GUIDE

WHAT TO KNOW, WHAT TO ASK, WHAT TO SAY: A GUIDE FOR HEALTHCARE PRACTITIONERS AND DECISION-MAKERS

Artificial intelligence (AI) refers to computer systems that can perform tasks like problem-solving and decision-making without being programmed step by step for each one. AI offers new tools that can support patient care and administrative tasks. As medical professionals, understanding the implications of these technologies is crucial for maintaining ethical standards, ensuring patient safety, and upholding trust.

This guide is designed for healthcare practitioners and decision-makers navigating the integration of AI into health care. It provides recommendations grounded in a rights-respecting framework to make sure the technology enhances the patient-caregiver relationship.

What You Need to Know

Research has identified four foundational pillars for AI use in a medical setting:

KEY PRINCIPLES OF ETHICAL AI IN HEALTH CARE:

1. Respect for Autonomy and Consent: Patients have the right to know and consent if AI is involved in decisions made about their care.
2. Human-centred: AI can be a part of decision-making, but never a replacement for human judgment and interaction.
3. Trust and Transparency: Open communication about AI risks, benefits, and limitations is essential.
4. Data Privacy and Protection: Strict safeguards over patient health information must be maintained.

UNDERSTANDING THE LANDSCAPE: CONCERNS WITH AI USE IN HEALTH CARE

While AI offers promise for efficiency and more patient-focused care, medical professionals must be aware of concerns regarding:

1. Privacy and Data Security

AI systems rely on large amounts of data. They are initially trained on very large data sets, but because many also continue to train themselves while they're being used, they may also collect data on the people who use them and on whom they're used.

Make sure you are familiar with federal, provincial and/or territorial legislation that relates to data collection and use, particularly in the context of patient or health care data. Your professional organization and/or your institution may also have policies that you should be aware of.

Even if this data is not shared or sold by the organization collecting it, there are always risks from hacking and security breaches. Patient data is highly personal, especially sensitive information like mental health or genetic data, and should never be used without clear consent. Youth in particular are sensitive about their health data being used without permission and about blurring the distinction between anonymous and identifiable data.

2. Accuracy and Bias

AI models learn from past data, and if this data reflects societal biases related to categories such as gender or race the AI can perpetuate stereotypes and lead to unfair or inaccurate recommendations and/or treatment.

AI can sometimes provide incorrect but confident answers or cite data and sources that don't exist. Patients tend to be less willing to forgive AI errors compared to human errors.

3. Loss of Human Connection and Oversight

Patients strongly believe AI should support, not replace, medical professionals. They want to maintain human interaction and place primary responsibility for decisions on health care workers.

AI should only ever be used as a supplement to care by medical professionals. Be clear when and how AI tools are being used and at what points in the process a human is "in the loop."

AI-drafted communications may seem impersonal or unsatisfying to patients. Avoid using them for personal or important communications and, when using them, always provide a way to reach an actual person. AI should never be used for sensitive communications, such as difficult diagnoses or end-of-life discussions.

Issues Relating Specifically to Youth in Healthcare

Youth have distinct ethical, social, and practical concerns regarding the integration of AI into their medical care. The issues center on the preservation of their autonomy, the security of their long-term digital footprint, and the maintenance of human empathy in clinical settings.

Research with youth has shown they generally have these opinions about the use of AI in healthcare:

- Autonomy and Consent: Youth strongly believe that age should not be a barrier to consent, and that even young children have a right to privacy and should be consulted directly about their medical records and treatment paths, rather than having parents or doctors make all the decisions.
- Granular Control of Data: Young people advocate for "unbundled" and "just-in-time" consent. They want the ability to agree to specific uses of their data (e.g., using it for their own diagnosis) while potentially refusing others (e.g., using it to train future AI models) at the exact moment the data is needed.
- The "Right to Be Forgotten": A particular concern for youth is the permanence of digital data. They worry that medical data profiles could influence their future education or employment opportunities. Consequently, they want a clear "right to be forgotten," which allows them to withdraw or erase their personal data once it is no longer needed for their immediate care.
- Fear of De-identification Failure: Youth recognize that medical data, particularly genetic information, is highly sensitive and difficult to truly anonymize. They are concerned that a security breach could not only "out" their own health status but also affect genetically connected family members.

- Loss of Human Connection: There is a significant fear that AI will replace the human interaction and empathy provided by medical staff. Youth describe the idea of being treated solely by a computer as "weird" and "treating patients like numbers." They insist that AI must remain a supplementary tool while medical staff maintain final responsibility for care.
- Algorithmic Bias and Fairness: Young patients are highly aware of the potential for AI to perpetuate societal stereotypes. They express concern that if an AI is trained on non-representative data (e.g., mostly urban or non-racialized patients), it may lead to inaccurate care for Indigenous, northern, or marginalized youth.
- Invasive Data Collection: Youth find the use of non-traditional health data—such as information from social media (mood, eating habits) or wearable devices (heart rate)—to be an invasion of privacy if used by healthcare practitioners without explicit, separate permission.

