Micromedex

Mastering AI in clinical decision support: Practical insights and evaluation strategies

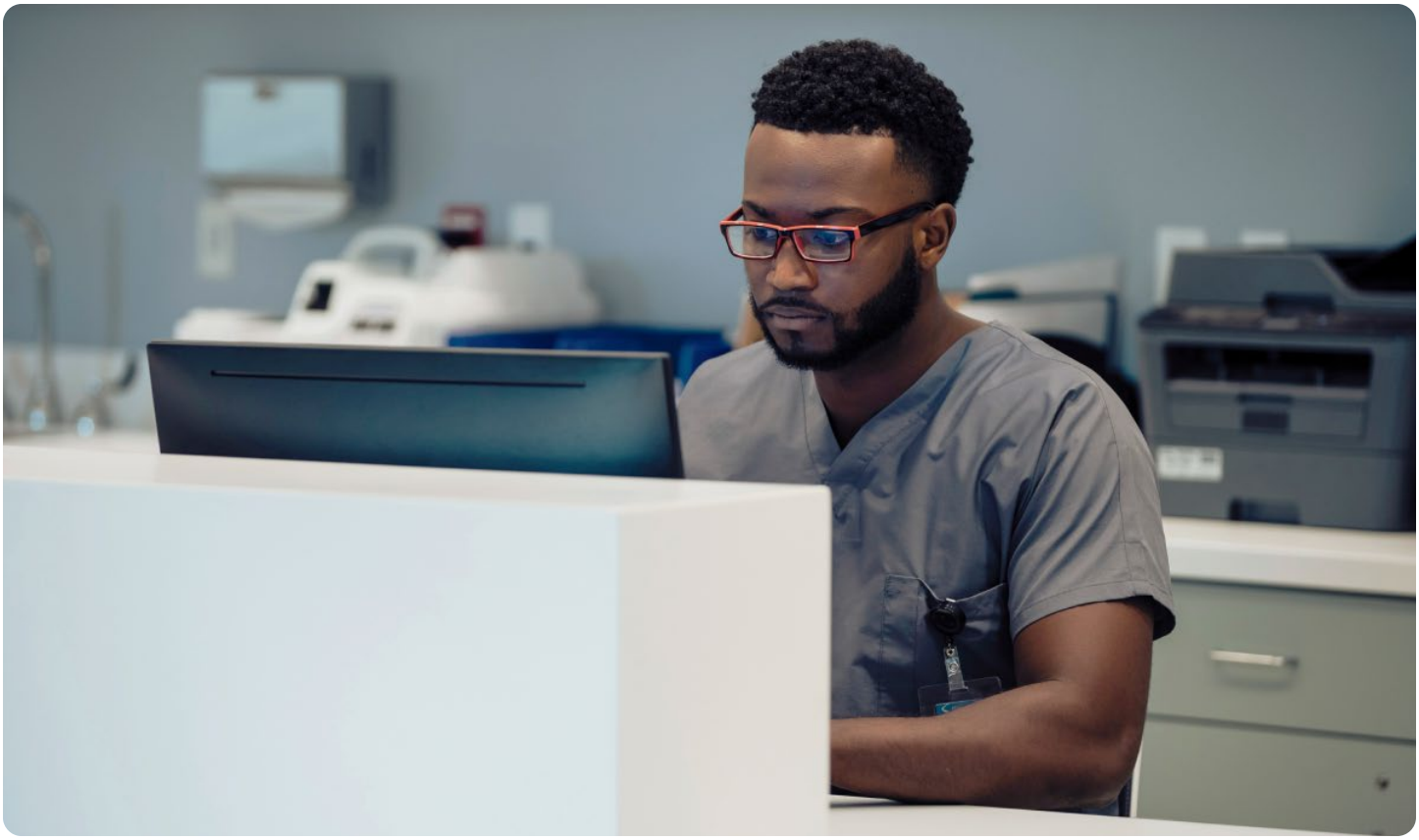
Artificial intelligence (AI) is already making its mark in healthcare, but keeping up with its rapid evolution can be challenging. This paper provides practical comparisons of AI technologies and a framework to help you critically assess AI tools to ensure they're the right fit for your team and organization, prioritizing safety and considering the potential risks.

