentioned asking an AI chatbot the most likely diagnosis for a man in his 20s with sore throat and swollen lymph nodes in the neck. For a white patient, the suggested diagnosis was infectious mononucleosis. When the race was changed to Black, gonococcal pharyngitis (i.e., a sexually transmitted disease) was offered as a diagnosis to consider even though it had not been mentioned when a patient was labeled as white. As one expert noted, the diagnostic suggestions provided by AI tools reflect probabilities based on what has been said online ("the epidemiology of the Internet") and not on the actual epidemiology of disease for relevant patient populations.

One journal article provides another example of racially biased recommendations from a genAI tool. The authors asked ChatGPT about the most appropriate analgesia for a 50-year-old man presenting to the emergency department with chest pain. When the man's race was white, ChatGPT recommended "a strong opioid, such as morphine or fentanyl." When the race was changed to Black, ChatGPT stated that the appropriate choice depended on the cause of the chest pain (a qualification not made for the white patient) and then went on to recommend a non-steroidal anti-inflammatory such as ibuprofen or naproxen. As previously mentioned, biased datasets can also affect the accuracy of AI-based diagnosis of radiographs and skin lesions.[63]

It is also worth noting that decision support typically takes the form of pop-up alerts and reminders, and these have generally had as many failures as successes, with the average effect consisting of small improvements in the percentages of patients receiving recommended care.[64] There is also the issue of alert fatigue due to alerts too often firing inappropriately. Some older rules-based decision support tools have often had false positive rates of 90 percent or higher.[65,66] In another study, an AI-driven pneumonia decision support tool had a lower false positive rate of 50 percent.[67] But a false alarm rate of 50 percent may still contribute to patient harm and burnout when clinicians are seeing multiple other forms of decision support, each with their own false alarm rates.

Finally, even if there are reasonable grounds for excitement over high-performing AI-driven decision support tools becoming available soon, fraudulent overpromising has also occurred. For instance, Babylon Health, a UK start-up claiming to put AI tools to work to "put an accessible and affordable health service in the hands of every person on earth," went from rapid expansion and being lauded in 2019 to shuttering its US headquarters in 2023.[68,69]

Example

- Intermountain Health is rolling out the ePneumonia app to all hospitals and urgent cares in seven states as announced in spring 2023.[70] Using an AI Bayesian network, the application helps clinicians determine the best course of treatment for patients and has showed improved outcomes for patients with pneumonia in a randomized controlled clinical trial in 16 community hospitals.[71]

3. Chatbots That Provide Patient Support

A popular, mainstream application of genAI, chatbots could also play a role in supporting communication with patients and managing clinical intake. While all genAI applications must be compliant with regulations such as the Health Insurance Portability and Accountability Act (HIPAA), functions could include:

- Personalized patient communication; and
- Triaging patient check-in information for in-person clinical care settings.

Patient Communication

Personalized patient communication via AI tools can occur both online and in person. Many experts anticipate using AI tools to respond to patient inquiries through online patient portals, provide patients with answers to basic medical questions, or support management plans, particularly for patients with chronic illnesses.[72,73] In-person AI applications for communication might include virtual chatbots available to patients that answer basic questions, set reminders, and facilitate communication with the care team.[74]

Similar to non-automated messages, clinicians will need to review drafts crafted by chatbots (then decide to modify or leave the draft as is, or create a unique response), but this AI application could still reduce the time spent on responses, particularly if the draft includes accurate and appropriate medical information. Moreover, like patient-facing documentation,

patient-oriented chatbots could also communicate in simple language (or at the patient's reading proficiency level) as well as provide language translation for patients.

Triage Patient Check-In and Information

While chatbots may already assist prior to an appointment by supporting timely check-in in ambulatory care, urgent care, or emergency department settings, there is potential for genAI to triage patient information prior to a patient-clinician interaction.[75] Chatbot-supported check-in could collect basic information on the patient, including reason for visit and patient history, and synthesize this information along with a patient's existing medical history to provide a triaged assessment of the patient for the provider to review prior to the encounter.

Potential Benefits and Risks

The use of chatbots for patient communication and triaging could reduce clinical workforce burden, especially EHR inboxes overflowing with patient messages awaiting responses. The use of AI support tools can also lead to quicker response times for simple patient questions and more personalized patient-clinician communication, which may improve the patient care experience.

One study has shown that chatbots can answer questions from patients with comparable quality to answers generated by physicians — and, perhaps because the chatbot does not fatigue or uses standardized language, its responses were judged by patients as more empathic than clinicians.[76] Yet, these AI applications will also contain some hallucinations and other errors, so health care organizations will need to consider strategies to audit the accuracy of applications. Even when implementation includes a step for clinicians to edit AI-generated responses to patients, there will be a natural tendency for many clinicians to omit this step — a concern raised by many interviewed experts, which is also relevant for AI-generated medical notes. Organizations may need to engage a cadre of clinicians who volunteer to critically review chatbot messages to patients to identify the true frequency of incorrect or misleading answers.

Further, messages must be edited by clinicians prior to dissemination to confirm accuracy and credibility and ensure that the response is unique to the patient's concerns, questions, conditions, and preference, including accessibility (e.g., use of simple language, provided in the primary language of the patient). Poor communication through the chatbot could negatively affect patient-provider relationships such as the patient or provider mistrusting one another or the technology. This is why the chatbot must be tested and validated by adequately skilled users who verify the accuracy of the AI-generated information. If automatic messages are utilized, health care systems may need to consider using disclaimers for liability purposes.

Example

- In the summer of 2023, UNC Health began piloting Epic's AI tool, Inbasket, which drafts clinician responses to patient's questions or comments and aims to reduce administrative burdens for staff.[77] Through this pilot test, UNC Health aimed to keep clinicians in the loop with new AI tools and hoped to offer a friendly introduction to AI tools for patients and clinicians with the reminder that AI is there to help and not replace clinicians. Additionally, while not a patient-facing application, in June 2023, UNC Health

announced the rollout of an internal chatbot tool for a small group of clinicians and administrators, with plans to offer it more broadly at a later date.[78]

Table 3. Generative AI Clinical Applications and Potential Risks

Note that not all potential risks and challenges are accounted for in the table.

GenAI Application	Description	Potential Risks and Challenges
Documentation support and scribe function	• Record visit or encounter notes during patient-clinician interactions • Summarize patient information including demographics, medical history, and medical records • Resolve inaccuracies and redundancies in a patient's electronic health record	• Errors caused by ambient noise or multiple speakers • Factually incorrect information ("hallucinations") • Need for local calibration or performance testing • Clinical deskilling (e.g., automation bias, complacency) • Patient privacy and data rights (e.g., data retention policy)
Clinical decision support	• Provide diagnostic support in identifying potential diagnoses and reducing diagnostic error, including improvements in imaging analysis • Deliver early detection on patient condition, including deterioration, and risk assessment • Serve as a virtual assistant to clinicians during treatment • Develop potential treatment plans, including suggested medications and tests	• Unknown performance for most current AI applications • Clinical overreliance and deskilling • Biased datasets leading to inappropriate output for patients from racialized or underserved groups • Worsening alert fatigue
Chatbots that provide patient support	• Provide personalized patient communication • Triage patient check-in and synthesize check-in information for in-person clinical care settings	• Inaccurate data or responses • Patient preference to interact with humans and/or avoid use of these tools • Lack of accessibility for some patients due to lack of facility with using chatbots • Falsehood mimicry and need for adequately skilled user • Poor performance could lead to mistrust of technology or providers

Future Considerations

With heightened excitement, expectations, and rapid adoption of genAI tools in health care, how do we further emphasize caution and safety in this ever-growing space to ensure that patient and workforce safety and other considerations remain at the forefront of consideration, and are not overshadowed?

Begin with Governance

The experts that IHI interviewed emphasized that health care organizations need to begin by building and fortifying governance structures and policies pertaining to genAI. This could include either creating a new governance or oversight body or by tasking a related, existing body with AI governance.

Regardless of where the oversight structure is located within a health care system, it is important to begin efforts by designating a system leader with responsibility for AI activities and ensuring that any oversight body includes a cross-functional team with staff from various roles and responsibilities.[79] This can include executive leaders; clinical representatives; end users (including physicians, nurses, pharmacists, safety/quality/risk leaders, patient and family advisors, administrative or operational staff), IT staff, and members with AI expertise (which could be an external partner if the health system lacks the appropriate internal expertise).

