Patient Safety and Artificial Intelligence
Considerations for Key Groups

Institute for Healthcare Improvement

GenAI Developers

In the dynamic landscape of health care and technology, generative artificial intelligence (genAI) developers must balance innovation with human well-being and safety. The IHI Lucian Leape Institute offers the following recommendations for genAI developers:

- **Prioritize transparent and explainable AI design to build trust:** Develop genAI tools that are transparent in their operations and decisions, so patients and clinicians can understand, and build trust in, the outputs. Ideally, once implemented, AI tools provide users with an indication of how confident users should be with each answer or output and cite evidence-based sources, when appropriate. Conduct ongoing evaluations to ensure that AI tools are performing well and to implement improvements.

- **Establish two-way dialogue with users:** Collaborate closely with health care professionals, safety leaders, patients, and patient advocates to gather feedback and iteratively improve AI tools, ensuring they align with existing workflows and enhance patient care and safety. AI developers must be proactive with these efforts, continually iterating on improvements by developing effective feedback channels. In addition, implement mechanisms that give designers signals when AI results are flawed (e.g., frequent rejection by clinicians of a specific piece of AI decision support). Finally, developers need to build tools that increase the probability that "human in the loop" double-checks provide real safety, not the illusion of safety.

- **Commit to learning about and mitigating patient safety risks:** To create effective, ethical, and safe health care solutions, AI developers must ensure that they are knowledgeable about existing and potential patient safety risks. Developers also

IHI Lucian Leape Institute Expert Panel Report on Patient Safety and AI

In January 2024, the IHI Lucian Leape Institute convened an expert panel to further explore the promise of generative artificial intelligence (genAI) and its potential risks for patient safety.

The panel reviewed the literature on AI and patient safety and engaged in a robust discussion that focused on three likely use cases for genAI in health care: documentation support, clinical decision support, and patient-facing chatbots.

The panel also reviewed considerations for key groups and provided specific recommendations and mitigation strategies for these audiences.

Visit **ihi.org/LLISafetyAI**

need to contribute to ongoing monitoring of their AI-based tools, including continually testing AI tools to ensure safe and appropriate use and contributing to central monitoring systems to identify and mitigate risks across health care systems.

- **Focus on avoiding bias:** Because genAI results generally mirror past practices, they may produce biased results if these practices were biased. Also, if a genAI tool uses a dataset containing a biased sample (e.g., too few women or minoritized patients), the AI output will reflect these biases. Finally, because genAI results are likely to be persuasive, this increases the possibility that self-interested parties (e.g., corporations selling a relevant product) will try to insert themselves into the process to influence AI-generated recommendations. AI developers need to work to overcome and avoid these potential biases at all stages, including the creation of datasets, employing appropriate computational methods to correct for biases, and being transparent about and avoiding corporate influences on AI outputs.

- **Incorporate robust data protection:** Integrate advanced data security measures to protect sensitive patient data and address concerns about privacy and data ownership head-on. Developers must follow existing guidelines for cybersecurity across their product's life cycle and proactively identify potential risks with data, such as the emergence of new data types (e.g., audio recordings produced by digital scribes) and new threat actors and their tactics.[1]

- **Align with regulatory and ethical standards:** Stay informed about and compliant with health care regulations and guidelines to ensure AI tools meet safety, privacy, and ethical standards set by governing bodies. To further ensure that genAI is designed and implemented safely, developers must also consider other suggested or voluntary actions they can take, including the World Health Organization's *Regulatory Considerations on AI for Health*,[2] the National Academy of Medicine's Health Care Artificial Intelligence Code of Conduct,[3] and the Biden-Harris Administration's Executive Order 14110[4] and voluntary commitments[5].

References

[1] Farlow CS, Jump ML, Seeberger MS, Fitzgerald BJ. ANSI/AAMI SW96: Raising the Bar for Medical Device Security Risk Management. *Biomedical Instrumentation & Technology*. 2023;57(2):40-43.

[2] Regulatory Considerations on Artificial Intelligence for Health. World Health Organization; 2023. https://iris.who.int/handle/10665/373421

[3] 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/

[4] Sendak M. Health AI Partnership: An Innovation and Learning Network for Health AI Software. Duke Institute for Health Innovation Blog. December 23, 2021. https://dihi.org/health-ai-partnership-an-innovation-and-learning-network-to-facilitate-the-safe-effective-and-responsible-diffusion-of-health-ai-software-applied-to-health-care-delivery-settings/

[5] *ANSI/AAMI SW96:2023; Standard for Medical Device Security—Security Risk Management for Device Manufacturers*. Association for the Advancement of Medical Instrumentation; 2022. https://array.aami.org/doi/10.2345/9781570208621.ch1