Authors:
Angela Anderson, MSN, RN-BC, Senior Director of Content Optimization & Innovation at Micromedex
Brendan Bull, Principal Data Scientist at Merative
Petra Schultz, PharmD, Editor in Chief at Micromedex
Fabien Wecker, MSN, MA, RN, MHFA, International Clinical Success Manager at Micromedex

10-minute read

merative™

Content



Introduction

A growing demand for AI in healthcare

Artificial intelligence (AI) continues to build momentum across industries, and healthcare is seeing growing demand. In clinical decision support systems, AI offers real value—from streamlining workflows to enhancing "findability" of clinical evidence. By reducing administrative burdens around repetitive, time-consuming tasks, AI can enable clinicians to focus on delivering high-quality, patient-centered care.

Breaking down barriers in AI complexity

Despite promising benefits, there is still some hesitancy around AI adoption in clinical decision support. This caution isn't unfounded. The rapid development of AI technologies, combined with a lack of unified guidance makes it challenging for healthcare leaders to navigate the evolving landscape. The complexity of AI models also raises concerns about their safe use in clinical settings, as understanding how these models generate information can be difficult.[1] Tool immaturity, regulatory uncertainty, and limited AI literacy are all cited as key barriers to AI adoption in healthcare.[2]

So, how can clinicians and healthcare leaders navigate this complex landscape safely and effectively? The key to preparedness lies in understanding the technology behind AI tools and being equipped with practical strategies to critically evaluate their use in clinical practice.

Actionable insights to guide your decisions

We aim to bridge the gap by offering actionable insights and a practical framework designed specifically for clinicians, healthcare leaders, and healthcare informatics professionals.

This paper is designed to:

- Guide you through the technology behind AI tools in clinical decision support
- Highlight the risks and safety considerations that must be addressed
- Provide critical strategies to ensure informed decision-making
- Address how to combat the inappropriate use of AI and dispel common misconceptions around what constitutes safe and ethical implementation

Whether you're a clinician exploring AI-supported tools for clinical decision support or a healthcare leader seeking new ways to integrate AI into your organization, this paper provides practical guidance to help you assess these technologies. With clear strategies and frameworks, you can ensure AI is both effective and ethical for your healthcare practice.

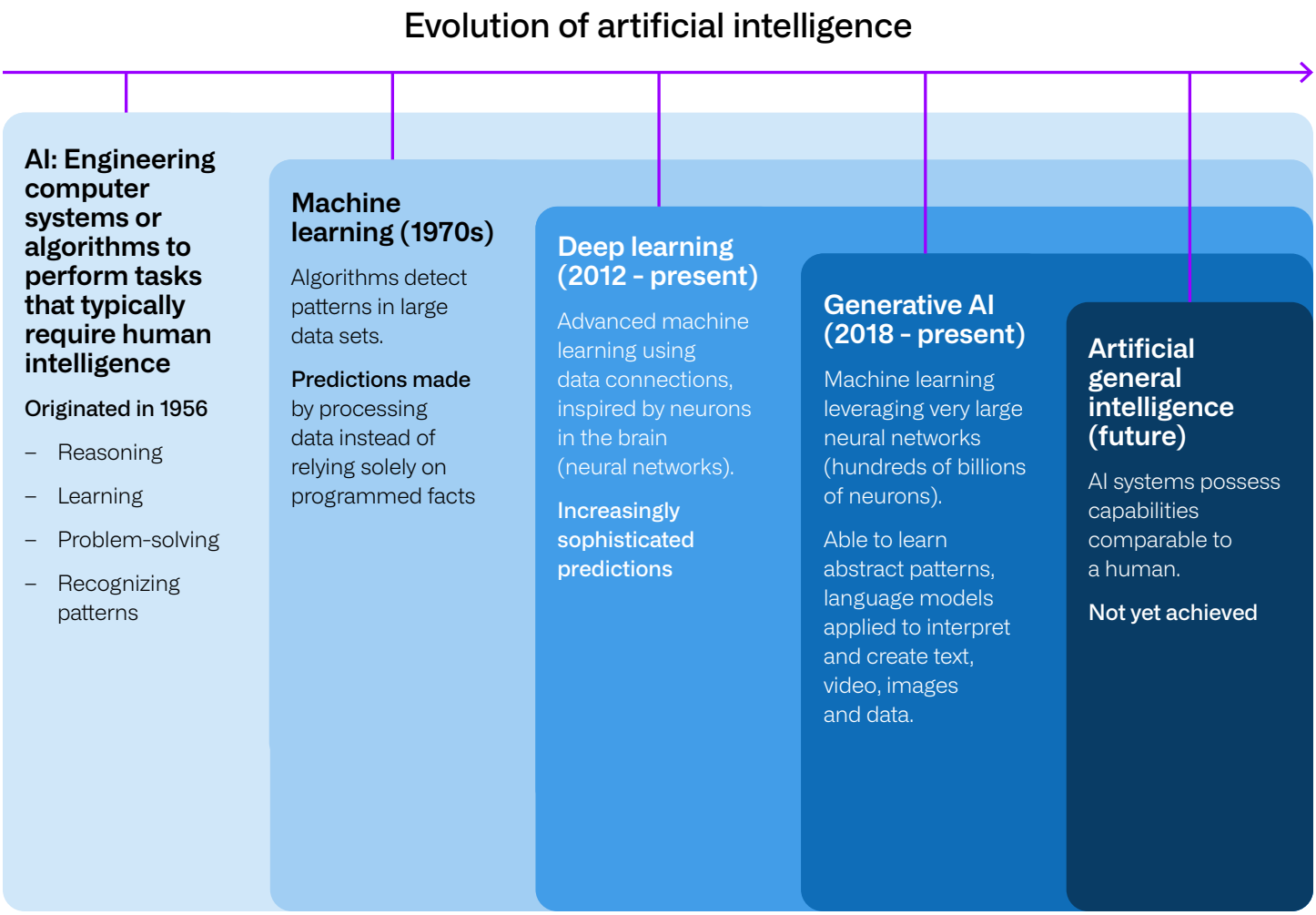
The evolution of artificial intelligence