Following the development of a governance body, organizations need to take these steps:

- Develop a clear organizational approach for AI integration that aligns with existing system goals and objectives;
- Improve AI literacy for governance members and staff, including foundational education on AI, applications in health care,

Begin with Governance

- Establish a clear and intentional organizational strategy to support a coordinated approach to AI implementation.
- Create guardrails to ensure the appropriate, effective use of AI applications.
- Conduct a critical assessment of system needs, readiness, and capacity.

Proceed with Caution

- Ensure that genAI models are validated, regulated, and monitored for quality, safety, and ethics.
- Remain up to date on regulations and on collaborative efforts to standardize and develop guidelines and guardrails.

Start Simple and Meet End Users' Needs

- Ensure that AI applications are in the best interests of ensuring patient safety.
- Communicate and engage staff in organizational genAI efforts.

design and implementation risks, and relevant regulatory guidance;
- Develop decision, testing, workflow, and support criteria to inform selection, vetting, value, and ongoing support of AI solutions;
- Create guardrails with safety management systems that include measurement and monitoring plans to ensure that AI is used safely, effectively, and equitably within the organization;
- Ensure that the benefits and risks of AI solutions for specific use cases are clearly defined and addressed prior to use; and
- Review existing operations and infrastructure to determine the health system's current capacity and identify potential improvements and investments needed to integrate AI.

Establish a clear and intentional organizational strategy to support a coordinated approach to AI considerations and potential implementation.

This strategy needs to align with a health care system's existing goals and objectives to ensure continuity in organizational vision. Through this process, systems can further establish priorities and processes prior to AI design and implementation to ensure a systematic and transparent approach to AI integration into health care delivery. Developing an informed and cohesive strategy, though, can be challenging in a rapidly changing technology and regulatory landscape, so it is important that governance members, as well as staff and end users, are educated and gain proficiency in AI in two aspects: 1) gain foundational knowledge on how AI is developed and how it functions, and 2) understand the applications and risks of the technology in a health care setting.[80] Organizations can provide learning opportunities through training sessions or tutorials, brainstorming sessions, and hands-on experiences in a controlled setting (e.g., application sandboxes or simulations). Additionally, an organization's workforce must be apprised of the organizational approach to AI, once identified, and involved in the co-design and implementation of genAI technology.

Create guardrails to ensure the appropriate and effective use of AI applications.

Develop rigorous evaluation methods and assessment frameworks for AI applications, audit tools for applications and datasets, as well as workforce feedback mechanisms such as user groups.[81] Safety guidelines must be detailed, including warnings on AI application risks, reminders of patient privacy (e.g., not sharing patient information on open-source AI), and emphasizing that AI tools must be used in accordance with accepted standards of care and conform with regulatory standards and guidance; following the accepted standard of care can reduce the likelihood of patient harm and liability for clinicians and health care delivery organizations.[82]

Conduct a critical assessment of system needs, readiness, and capacity.

Health care systems seeking to integrate AI must first assess their needs.

- Ask clinicians and staff what problems they need solved to guarantee that any new technology is relevant (e.g., asking, “What is important to you for the organization to improve?” or “What is a major pain point in your daily workflow?”).
- Ensure outreach to a diverse group of employees with various roles and responsibilities for input.
- Reflect on whether identified issues or solutions align with the organizational vision or aims. Note that, if a need identified by staff does not align with an organization’s aims or vision, this does not mean the need should be disregarded; this could serve as an opportunity for the organization to assess whether or not the aims or vision are also meeting workforce needs.

Assessing for system capacity specifically refers to a health care system’s IT infrastructure. This assessment should aim to answer questions such as:

- What is the current capacity of a health care system IT and data infrastructures?
- What types of AI implementations can the infrastructure support in its current state?
- Is the necessary expertise and talent onboard or able to be resourced to support system and end user needs?
- Have data fidelity and accuracy needs been assessed to determine whether it is fit for use with AI tools?
- Do we understand data risks and do we have the necessary infrastructure to ensure protection of data?
- What improvements and investments are needed to ensure the system can function with new AI applications, if any?

These suggested steps in developing a governance framework for AI applications are not linear or comprehensive. Interested organizations are encouraged to review other sources — such as Duke Health AI’s Algorithm-Based Clinical Decision Support Oversight[83] or a proposed governance model for AI in health care[84] — and establish learning system practices across their organization and with other organizations.

Proceed with Caution

It is critical to remember that many genAI models are not ready for full-scale deployment in a clinical care delivery setting. As our research and interviews reveal, and as the World Health Organization acknowledged,[85] while many are enthusiastic about the use of these technologies, there are varying opinions and concerns about the level of caution that is being exercised. Risk identification and mitigation is key to using AI-based applications in health care and, with a rapidly changing technological landscape, interested parties need to demonstrate restraint and prioritize an ongoing commitment to patient and workforce safety and quality.[86]

Health care systems need to exercise care when considering and adopting genAI in clinical care settings (as well as non-clinical settings), with attention to research and testing to ensure that genAI models are validated, regulated, and monitored for quality, safety, and ethics. This includes remaining up to date on collaborative efforts to standardize and develop guidelines and guardrails to promote the fair use of AI systems in health care, such as the National Academy of Medicine's Artificial Intelligence Code of Conduct[87] and the Coalition for Health AI (CHAI).[88]

Organizations must also remain current on pending and authorized regulations at multiple levels. For example, the US Food and Drug Administration released draft guidance tailored to AI- and ML-enabled devices as they further develop a regulatory approach to protect patients and increase the effectiveness of devices.[89] Using caution and staying aware of the regulatory landscape can ensure that patients and the health care workforce do not bear the burden of poorly executed design, implementation, or integration of AI applications into health care.

Start Simple and Meet End Users' Needs

For health care organizations interested in developing or implementing genAI solutions, it is important to start simple (e.g., small tests of change) and to ensure that applications meet end users' needs and are in the best interests of ensuring patient safety, including streamlined integration into the daily workflow and experience of patients and the health care workforce.

Questions to consider:

- What problems do clinicians and staff want solved?
- Is there a genAI solution to meet this need that offers objective evidence of safety and effectiveness?
- Is an AI application feasible for our setting and is there a relevant use case?
- How will this AI application impact the safety, experience, and outcomes of patients?
- Have we assessed the potential reward and risk related to this potential AI solution? If not, how will be assess for reward and risk (e.g., a failure mode and effects analysis, if applicable)?

Organizations also need to ensure that any genAI application considered for implementation has been rigorously validated, tested, and piloted prior to scaling, including small tests of change, and that the proper processes are in place to monitor the application's performance on a consistent basis. For example, one organization engaged in our research shared that their AI oversight committee reviews data and algorithm performance every six months to ensure high-quality performance. Health systems must remember that AI technology is not simply plug-and-play, but instead requires ongoing monitoring to ensure performance quality and identify opportunities for improvement.

Moreover, how organizations communicate with staff and disseminate information on organizational genAI efforts is critical to implementation success. Clinicians and staff may have concerns about loss of autonomy in their role, staff reductions, and job loss with the introduction of AI applications.[90] Because of this, health system executives and governance

members need to reassure staff that AI applications are complementary tools that must be supplemented with clinical judgment provided by humans.[91] As the literature notes, the implementation of AI tools aims to improve the patient experience and improve clinical safety and reliability as well as allow clinicians to upskill and practice at the top of their license.[92,93] Likewise, organizations must not underestimate the importance of change management when introducing AI tools to support workflows and clinical care delivery.

Risk Identification, Mitigation, and Monitoring

Experts encourage health care organizations to develop or expand the needed governance and oversight bodies to support AI implementation in two fundamental ways.

- Identify and mitigate risks and unintended consequences associated with AI tools.
- Develop a cohesive and transparent organizational approach to AI that includes:
  - Education and training;
  - Guardrails to ensure safe use;
  - Processes for design, implementation, and monitoring; and
  - Mechanisms for collaborative learning within the organization and broader health care sector.