Artificial intelligence (AI) has been around for many years and has evolved from an ambitious concept into a powerful tool with rapid advances, offering opportunities in healthcare to improve patient outcomes, streamline operations, and unlock new insights from complex datasets.

Figure 1 outlines key milestones in the development of AI, from its origins to advanced generative models and the vision of artificial general intelligence. These milestones highlight how AI has progressed from simple reasoning tasks to sophisticated models capable of learning and creating.[3]

Today, AI technologies are transforming clinical workflows, enabling healthcare operations, patient engagement, research and development, and much more. Understanding this evolution allows healthcare leaders, clinicians and clinical informatics specialists to be informed consumers as they harness AI to innovate and shape the future of medicine.

Figure 1:



Understanding the technology behind AI tools

As AI becomes more widely adopted in healthcare, the technology continues to advance, and misconceptions about the capabilities of AI tools can hinder their effective use. Among these technologies, Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG) systems have gained prominence due to their ability to handle complex data and language tasks. Each operates differently, with unique strengths and limitations. Understanding these differences is important for clinicians, healthcare leaders, and informatics specialists aiming to integrate AI into their work and organizations effectively.

Large Language Models (LLMs)

Key features

LLMs are stand-alone generative AI models designed to respond to questions or prompts with fluid, human-like communication. These systems operate by analyzing statistical patterns in their training data to predict and assemble responses. While highly sophisticated in their language output, LLMs do not critically appraise the validity of the information they use to create a response. Examples of LLMs include ChatGPT and Claude, among others.

Training Process

LLMs are trained using billions of data points, often referred to as a "black box" due to the vast, unverified scope of information involved. It is widely assumed that these training datasets include diverse sources from across the internet. However, this training lacks transparency, leaving gaps in the understanding of specific input data. Figure 2 details the benefits and limitations of LLMs.

Figure 2:

Benefits of LLMs	Limitations of LLMs
Excellent for idea generation and drafting nonclinical content	Has no knowledge of information or events arising after the date of their last training. May be unaware and unable to process new studies, findings, or recent developments
Proficient at summarizing and organizing text succinctly	Training data can contain inaccuracies or biases from unverified sources, such as social media posts or personal blogs
Useful as an editorial assistant or a "word calculator," helping refine written communication	Lack reliable citations for their outputs and struggle with producing precise calculations
Built for general-purpose applications, making them adaptable to a variety of needs	Parrot information without critically synthesizing it, increasing the risk of "hallucination" (the generation of false information)
	Perform inconsistently in non-English languages, creating the illusion of fluency when training data is primarily in English
	Not specifically trained or fine-tuned for clinical or healthcare-related tasks, limiting their utility in specialized scenarios

Retrieval-Augmented Generation (RAG) systems

Key features

RAG systems take a distinct approach by pairing AI LLMs with search capabilities. These systems are designed to retrieve and summarize information from a designated dataset in real time. With this method, RAG systems prioritize accuracy and currency by using a controlled corpus as their information source. The models used as components of RAG are typically smaller and specialized, each assigned a specific role to maintain precision and focus. They can be adapted to accommodate specific types of content. RAG systems require ongoing quality and performance monitoring. Figure 3 details the benefits and limitations of RAG systems.

Figure 3:

Benefits of RAG Systems	Limitations of RAG Systems
Responses are sourced directly from a defined information corpus, ensuring reliability and relevance	The accuracy and quality of responses depend entirely on the integrity and composition of the underlying content
Answers are generated in real time, reflecting the most current information available in the source data	Requires ongoing monitoring and maintenance to ensure optimal performance and data relevancy
The system cites its sources, providing transparency and traceability to its output	
Offers customization by adapting to accommodate specific types of content or datasets.	
Reduces the risk of hallucinations or biased outputs by answering only when validated data is available	

Choosing the right tool

When integrating AI technologies into healthcare, it's essential to match the tool to the task. LLMs may serve as efficient assistants for creative and administrative roles, while RAG systems, with their precision and transparency, can fill the gaps for clinical and evidence-backed tasks. Both systems have transformative potential, but their impact depends on informed usage and an understanding of their inherent limitations. Figure 4 clarifies the differences between RAG models and LLMs.