Collaboration and the need for learning health systems were consistently highlighted by experts as an important consideration for AI's integration into health care. Although data, algorithms, and applications are often proprietary to AI vendors, siloed efforts may not only stymie technological advances to improve care, but could also obscure repeated and preventable adverse events. Sharing experiential learning from design and implementation efforts across health care systems and the industry could help mitigate potential harm and bolster monitoring efforts to ensure that genAI is applied safely. Further, collaboration may help prevent variation. As one expert noted from their experience with new technologies, variation can lead to a lack of interoperability and complicate the landscape, opening the door for errors to occur.

For internal mitigation strategies, experts suggested the development of audit tools, including, but not limited to, algorithmic auditing, use of patient and clinician user feedback and rating of AI-generated results, and ongoing independent testing with validation sets. They cited the need for establishing accepted limits to the technology prior to implementation, such as hard end points (e.g., mortality), performance metrics, and accepted error rates for technology, if any.

This IHI Innovation Report identifies two main perspectives on potential harms arising from use of genAI tools in health care.

- The first perspective highlights the upsides of genAI tools and regards the downsides as addressable and/or not qualitatively different from existing problems in health care without AI. For instance, chatbots may at times produce medical notes with factually incorrect information (i.e., hallucinations or confabulations) and many clinicians may skip checking and correcting such inaccuracies. Some ask whether and how this differs from traditional dictated notes that often contain incorrect transcriptions and other errors that remain uncorrected in the medical record, which may then be copied and pasted as inaccurate notations without review. The optimistic perspective also points out that with use of AI tools, on balance, we might come out ahead with improvements in patient and workforce safety and well-being.

- The second perspective takes a more skeptical view on whether the promise of AI will materialize to provide benefit for the quality and safety of care and caring. After all, EHRs were also supposed to make clinicians' lives easier. Yet, a noted advantage in favor of AI is that it can address clinician burnout attributed to EHRs. This uncertain and critical stance toward AI in health care also worries about the pace at which massive and potentially irreversible changes may be instated in health care. These changes impact how clinicians and patients interact, displace the human workforce, and open the door for enormous privacy breaches, the commercialization of private health care information, and long-lasting implications of care delivery and clinical practice patterns.[94] As noted at the outset of this report, health care is posed to adopt AI tools at the same time that computer science experts, including some of the researchers who pioneered the development of these tools, have called for a pause on further AI tool development to allow time to establish regulations and other safeguards to protect society from potential risks from the technology.[95,96]

Given both the promise and apprehension surrounding AI-related developments in health care, organizations seeking to explore genAI tool integration are strongly encouraged to consider the recommendations that follow.

- **Be cautious and critical.**

  Health care systems and workers must be steadfast in their dedication to patient safety, quality care, and evidence-based practices. There are many risks detailed throughout the report (and in Appendix A) that emphasize the importance of a thoughtful and vigilant approach to AI integration in health care. While AI can relieve existing burdens within health care, its integration needs to be based on rigorous evaluation and peer-reviewed evidence.

- **Be strategic and collaborative.**

  The design, implementation, and use of genAI should be strategically constructed to align with existing organizational goals or roadmaps, meet workforce needs, fit into existing workflows, and safely support clinical care delivery. Health care systems must ensure that AI tools and applications are not prioritized over workforce and patient well-being and safety.

  As with any major new technology or process change, organizations will need to develop teams to govern, plan, and support the design, implementation, and monitoring of AI tools, with representation from different health professions (e.g., physicians and nurses), clinical informaticists, quality and safety personnel, and patients and families, among others. Stakeholder diversity is important to ensure that applications are designed and implemented to continually meet end users' needs and mitigate risks for safety and equity. Once a team or teams are established, it is important for organizations to develop standard processes for considering, designing, and/or implementing AI solutions as well as guardrails or standard practices for the professional use of AI. It is likely that these processes and guardrails will change as rapidly as the AI landscape in health care, but it necessary to ensure that structures are in place to regulate the design, implementation, and use of AI for the safety of the patients, the workforce, and the health care system.

  Further collaboration is encouraged between health care systems, academic institutions, technology companies, and others to ensure that genAI tools and applications are used safely and ethically and to pursue opportunities to improve the quality of care with AI assistance. This collaboration can also help establish industry standards, practices, and guidelines.

- **Be prepared and intentional.**

  It is important for health care systems to appropriately prepare for AI integration. This includes developing robust governance and oversight bodies (including considerations for evaluation and monitoring), educating the workforce (e.g., how AI works, appropriate use and risks of AI tools), and evaluating organizational infrastructure to ensure that a health care system can safely support the technological or computational demands of potential AI tools and related infrastructure needs. Taking the time and resources to prepare a health care system for AI integration demonstrates an intentional and proactive approach to this emerging technology, which can help mitigate potential risks or unintended consequences of integration.

Conclusion

The emergence of genAI has reignited widespread interest in and enthusiasm for implementing AI tools in health care. Yet, this enthusiasm should not distract us from identifying and addressing serious potential harms.

The genAI applications health care organizations are most likely to see implemented now or in the immediate future include documentation support and chatbots to help answer questions from patients. Given the fast-paced and dynamic nature of AI in health care in addition to the tension of the perspectives detailed above, this report does not offer any specific recommendations on how to design, implement, or regulate this technology. Rather, we stress the need for further research and evaluation to assess the quality and safety of emerging AI technologies to ensure maintenance (and even improvement) of clinical care delivery and workforce well-being.

AI in health care is a promising but complex field. As such, it requires careful consideration of ethical, legal, social, and technical issues. We strongly encourage organizations to consider building systems to support AI application design and implementation in a safe manner by building governance structures, using caution when designing, evaluating, implementing, and monitoring AI, and starting simple with small tests of change that meet end users' needs.

Appendix A: Risks, Challenges, and Unintended Consequences of Generative AI Clinical Applications in Health Care

Table 4. Risks, Challenges, and Unintended Consequences of Generative AI Clinical Applications in Health Care (listed alphabetically by category of concern)

Data Concerns	
Biased, inaccurate, aged, or inappropriate training data	Models rely heavily on their training datasets to learn how to function. Yet, datasets provided to models or algorithms may codify existing human biases or simply not include enough data involving underserved populations, resulting in inaccurate or misleading outputs. For instance, one study demonstrated higher false negative rates (i.e., missing signs of disease) in AI-interpreted radiographs involving racialized patients.[97]
Falsehood mimicry	Occurs when the user inputs a factually incorrect prompt, leading to the generative AI tool producing an erroneous output instead of first asking clarifying questions that might lead to factual correction.[98]
Hallucinations or confabulations	False or made-up statements presented as factual. "Content that is nonsensical or untruthful in relation to certain sources."[99] While hallucinations remains the commonly encountered term, confabulations is the better term for logical sounding, semantically coherent statements not based on distorted or mistaken sensory input.[100,101]
Erosion of human oversight (preserving a "human in the loop")	Implementing AI algorithms and tools without human supervision creates the potential for risk. Maintaining a "human in the loop" can help ensure that AI tools are functioning properly with reliable and accurate outputs and meeting users' needs.
Lack of AI transparency and explainability (the "black box")	AI algorithms and models are often unable to disclose how outputs (or results) are reached, hence the term "black box." This lack of transparency and explainability may lead to skepticism or mistrust among clinicians and patients.[102]
Poor quality, inaccurate, or biased data and limited data quantity	Existing data that AI tools process, such as EHR data, may be inconsistent, incomplete, missing, or incompatible. Thus, the AI output may be unreliable due to the quality of the data input (e.g., visit notes, discharge summaries).[103,104] This can also include the exclusion of relevant data (e.g., failure to incorporate variables to improve outcome accuracy). Further, available data quantity may be limited for training, testing, and auditing purposes, which may decrease the effectiveness of an AI algorithm.

Ethical Concerns	
Increase inequities related to design and access to quality AI solutions	AI technology is a double-edged sword for health equity. To improve health equity, AI can transform decision-making and medical treatment, particularly in primary and preventive care, as well as support patient-centered care.[105] The technology can also support marginalized and underserved communities such as rural communities and non-dominant language speakers.[106,107] However, AI-related barriers to health equity include codified algorithmic or dataset bias (including training and validation), inequitable accessibility or affordability of AI tools, and practice bias (e.g., clinician bias in using AI tools or interpreting/applying AI outputs, underperformance in certain groups).[108] This, in combination with concerns about cost, could further drive inequities between health care systems and their ability to deliver quality care, and patients' ability to access and afford quality care.
Limited guidance on standard operating practices and regulations	At the time of publication, standard practices and regulations such as SaMD that are relevant to AI applications in US health care remain unestablished. Moreover, some experts caution against relying on HIPAA compliance as fulfilling the duty to maintain patient confidentiality.[109]
Prioritization of commercial interests	A handful of experts noted that technology companies were pushing them to move faster than they were comfortable with, and that the Silicon Valley paradigm does not apply to health care ("move fast and break things").[110] With many health care systems partnering with technology companies and corporations, it is critical that health care prioritize patient well-being and safety. This also acknowledges that prioritization of commercial interests, including those of health care systems, could inhibit efforts to collaborate on AI application design and implementation across health care systems and across industries.
Public trust	Credibility is important for AI to be successfully integrated in health care.[111] The public, including the workforce and patients, must trust the AI tools and clinical use of the tools. Losing public trust is risky as second chances are unlikely should AI cause harm in health care ("user and regulator trust are easy to lose and very hard to regain").[112] Clinician uptake of AI applications for care delivery will rely on successful and evidence-based demonstrations that the tools are safe.[113]
"Shiny object" syndrome	AI and genAI have captured attention, but the hype should not be overstated. While this may be a worthwhile new pursuit, health care systems must note the risks involved in what could be a distraction or "shiny object."
Managerial and Operational Concerns	
Cost	While genAI solutions in health care may save health care systems money, the upfront investment may not be feasible for some organizations, including under-resourced systems.[114] This could have implications for equitable access to quality health care.

Data privacy, protection, and security	Health care data is sensitive and personal and must be protected from unauthorized or inappropriate access, theft, or misuse. AI introduces new data and infrastructural demands, raising concerns about how health care can use this advanced technology while ensuring that data is protected and secure. Risks to consider include unauthorized access, use, and control of patient data by private entities, as well as external threats (e.g., threat actors) of data breaches or leaks.[115] Additionally, patients need to be confident that their data, if they consent to share it, is being used ethically and securely.[116]
Data storage and the cloud	Health care data storage capabilities must keep up with the data demands of advanced technology such as AI. Because of this, questions emerge on whether on-site data storage is still viable or whether cloud storage should be invested in. Yet, multiple risks arise with the move to the cloud, including existing organizational policy for on-site storage of health care data, data ownership and user agreements with third parties, and patient privacy and consent to storage. Further, the cloud may not offer adequate safeguards for data encryption, backup, or recovery. Experts provided mixed responses on whether moving health care data to the cloud posed an increased risk to data security.
Liability for health care systems and clinicians	GenAI is a grey area for legal matters in health care, particularly liability. Legal liability for a health care system or clinician could include cases of data breaches, intellectual property rights, and patient consent and comprehension (e.g., if a patient does not comprehend the implications of their agreement).[117,118] Further, malpractice could cause public mistrust of new technological advances and necessitate new regulations and licensing.
Lack of interoperability	Lack of interoperability is applicable to both data and infrastructure. First, data may not be compatible or standardized across different systems or platforms. Second, with multiple applications and vendor options (as well as internal design and implementation), there is increased variation in AI application design and functionality, which may not be compatible with other applications or existing hardware and software in use in a health care system.
Lack of strategic planning	The successful design, implementation, and continued use of new technology such as AI relies on careful planning by leadership. “AI initiatives often fail because they are implemented as standalone projects. To ensure the success of AI efforts, AI must be integrated into the strategic planning process so that it aligns with the goals and objectives of the organization.”[119] Without proper strategic planning, organizations risk poor outcomes, including low clinical use or uptake and poor workflow integration, which could dampen future efforts to introduce AI applications.

Limited existing infrastructure	AI applications can require a significant amount of computational power to run. Because of this, health care systems may create risk if they implement AI without the proper IT infrastructure in place. Lack of supporting infrastructure could lead to delays in clinical care or gaps in accessing technological advances to improve care.
Poor governance and oversight	Inadequate governance or oversight of emerging technologies such as genAI could increase the potential for other risks listed in this table to translate into real-world harm, including ethical concerns on privacy, trust, and bias as well as regulatory concerns.[120]
Potential for downtime	All technology can malfunction and must be maintained and updated. As genAI solutions are integrated, systems must consider the risk for planned and unplanned downtime to ensure continuity of care.
Patient and Workforce Concerns	
Alert fatigue	Refers to the desensitization of clinicians to safety alerts, which causes clinicians to ignore or fail to appropriately respond to warning systems.[121]
Automation bias, overreliance, clinical complacency, and deskilling	These terms collectively refer to the possibility that the use of genAI tools in health care will lead to a decrease in workforce competency. Overreliance "... occurs when users excessively trust and depend on the model, potentially leading to unnoticed mistakes and inadequate oversight."[122] This can lead to automation bias, which is "the tendency to over-rely on automation."[123] Clinical complacency and deskilling, then, can occur as clinicians may no longer employ or improve their skills.[124] This can lead to a loss of skills, inability to reliably detect mistakes or errors, and reduction in decision-making quality, diagnostic reasoning, and patient communication.
Difficulty integrating into daily clinical workflow	For genAI tools and applications to be successful, they must integrate into existing clinical workflows. An evidence-based and clinically tested genAI tool will be underutilized if it is inconvenient to a clinician or staff member's daily workflow. Risks include improper use or underutilization of tools that could improve care delivery and patient outcomes.
Fear that care will be depersonalized	Refers to the concern that the implementation of AI in clinical care settings will negatively impact the patient-clinician dynamic. This includes the possibility that clinicians may prioritize the clinician-technology relationship over the patient.
Lack of patient education on AI	The general patient population may be unfamiliar with the use of genAI tools and applications in health care. Health care systems and clinicians must ensure that patients are aware of the uses and risks of genAI tools as well as their patient rights, including those pertaining to patient data. Further, clinicians need to be aware that, even though patients may lack education on AI, they may be using it in their daily lives, including asking applications like ChatGPT for medical advice.

Lack of proactive safety oversight and engagement of related expertise	Implementing AI requires engaging multiple stakeholders, including for planning, evaluation, and decision-making. The need for multiple stakeholders can be challenging, but the process should not advance without adequate resources and expertise from safety, quality, and risk teams. Further, use a proactive safety approach through design and integration along with appropriate expertise.
Misalignment with patient preference	Patients may prefer not to receive clinical care assisted by genAI tools. In 2023, Pew Research found that 60% of surveyed Americans were uncomfortable with their health care provider using AI.[125] Health care systems must consider how to handle patients who opt out of AI tools, and adequately educate patients on AI tool use for those who opt in.
Workforce resistance	With the introduction of genAI tools and applications, the health care workforce may be resistant to this integration for multiple reasons, including fear of replacement and workforce displacement.[126,127] Resistance by the workforce could reduce the likelihood of successful implementation of tools aimed at improving care quality and workplace environment and workload.

Appendix B: Non-Clinical Applications of Generative AI in Health Care

More detailed information for each genAI application follows Table 5.