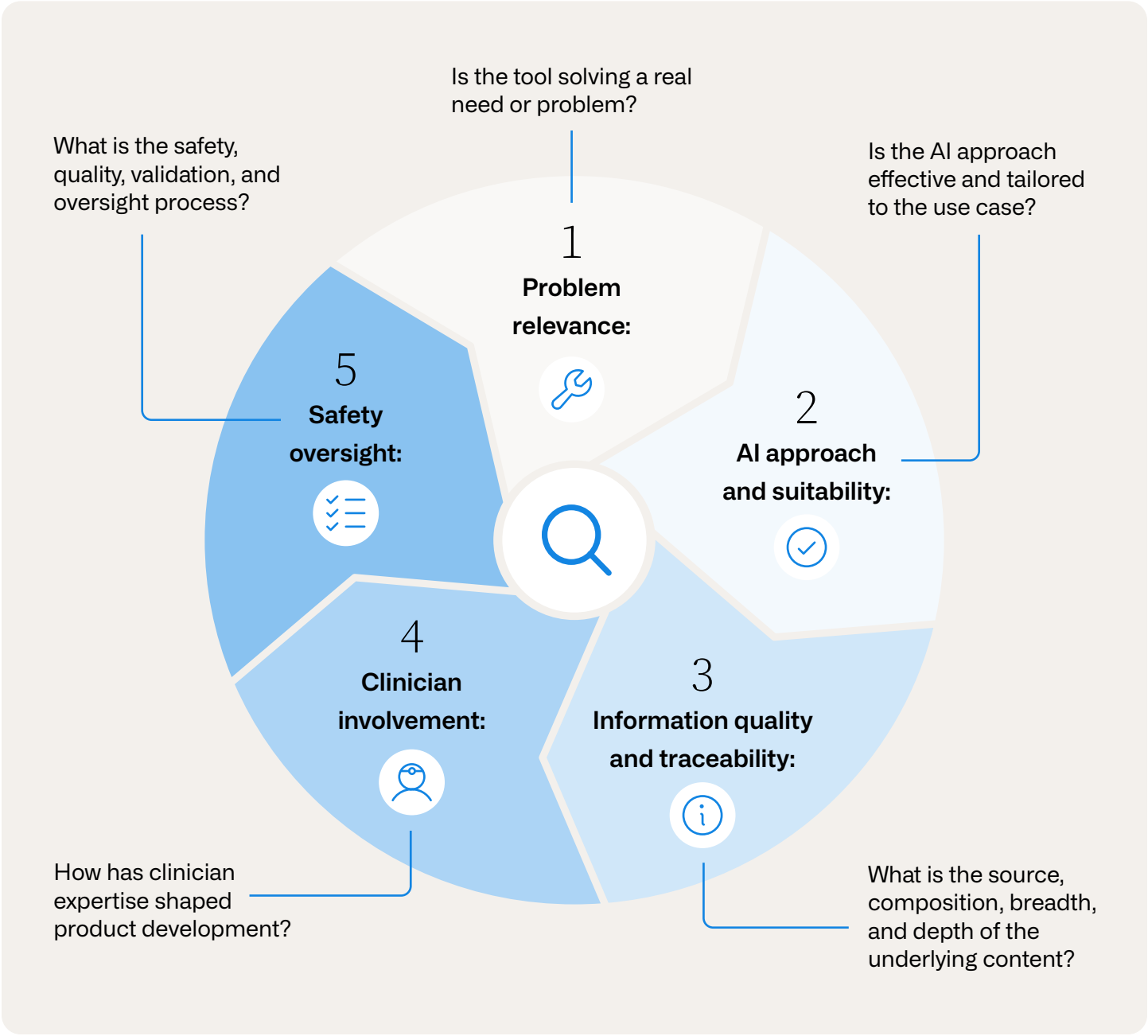
Figure 4: RAG systems vs. LLM

Theme	RAG	LLM
Source and quality of information	– Underlying corpus is its sole information source	– Training data is its information source and may contain unreliable or biased information (Reddit, blogs, social media)
Access to current information	– Responses are generated in real time; has access to the most current content in its corpus	– Has no knowledge of information published after the last training date (may be out of date)
Citation transparency	– Cites its sources	– Cannot reliably cite its sources
Risk of error	– Significantly lower risk of hallucination or bias – Tasked with answering only if it can find the information in the source data	– Sometimes puts together false information in its effort to create a response (hallucination)

Key tips for evaluating AI tools for clinical decision support

Not all AI tools are built for purpose, making it important for organizations to critically assess and evaluate these tools to ensure they align with their goals and standards. We have built a **practical framework of 5 key factors** to consider as you evaluate AI tools for clinical practice in your team or organization:

Figure 5:



1. Problem relevance

Is the tool solving a real need or problem?

The application of AI must address real, tangible challenges to be meaningful. Healthcare and IT leaders need to ask whether the AI tool is specifically designed to solve a real need or problem, such as speeding up access to information, accelerating decision-making processes, or reducing administrative burdens. Only by targeting real pain points will AI tools empower healthcare teams to offer enhanced care and foster better patient outcomes.

2. AI approach and suitability

Is the AI approach effective and tailored to the use case?

When evaluating tools for clinical practice, it's essential to understand how the AI is used, whether it's been designed for healthcare-specific needs, and if AI is indeed truly the best solution for the task at hand. General AI models, like standalone LLMs, aren't ideal for clinical use without refinement. To be effective for clinical decision support, LLMs must be paired with a validated content set that ensures reliable information, fine-tuned appropriately for the specific types of questions that clinicians ask, and equipped with guardrails to handle questions the AI may struggle to answer accurately.

General-purpose LLMs can be helpful for low-risk, administrative tasks. **RAG technology is a safer alternative for clinical scenarios that require precision, such as calculating drug dosages or determining intravenous (IV) compatibility.** Focused attention and specialized testing can decrease risk of errors, such as missing critical steps in a procedure, or unintentional combining of drug indications. Always evaluate the tool's design, purpose, and testing process to ensure it meets the specific needs of your team or organization.

3. Information quality and traceability:

What is the source, composition, breadth, and depth of the underlying content?

The information AI tools generate is highly dependent on its underlying knowledge source. For clinical AI tools to be useful, accurate, and trustworthy, special attention must be given to the content they search or are trained with. Consumers should examine the depth, breadth, and composition of the content being leveraged. Is it expert-curated? What is the update frequency? Is it full text or is the tool relying on abstracts/partial information? How are changes, errata, or corrections handled? Rigorous content data curation rooted in peer-reviewed research and clinical guidelines ensures credibility.