Table 5. Overview of Non-Clinical Generative AI Applications in Health Care

GenAI Application	Potential Benefits	Considerations
Publications	• Assist with literature reviews • Create or edit written drafts	• Intellectual property rights • Authorship • Hallucinations and confabulations such as with content and references
Clinical trials and medical research	• Generate and synthesize data • Propose clinical trial design • Offer outcome predictions • Detect the likelihood of adverse effects or events • Streamline and improve appropriateness of clinical trial recruitment	• Regulatory compliance • Overreliance
Education for medical and other health care students Education for patients	• Produce medical simulations or narrative cases for medical students • Provide an assessment tool for medical students • Provide patient education about their condition(s), medication(s), and treatment plan • Support health literacy for the general population	• Ensuring appropriate use • Plagiarism • Falsehood mimicry (see Appendix A) • Hallucinations and confabulations • Language and health literacy
IT support	• Provide middleware to support interoperability of existing platforms and systems • Predict maintenance for technology	• Potential for downtime • Aged datasets
Administrative assistance	• Streamline appointment scheduling • Act as a digital assistant	• Potential for downtime
Medical billing, coding, and claims	• Automate the process of coding medical procedures and services • Identify and correct coding errors • Optimize revenue and ensure compliance • Draft documentation for pre-authorization, claims, reimbursements, and appeals	• Poor data quality (or need for data fidelity) • Denial of legitimate claims • Increased clinician time to challenge denials

Safety and quality improvement practices	• Provide data analysis, data visualization, process analysis, qualitative analysis, and guidance for developing and testing changes	• Safety, accuracy, and reliability of LLMs in health care • Data concerns (e.g., biased, inaccurate, out-of-date, or inappropriate training) • Hallucinations • Data privacy and security concerns • Liability or copyright implications • Workforce displacement • Environmental impact (e.g., high greenhouse gas emissions due to demand of computing power and enormous use of water for cooling data centers)

Publications

GenAI has the demonstrated ability to assist with research and writing and may be useful for clinical publications. The technology could be used to support medical writing (including peer-reviewed publications and practice standards), assisting with literature reviews, and editing human-written drafts or providing written drafts for humans to edit.[128,129]

Concerns about this usage have been raised, including intellectual property rights, authorship, hallucinations, and confabulations, including fabricating nonexistent references. Because of this, journals such as *JAMA* have published guidance for authors, peer reviewers, and editors on the use of these technologies, including AI, LLM, and chatbots.[130]

Clinical Trials and Medical Research

GenAI could reshape clinical trials and medical research processes by accelerating timelines and enhancing outcomes, especially for drug discovery and development. Using historical clinical trial data and established trial parameters (e.g., desired outcomes), genAI could support clinical trials by:

- Generating and synthesizing data;
- Proposing clinical trial design;
- Offering outcome predictions;
- Detecting the likelihood of adverse effects or events; and
- Streamlining trial recruitment and targeting potential enrollment opportunities based on predefined criteria.

Yet, the use of genAI in this capacity is untested and must be rigorously evaluated and validated prior to implementation. If utilized, applications need to comply with regulatory requirements, including HIPAA, existing or pending Food and Drug Administration regulations, and other national or international standards and regulations. Further, even with genAI support, clinical researchers cannot become complacent or overreliant on technological assistance; a human must always be in the loop if technological applications are used.

Education for Medical and Other Health Care Students

GenAI could enrich educational opportunities for medical students. As an experiential tool, genAI can be used to produce medical simulations or narrative cases for medical students, which could be offered on demand, to work through to test and enhance their critical thinking and clinical reasoning.[131] Although untested as an assessment tool, genAI, specifically ChatGPT, has demonstrated the ability to comprehend and respond to standardized tests in higher education. For example, one study found that ChatGPT performed above the 60 percent threshold on the United States Medical Licensing Examination Step 1 (using publicly available sample questions).[132] While only theoretical, it is possible that genAI could be used as an assessment tool for medical students, in conjunction with human proctors to ensure that questions and responses are accurately communicated. This function may also assist with producing appropriate questions for standardized tests and could be used for continuing education for health care professionals.

Implications for genAI in education include ensuring appropriate use by faculty and students, concerns about plagiarism (including inadvertent cases), and application issues such as false mimicry, hallucinations, and confabulations. Many universities have made an effort to provide guidance for the use of genAI both for students and faculty, including Harvard University, the University of North Carolina–Chapel Hill, and Tufts University.[133,134,135]

Education for Patients

GenAI could provide patients with education on their conditions or illnesses, medications, and treatment plans. GenAI applications, like ChatGPT, have the capacity to provide patient education. For example, a group of physicians assessed ChatGPT on its knowledge of colonoscopies and found that it provided adequate and original responses to frequently asked questions about the procedure. Further, some companies have taken strides to support patient education, such as Vital's AI-powered video education feature that auto-prescribes education to patients in the emergency department and inpatient settings.[136,137]

IT Support

GenAI may support two functions for health care IT:

- GenAI applications may be capable of providing middleware to support interoperability between systems and platforms.
- These applications can provide predictive maintenance for technology.[138]

Beyond routine maintenance, genAI can be used to anticipate machine malfunctions and to plan maintenance prior to errors occurring, which could decrease delay and replacement costs. Even with technological support for IT maintenance and predictive malfunction, manual maintenance must still be routine, including maintenance and auditing of any genAI supporting health care IT.

Administrative Assistance

There are several potential benefits of genAI for health care administration:

- GenAI can optimize scheduling processes for health care services.[139] GenAI applications can automate the tasks of scheduling appointments, centralizing scheduling across multiple locations and departments, managing call centers, balancing workload among staff members, and forecasting appointment demand.[140]
- GenAI can enhance the role of digital assistants in health care settings by enabling digital assistants to perform diverse functions, such as reviewing daily schedules, evaluating appropriateness of visit type, answering queries, providing guidance, collecting feedback, and generating reports.

Downtime options and backup procedures must remain available in case of emergencies or routine maintenance interruptions. Staff using genAI for scheduling or related responsibilities need to be trained appropriately, both in manual and genAI supported methods, to ensure efficient and safe use of AI tools for patients and to increase workforce satisfaction with supplemental support. In all cases, the use of AI should be closely monitored to ensure that users do not become overly reliant on digital assistance.

Medical Billing, Coding, and Claims

GenAI can streamline billing, coding, and claims for health care systems. Generative AI can help automate the process of coding medical procedures and services, identify and correct coding errors, optimize revenue and ensure compliance, and draft documentation pertaining to pre-authorizations, claims, reimbursements, and appeals.[141,142] If genAI is used for medical billing, coding, and claims, it is important to ensure data fidelity prior to implementation, otherwise the application may struggle to output accurate information. Even if data fidelity is established, an appropriately trained human needs to remain in the loop and the AI algorithms and data need regular auditing. Established risks to these applications include the inability of AI to make informed, evidence-based decisions, and the high level of denials of prior authorizations has created extensive challenges, leading to patient harm and excessive burden and costs for clinicians to advocate for appropriate care.[143]

Safety and Quality Improvement Practices

Safety and quality improvement efforts in health care may benefit from genAI tools and applications, including data analysis, data visualization, process analysis, qualitative analysis, and guidance for developing and testing changes. These uses can support the maintenance and

improvement of the quality and safety of clinical care delivery. Notable challenges and risks include concerns on the safety, accuracy, and reliability of LLMs in health care; data concerns (e.g., biased, inappropriate training data, out-of-date data); hallucinations; data privacy and security concerns; liability or copyright implications; workforce displacement; and environmental impact (e.g., high greenhouse gas emissions due to demand of computing power and enormous use of water for cooling data centers).

References

[1] Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digit Health*. 2023;2(2):e0000198.

[2] Nayak A, Alkaitis MS, Nayak K, Nikolov M, Weinfurt KP, Schulman K. Comparison of history of present illness summaries generated by a chatbot and senior internal medicinerResidents. *JAMA Intern Med*. 2023;183(9):1026.

[3] Ayers JW, Zhu Z, Poliak A, et al. Evaluating artificial intelligence responses to public health questions. *JAMA Netw Open*. 2023;6(6):e2317517.

[4] Ayers JW, Poliak A, Dredze M, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Intern Med*. 2023;183(6):589.

[5] Weiser B. Here's what happens when your lawyer uses ChatGPT. *New York Times*. May 27, 2023. https://www.nytimes.com/2023/05/27/nyregion/avianca-airline-lawsuit-chatgpt.html

[6] Donker T. The dangers of using large language models for peer review. *The Lancet Infectious Diseases*. 2023;23(7):781.

[7] Lee P, Bubeck S, Petro J. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. *N Engl J Med*. 2023;388(13):1233-1239.

[8] Lee P, Bubeck S, Petro J. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. *N Engl J Med*. 2023;388(13):1233-1239.

[9] Tschandl P. Risk of bias and error from data sets used for dermatologic artificial intelligence. *JAMA Dermatol*. 2021;157(11):1271.

[10] Seyyed-Kalantari L, Zhang H, McDermott MBA, Chen IY, Ghassemi M. Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. *Nat Med*. 2021;27(12):2176-2182.

[11] Leslie D, Mazumder A, Peppin A, Wolters MK, Hagerty A. Does "AI" stand for augmenting inequality in the era of Covid-19 healthcare? *BMJ*. 2021 Mar 15:372:n304.

[12] Johnson-Mann CN, Loftus TJ, Bihorac A. Equity and artificial intelligence in surgical care. *JAMA Surg*. 2021;156(6):509.

[13] Parikh RB, Teeple S, Navathe AS. Addressing bias in artificial intelligence in health care. *JAMA*. 2019;322(24):2377.

[14] Tschandl P. Risk of bias and error from data sets used for dermatologic artificial intelligence. *JAMA Dermatol*. 2021;157(11):1271.

[15] Tschandl P. Risk of bias and error from data sets used for dermatologic artificial intelligence. *JAMA Dermatol*. 2021;157(11):1271.

[16] Chan CYT, Petrikat D. Strategic applications of artificial intelligence in health care and medicine. *Journal of Medical and Health Studies*. 2023;4(3):58-68.

[17] Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med*. 2019;380(14):1347-1358.

[18] Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115-118.

[19] Ehteshami Bejnordi B, Veta M, Johannes Van Diest P, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA*. 2017;318(22):2199.

[20] Ting DSW, Cheung CYL, Lim G, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA*. 2017;318(22):2211.

[21] Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402.

[22] Manning C. *Artificial Intelligence Definitions*. Stanford University, Human- Centered Artifical Intelligence; 2020:1-2. https://hai.stanford.edu/sites/default/files/2020-09/AI-Definitions-HAI.pdf

[23] Favaretto M, De Clercq E, Schneble CO, Elger BS. What is your definition of Big Data? Researchers' understanding of the phenomenon of the decade. *PLoS One*. 2020;15(2):e0228987.

[24] Toner H. What Are Generative AI, Large Language Models, and Foundation Models? Center for Security and Emerging Technology Blog. May 12, 2023. https://cset.georgetown.edu/article/what-are-generative-ai-large-language-models-and-foundation-models/

[25] Harrer S. Attention is not all you need: The complicated case of ethically using large language models in health care and medicine. *eBioMedicine*. 2023;90:104512.

[26] Toner H. What Are Generative AI, Large Language Models, and Foundation Models? Center for Security and Emerging Technology Blog. May 12, 2023. https://cset.georgetown.edu/article/what-are-generative-ai-large-language-models-and-foundation-models/

[27] Javaid M, Haleem A, Singh RP. ChatGPT for health care services: An emerging stage for an innovative perspective. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*. 2023;3(1):100105.

[28] Toner H. What Are Generative AI, Large Language Models, and Foundation Models? Center for Security and Emerging Technology Blog. May 12, 2023. https://cset.georgetown.edu/article/what-are-generative-ai-large-language-models-and-foundation-models/

[29] Brown S. Machine Learning, Explained. *Ideas Made to Matter*. MIT Management Sloan School. April 21, 2021. https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained

[30] Gruetzemacher R. The power of natural language processing. *Harvard Buisness Review*. April 19, 2022. https://hbr.org/2022/04/the-power-of-natural-language-processing

[31] Kwan J. Working smarter: AI's growing role in nursing care. *Voice of Nursing Leadership*. 2022;20(60):8-10.

[32] Neural networks: What they are and why they matter. SAS. https://www.sas.com/en_in/insights/analytics/neural-networks.html

[33] Harrer S. Attention is not all you need: The complicated case of ethically using large language models in health care and medicine. *eBioMedicine*. 2023;90:104512.

[34] Harrer S. Attention is not all you need: The complicated case of ethically using large language models in healthcare and medicine. *eBioMedicine*. 2023;90:104512.

[35] Norden J, Wang J, Bhattacharyya A. Where generative AI meets healthcare: Updating the healthcare AI landscape. *AI Checkup*. June 22, 2023. https://aicheckup.substack.com/p/where-generative-ai-meets-healthcare

[36] Kulkarni PA, Singh H. Artificial intelligence in clinical diagnosis: Opportunities, challenges, and hype. *JAMA*. 2023 Jul 25;330(4):317-318.

[37] Bates DW, Levine D, Syrowatka A, et al. The potential of artificial intelligence to improve patient safety: A scoping review. *NPJ Digit Med*. 2021;4(1):54.

[38] Zhang Y, Pei H, Zhen S, Li Q, Liang F. Chat Generative Pre-Trained Transformer (ChatGPT) usage in health care. *Gastroenterology & Endoscopy*. 2023;1(3):139-143.

[39] Hswen Y, Voelker R. Electronic health records failed to make clinicians' lives easier—will AI technology succeed? *JAMA*. 2023 Oct 24;330(16):1509-1511.

[40] Nguyen OT, Jenkins NJ, Khanna N, et al. A systematic review of contributing factors of and solutions to electronic health record–related impacts on physician well-being. *J Am Med Inform Assoc*. 2021;28(5):974-984.

[41] Strudwick G, Tajirian T, Kemp J, et al. Utilizing an Informatics Engagement Strategy as an Approach to Sustain and Retain the Nursing Workforce. *Nurs Leadersh (Tor Ont)*. 2023;35(4):42-54.

[42] Melnick ER, Dyrbye LN, Sinsky CA, et al. The association between perceived electronic health record usability and professional burnout among US physicians. *Mayo Clinic Proceedings*. 2020;95(3):476-487.

[43] Kwan J. Working smarter: AI's growing role in nursing care. *Voice of Nursing Leadership*. 2022;20(60):8-10.

[44] Ambient Listening Helps Primary Care Doctors Finish Notes Faster. Epic Share. January 2, 2023. https://www.epicshare.org/share-and-learn/umhw-ambient-listening-notes

[45] Schwartz N. HCA piloting AI for clinical documentation. *Becker's Health IT*. April 20, 2023. https://www.beckershospitalreview.com/innovation/hca-piloting-ai-for-clinical-documentation.html

[46] Diaz N. 7 health systems using AI to ease clinicians' documentation burden. *Becker's Health IT*. May 15, 2023. https://www.beckershospitalreview.com/innovation/7-health-systems-using-ai-to-ease-clinicians-documentation-burden.html

[47] Lee P, Bubeck S, Petro J. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. *N Engl J Med*. 2023;388(13):1233-1239.

[48] Marks M, Haupt CE. AI chatbots, health privacy, and challenges to HIPAA compliance. *JAMA*. 2023;330(4):309.

[49] March 20 ChatGPT outage: Here's what happened. OpenAI. March 24, 2023. https://openai.com/blog/march-20-chatgpt-outage

[50] Marks M, Haupt CE. AI chatbots, health privacy, and challenges to HIPAA compliance. *JAMA*. 2023;330(4):309.

[51] Generative AI: How Houston Methodist Is Leading the Charge to Unburden Care Staff. Pieces Webinar; 2023. https://go.piecestech.com/generative-ai-houston-methodist-webinar

[52] Kennedy S. HCA Health care Launches AI-Enabled Ambient Documentation Partnership. *HealthITAnalytics*. April 20, 2023. https://healthitanalytics.com/news/hca-healthcare-launches-ai-enabled-ambient-documentation-partnership

[53] HCA Healthcare Collaborates with Google Cloud to Bring Generative AI to Hospitals. HCA Healthcare. August 29, 2023. https://investor.hcahealthcare.com/news/news-details/2023/HCA-Healthcare-Collaborates-With-Google-Cloud-to-Bring-Generative-AI-to-Hospitals/default.aspx

[54] Schwartz N. Epic gets new generative AI "pal." *Becker's Health IT*. August 16, 2023. https://www.beckershospitalreview.com/ehrs/epic-gets-new-generative-ai-pal.html

[55] Abridge becomes Epic's First Pal, bringing generative AI to more providers and patients, including those at Emory Healthcare. Emory University. August 16, 2023. https://news.emory.edu/stories/2023/08/hs_ehc_abridge_epic_ai_ambient_listening_tool_16-08-2023/story.html

[56] Schwartz N. Epic gets new generative AI "pal." *Becker's Health IT*. August 16, 2023. https://www.beckershospitalreview.com/ehrs/epic-gets-new-generative-ai-pal.html

[57] Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: Benefits, risks, and strategies for success. *NPJ Digit Med*. 2020;3(1):17.