Additionally, transparent access to the specific text being cited facilitates explainable AI (XAI). It allows clinicians and stakeholders to fact check AI-generated answers with the original sources ensuring results are reliable and aligned with real-world medical practices.

4. Clinician Involvement:

How has clinician expertise shaped product development?

Are there clinicians on the product development team? In what capacity? A partnership between clinicians and data scientists across the development and deployment stages of AI tools is non-negotiable. Clinicians need to actively contribute their expertise to shape both algorithms and workflows, ensuring tools align with the realities of patient care. The language of healthcare is nuanced. Beyond development, clinicians should support ongoing evaluation by testing the tools in diverse clinical scenarios, reinforcing their practicality and reliability. This partnership ensures that AI supports and enhances human judgment rather than attempting to replace it.

5. Safety oversight:

What is the safety, quality, validation, and oversight process?

Safety oversight is a critical factor for any clinical AI tool. Each stage of development should incorporate validation processes to minimize risks and maximize accuracy. Regular performance assessments, both automated and clinician expert-led, are key to ensuring quality does not falter over time or with changes to the system. Ask vendors about their initial and ongoing safety and quality program. A deliberate focus on safety ensures the use of AI contributes positively to patient outcomes while maintaining the integrity and quality expected in clinical environments. This is especially important given current gaps in defined standards or thresholds related to use of AI for non-medical device decision support solutions.

Conclusion

While AI is not a magic bullet, its potential to enhance clinical decision-making is undeniable when implemented thoughtfully and responsibly. However, in a rapidly evolving technological landscape, clinicians and healthcare leaders must approach AI tools critically and with a well-informed perspective to prevent inappropriate use and ensure ethical, safe practices.

By asking the right questions, healthcare teams can ensure that the AI tools they adopt leverage high-quality data sources, undergo robust clinical validation, and adhere to strict safety oversight processes. When implemented appropriately, AI has the potential to streamline workflows in clinical decision support systems, helping clinicians access the right information, faster.

Ultimately, AI should be seen as a tool that complements—not replaces—the expertise of healthcare professionals. By maintaining this balance, healthcare organizations can create models of care that meet the needs of clinicians while ensuring safety and trust for the patients they serve.

What's next?

Learn more about the Micromedex approach to AI at merative.com/clinical-decision-support/ai

Got 15 minutes? Watch the webinar for focused insights:

Access the webinar on-demand to explore how AI can unlock greater value from clinical evidence for a deeper understanding of its potential in daily clinical practice.

Access the webinar →

About Micromedex

Micromedex by Merative is trusted by healthcare professionals in over 80 countries to provide award-winning clinical decision support solutions for drug and disease information and patient education. Micromedex was named Best in KLAS 2023 (clinical decision support: point of care clinical reference), is recognized as a CMS compendium, and has been recognized for its robust editorial process.

Learn more at merative.com/micromedex

About Merative

Merative provides data, analytics, and software for healthcare and government social services. With focused innovation and deep expertise, Merative works with providers, employers, health plans, governments, and life sciences companies to drive real progress. Merative helps clients orient information and insights around the people they serve to improve decision-making and performance.

Learn more at merative.com

References

1. Alattal, D., Azar, A. K., Myles, P., Branson, R., Abdulhussein, H., & Tucker, A (2025). *Integrating Explainable AI in Medical Devices: Technical, Clinical and Regulatory Insights and Recommendations.* arXiv: 2505.06620. Retrieved from https://arxiv.org/pdf/2505.06620
2. Poon, E. G., Lemak, C. H., Rojas, J. C., Guptill, J., & Classen, D. (2025). *Adoption of artificial intelligence in healthcare: survey of health system priorities, successes, and challenges.* Journal of the American Medical Informatics Association. Retrieved from https://academic.oup.com/jamia/advance-article/doi/10.1093/jamia/ocaf065/8125015
3. McKinsey & Company. (2024). *What is AI (artificial intelligence)?* Retrieved from https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-ai

Produced in the United States of America
June 2025

MDX-9201861539 Rev 1.0