[58] Fox A. ChatGPT scored 72% in clinical decision accuracy, MGB study shows. *Healthcare IT News*. August 29, 2023. https://www.healthcareitnews.com/news/chatgpt-scored-72-clinical-decision-accuracy-mgb-study-shows

[59] Kanjee Z, Crowe B, Rodman A. Accuracy of a generative artificial intelligence model in a complex diagnostic challenge. *JAMA*. 2023;330(1):78.

[60] Kwan J. Working smarter: AI's growing role in nursing care. *Voice of Nursing Leadership*. 2022;20(60):8-10.

[61] Kaur J. Generative AI in healthcare and its uses: Complete guide. Xenonstack. December 13, 2023. https://www.xenonstack.com/blog/generative-ai-healthcare

[62] Kaur J. Generative AI in healthcare and its uses: Complete guide. Xenonstack. December 13, 2023. https://www.xenonstack.com/blog/generative-ai-healthcare

[63] Tschandl P. Risk of bias and error from data sets used for dermatologic artificial intelligence. *JAMA Dermatol*. 2021;157(11):1271.

[64] Kwan JL, Lo L, Ferguson J, et al. Computerised clinical decision support systems and absolute improvements in care: meta-analysis of controlled clinical trials. *BMJ*. 2020 Sep 17:370:m3216.

[65] Shah SN, Amato MG, Garlo KG, Seger DL, Bates DW. Renal medication-related clinical decision support (CDS) alerts and overrides in the inpatient setting following implementation of a commercial electronic health record: Implications for designing more effective alerts. *J Am Med Inform Assoc*. 2021;28(6):1081-1087.

[66] Escobar GJ, Liu VX, Schuler A, Lawson B, Greene JD, Kipnis P. Automated identification of adults at risk for in-hospital clinical deterioration. *N Engl J Med*. 2020;383(20):1951-1960.

[67] Dean NC, Jones BE, Jones JP, et al. Impact of an electronic clinical decision support tool for emergency department patients with pneumonia. *Annals of Emergency Medicine*. 2015;66(5):511-520.

[68] Marr B. The amazing ways Babylon Health is using artificial intelligence to make health care universally accessible. *Forbes*. August 16, 2019. https://www.forbes.com/sites/bernardmarr/2019/08/16/the-amazing-ways-babylon-health-is-using-artificial-intelligence-to-make-healthcare-universally-accessible/?sh=75c3742e3842

[69] Adams K. Babylon Health's CEO "should spend some time with Elizabeth Holmes." *MedCityNews*. September 12, 2023. https://medcitynews.com/2023/09/babylon-health care-ai-uk/

[70] Fox A. Intermountain launches interoperable, real-time CDS platform. *Healthcare IT News*. April 24, 2023. https://www.healthcareitnews.com/news/intermountain-launches-interoperable-real-time-cds-platform

[71] Dean NC, Jones BE, Jones JP, et al. Impact of an electronic clinical decision support tool for emergency department patients with pneumonia. *Annals of Emergency Medicine*. 2015;66(5):511-520.

[72] Kwan J. Working smarter: AI's growing role in nursing care. *Voice of Nursing Leadership*. 2022;20(60):8-10.

[73] Javaid M, Haleem A, Singh RP. ChatGPT for health care services: An emerging stage for an innovative perspective. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*. 2023;3(1):100105.

[74] Kwan J. Working smarter: AI's growing role in nursing care. *Voice of Nursing Leadership*. 2022;20(60):8-10.

[75] Kwan J. Working smarter: AI's growing role in nursing care. *Voice of Nursing Leadership*. 2022;20(60):8-10.

[76] Ayers JW, Poliak A, Dredze M, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Intern Med*. 2023;183(6):589.

[77] Talaga R. How UNC Health's Epic InBasket pilot will inform the way AI is used in EHRs. *Becker's Health IT*. August 2, 2023. https://www.beckershospitalreview.com/ehrs/how-unc-healths-epic-chatbot-pilot-will-inform-the-way-ai-is-used-in-ehrs.html

[78] Fox A. UNC Health pilots generative AI chatbot. *Healthcare IT News*. June 26, 2023. https://www.healthcareitnews.com/news/unc-health-pilots-generative-ai-chatbot

[79] Lund F, Maor D, Spielmann N, Sukharevsky A. Four Essential Questions for Boards to Ask about Generative AI. QuantumBlack AI by McKinsey. July 27, 2023. https://www.mckinsey.com/capabilities/quantumblack/our-insights/four-essential-questions-for-boards-to-ask-about-generative-ai

[80] Harrer S. Attention is not all you need: The complicated case of ethically using large language models in health care and medicine. *eBioMedicine*. 2023;90:104512.

[81] Sezgin E. Artificial intelligence in health care: Complementing, not replacing, doctors and health care providers. *Digit Health*. 2023;9:20552076231186520.

[82] Duffourc M, Gerke S. Generative AI in health care and liability risks for physicians and safety concerns for patients. *JAMA*. 2023 Jul 25;330(4):313-314.

[83] What is ABCDS? Duke AI Health. https://aihealth.duke.edu/algorithm-based-clinical-decision-support-abcds/

[84] Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. *J Am Med Inform Assoc*. 2020;27(3):491-497.

[85] WHO Calls for Safe and Ethical AI for Health. World Health Organization. May 16, 2023. https://www.who.int/news/item/16-05-2023-who-calls-for-safe-and-ethical-ai-for-health

[86] Ellahham S, Ellahham N, Simsekler MCE. Application of artificial intelligence in the health care safety context: Opportunities and challenges. *Am J Med Qual*. 2020;35(4):341-348.

[87] Health Care Artificial Intelligence Code of Conduct. National Academy of Medicine. https://nam.edu/programs/value-science-driven-health-care/health-care-artificial-intelligence-code-of-conduct/

[88] Providing Guidelines for the Responsible Use of AI in Healthcare. Coalition for Health AI. https://www.coalitionforhealthai.org/

[89] *Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence/Machine Learning (AI/ML)-Enabled Device Software Functions: Draft Guidance for Industry and Food and Drug Administration Staff*. US Food and Drug Administration; April 2023. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/marketing-submission-recommendations-predetermined-change-control-plan-artificial

[90] Kwan J. Working smarter: AI's growing role in nursing care. *Voice of Nursing Leadership*. 2022;20(60):8-10.

[91] Sezgin E. Artificial intelligence in health care: Complementing, not replacing, doctors and health care providers. *Digit Health*. 2023;9:20552076231186520.

[92] Ross P, Spates K. Considering the safety and quality of artificial intelligence in health care. *Jt Comm J Qual Patient Saf*. 2020;46(10):596-599.

[93] *GPT-4 System Card*. OpenAI. March 23, 2023. https://cdn.openai.com/papers/gpt-4-system-card.pdf

[94] Marks M, Haupt CE. AI chatbots, health privacy, and challenges to HIPAA compliance. *JAMA*. 2023;330(4):309.

[95] Clarke L. Alarmed tech leaders call for AI research pause. *Science*. April 11, 2023. https://www.science.org/content/article/alarmed-tech-leaders-call-ai-research-pause

[96] Schneier B, Sanders N. The AI wars have three factions, and they all crave power. *New York Times*. September 28, 2023:1-7.

[97] Seyyed-Kalantari L, Zhang H, McDermott MBA, Chen IY, Ghassemi M. Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. *Nat Med*. 2021;27(12):2176-2182.

[98] Au Yeung J, Kraljevic Z, Luintel A, et al. AI chatbots not yet ready for clinical use. *Front Digit Health*. 2023;5:1161098.

[99] *GPT-4 System Card*. OpenAI. March 23, 2023. https://cdn.openai.com/papers/gpt-4-system-card.pdf

[100] Hatem R, Simmons B, Thornton JE. Chatbot confabulations are not hallucinations. *JAMA Intern Med*. 2023;183(10):1177.

[101] Edwards B. Why ChatGPT and Bing Chat are so good at making things up. *Ars Technica*. April 6, 2023. https://arstechnica.com/information-technology/2023/04/why-ai-chatbots-are-the-ultimate-bs-machines-and-how-people-hope-to-fix-them/

[102] Kaur J. Generative AI in healthcare and its uses: Complete guide. Xenonstack. May 11, 2023. https://www.xenonstack.com/blog/generative-ai-healthcare

[103] Kulkarni PA, Singh H. Artificial intelligence in clinical diagnosis: Opportunities, challenges, and hype. *JAMA*. 2023 Jul 25;330(4):317-318.

[104] Kulkarni PA, Singh H. Artificial intelligence in clinical diagnosis: Opportunities, challenges, and hype. *JAMA*. 2023 Jul 25;330(4):317-318.

[105] Gurevich E, El Hassan B, El Morr C. Equity within AI systems: What can health leaders expect? *Healthc Manage Forum*. 2023;36(2):119-124.

[106] Guo J, Li B. The application of medical artificial intelligence technology in rural areas of developing countries. *Health Equity*. 2018;2(1):174-181.

[107] Chonde DB, Pourvaziri A, Williams J, et al. RadTranslate: An artificial intelligence–powered intervention for urgent imaging to enhance care equity for patients with limited English proficiency during the COVID-19 pandemic. *Journal of the American College of Radiology*. 2021;18(7):1000-1008.

[108] Abràmoff MD, Tarver ME, Loyo-Berrios N, et al. Considerations for addressing bias in artificial intelligence for health equity. *NPJ Digit Med*. 2023;6(1):170.

[109] Marks M, Haupt CE. AI chatbots, health privacy, and challenges to HIPAA compliance. *JAMA*. 2023;330(4):309.

[110] Harrer S. Attention is not all you need: The complicated case of ethically using large language models in health care and medicine. *eBioMedicine*. 2023;90:104512.

[111] Zhang Y, Pei H, Zhen S, Li Q, Liang F. Chat Generative Pre-Trained Transformer (ChatGPT) usage in health care. *Gastroenterology & Endoscopy*. 2023;1(3):139-143.

[112] Harrer S. Attention is not all you need: The complicated case of ethically using large language models in health care and medicine. *eBioMedicine*. 2023;90:104512.

[113] Ross P, Spates K. Considering the safety and quality of artificial intelligence in health care. *Jt Comm J Qual Patient Saf*. 2020;46(10):596-599.

[114] Chan CYT, Petrikat D. Strategic applications of artificial intelligence in health care and medicine. *Journal of Medical and Health Studies*. 2023;4(3):58-68.

[115] Ellahham S, Ellahham N, Simsekler MCE. Application of artificial intelligence in the health care safety context: Opportunities and challenges. *Am J Med Qual*. 2020;35(4):341-348.

[116] WHO Calls for Safe and Ethical AI for Health. World Health Organization. May 16, 2023. https://www.who.int/news/item/16-05-2023-who-calls-for-safe-and-ethical-ai-for-health

[117] Javaid M, Haleem A, Singh RP. ChatGPT for health care services: An emerging stage for an innovative perspective. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*. 2023;3(1):100105.

[118] Zhang Y, Pei H, Zhen S, Li Q, Liang F. Chat Generative Pre-Trained Transformer (ChatGPT) usage in health care. *Gastroenterology & Endoscopy*. 2023;1(3):139-143.

[119] Chan CYT, Petrikat D. Strategic applications of artificial intelligence in health care and medicine. *Journal of Medical and Health Studies*. 2023;4(3):58-68.

[120] Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. *J Am Med Inform Assoc*. 2020;27(3):491-497.

[121] Agency for Healthcare Research and Quality. Alert Fatigue. Patient Safety Network. September 7, 2019. https://psnet.ahrq.gov/primer/alert-fatigue

[122] *GPT-4 System Card*. OpenAI. March 23, 2023. https://cdn.openai.com/papers/gpt-4-system-card.pdf

[123] Goddard K, Roudsari A, Wyatt JC. Automation bias: A systematic review of frequency, effect mediators, and mitigators. *J Am Med Inform Assoc*. 2012;19(1):121-127.

[124] Ross P, Spates K. Considering the safety and quality of artificial intelligence in health care. *Jt Comm J Qual Patient Saf*. 2020;46(10):596-599.

[125] Tyson A, Pasquini G, Spencer A, Funk C. *60% of Americans Would Be Uncomfortable With Provider Relying on AI in Their Own Health Care*. Pew Research Center; 2023. https://www.pewresearch.org/science/2023/02/22/60-of-americans-would-be-uncomfortable-with-provider-relying-on-ai-in-their-own-health-care/

[126] Javaid M, Haleem A, Singh RP. ChatGPT for health care services: An emerging stage for an innovative perspective. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*. 2023;3(1):100105.

[127] *GPT-4 System Card*. OpenAI. March 23, 2023. https://cdn.openai.com/papers/gpt-4-system-card.pdf

[128] Dave T, Athaluri SA, Singh S. ChatGPT in medicine: An overview of its applications, advantages, limitations, future prospects, and ethical considerations. *Front Artif Intell*. 2023;6:1169595.

[129] Zhang Y, Pei H, Zhen S, Li Q, Liang F. Chat Generative Pre-Trained Transformer (ChatGPT) usage in health care. *Gastroenterology & Endoscopy*. 2023;1(3):139-143.

[130] Flanagin A, Kendall-Taylor J, Bibbins-Domingo K. Guidance for authors, peer reviewers, and editors on use of AI, language models, and chatbots. *JAMA*. 2023;330(8):702.

[131] Gilson A, Safranek CW, Huang T, et al. How does ChatGPT perform on the United States medical licensing examination? The implications of large language models for medical education and knowledge assessment. *JMIR Med Educ*. 2023;9:e45312.

[132] Gilson A, Safranek CW, Huang T, et al. How does ChatGPT perform on the United States medical licensing examination? The implications of large language models for medical education and knowledge assessment. *JMIR Med Educ*. 2023;9:e45312.

[133] Office of the Provost. Guidelines for Using ChatGPT and other Generative AI tools at Harvard. Harvard University. https://provost.harvard.edu/guidelines-using-chatgpt-and-other-generative-ai-tools-harvard

[134] Office of the Provost. Generative AI Usage Guidance: Student Generative AI Usage. University of North Carolina, Chapel Hill. https://provost.unc.edu/student-generative-ai-usage-guidance/

[135] Technology Services. Guidelines for Use of Generative AI Tools. Tufts University. https://it.tufts.edu/guidelines-use-generative-ai-tools

[136] Heath S. ChatGPT continues to prove useful for patient education. *PatientEngagementHIT*. May 18, 2023. https://patientengagementhit.com/news/chatgpt-continues-to-prove-useful-for-patient-education

[137] Vital advances precision patient education with the introduction of the industry's first AI-powered video education feature. *Buisness Wire*. June 27, 2023. https://www.businesswire.com/news/home/20230627129241/en/Vital-Advances-Precision-Patient-Education-with-the-Introduction-of-the-Industry%E2%80%99s-First-AI-Powered-Video-Education-Feature

[138] Chan CYT, Petrikat D. Strategic applications of artificial intelligence in healthcare and medicine. *Journal of Medical and Health Studies.* 2023;4(3):58-68.

[139] Chan CYT, Petrikat D. Strategic applications of artificial intelligence in healthcare and medicine. *Journal of Medical and Health Studies.* 2023;4(3):58-68.

[140] Kwan J. Working smarter: AI's growing role in nursing care. *Voice of Nursing Leadership*. 2022;20(60):8-10.

[141] Aronson S, Lieu TW, Scirica BM. Getting generative AI right. *NEJM Catalyst*. May 2, 2023. https://catalyst.nejm.org/doi/full/10.1056/CAT.23.0063

[142] Bhasker S, Bruce D, Lamb J, Stein G. Tackling Healthcare's Biggest Burdens with Generative AI. McKinsey & Company. July 10, 2023. https://www.mckinsey.com/industries/healthcare/our-insights/tackling-healthcares-biggest-burdens-with-generative-ai

[143] Henry TA. Oversight needed on payers' use of AI in prior authorization. American Medical Association. June 14, 2023. https://www.ama-assn.org/practice-management/prior-authorization/oversight-needed-payers-use-ai-prior-authorization