What you Need to Ask: Essential Questions for Medical Professionals

BEFORE USING AN AI TOOL:

- What are the specific functions of this AI tool, and how does it assist in patient care? Is there a good reason to think it will significantly improve patient care?
- How does this AI tool work, and what is my role in overseeing its suggestions or outputs? Understanding how an AI was trained and maintaining human oversight is critical. Are you confident that you could explain how the tool works to a patient (or their caregiver)?
- What are the known benefits and potential risks of using this tool in any specific patient population, especially concerning algorithmic bias in subgroups (e.g., race, gender, age)?
- How much training will it take for me, others in my institution, to use this tool effectively? (Consider both initial and follow-up training.)
- How is the tool's performance monitored, and what are the mechanisms for detecting and correcting errors?
- How can I report problems and errors, and to whom?
- What data was this AI tool trained on, and how are patient privacy and data security ensured throughout its operation?
- Can the tool run fully locally (on a single device or local network) or does it have to connect to an outside server or cloud service?
- What patient data does the tool use? How will it be kept secure during and after using the tool?

WHEN YOUR INSTITUTION IMPLEMENTS AN AI TOOL:

What is the organization's policy regarding patient notification or consent for this AI tool? Organizations should have clear policies about whether to notify patients, seek consent, or neither, based on risk and patient agency.

- How will the organization ensure the safe, ethical, and responsible use of AI tools, including ongoing governance processes?
- What training and resources are provided to help understand the AI tool's workings, limitations, and potential biases? Clinicians need to understand the AI programs they use.

- What channels are available to question the use or outputs of the tool, or decisions influenced by it, if concerns arise? Is there a process to allow patients to opt out of the tool?
- Have special concerns for particularly vulnerable groups, such as youth, been taken into consideration?

WHAT TO TELL PATIENTS (OR THEIR PARENTS/ CAREGIVERS) WHEN USING AN AI TOOL IN THEIR CARE:

Clearly state that an AI tool is being used and explain its specific function (e.g., to help interpret images, draft communications, suggest treatment plans).

- Provide a basic description of how the AI tool works and emphasize your role in reviewing and validating its outputs. Reassure them that decisions remain under human control.
- Explain why the organization believes using this tool improves care (e.g., efficiency, accuracy).
- Offer a basic overview of how the organization monitors the tool's performance, including efforts to address differential performance across patient groups.
- Inform them of any choices they have regarding the use of the tool.
- Provide information about how their data will be collected, used, and protected, ensuring transparency regarding privacy measures.
- Tell them about any channels that are available for questioning or filing complaints about the tool or how it is used.

By proactively addressing these concerns and engaging in thoughtful communication, medical professionals can foster a safe, trustworthy, and patient-centered environment as AI continues to evolve in healthcare.

Implementing these Practices Before, During and After Clinical Care

BEFORE USING AN AI TOOL (PREPARATION):

Ensure Data Protection: If possible, choose or develop in-house AI tools over third-party platforms to keep data within a closed network. Prioritize tools that work locally rather than needing to connect to an online server. Make sure that any data collected will be protected and that breach notification protocols are in place.

Audit for Bias: Check if the tool was trained on diverse datasets to prevent algorithmic bias (e.g., ensuring it works for Indigenous or northern communities).

Evaluate Improvements: Ask if there is a clear reason to believe the tool will significantly improve care before implementation.

Prepare Literacy Materials: Have patient-friendly, non-technical resources ready to explain AI to patients.

WHEN USING AN AI TOOL (IMPLEMENTATION):

Obtain "Unbundled" Consent: If possible, allow patients to consent to specific data uses rather than a single "all-or-nothing" agreement.

Apply "Just-in-Time" Consent: If possible, re-obtain consent if the way the patient's data is being used changes during treatment.

Maintain Oversight: Ensure clinicians review all AI-generated notes or diagnostic suggestions; never allow AI to make a final decision without a human in the loop.

Prioritize Human Interaction: Use AI to speed up administrative tasks only, ensuring the "human connection" remains the focus of the visit.

Issue Bias Warnings: If a tool is known to be less accurate for certain groups, instruct clinical staff to inform the patient and use extra human scrutiny.

AFTER USING AN AI TOOL (FOLLOW-UP):

Honour the "Right to be Forgotten": Provide a simple process for patients to withdraw or erase their data if they change their mind.

Provide an Audit Trail: Allow patients to request a record of every instanceDocument all uses of AI in a patient's health record so they can understand how it was used in their care.

Monitor and Report: Develop processes for clinical staff to report instances of AI "hallucinations" or errors.

Secure Storage: Ensure all resulting data is stored within the organization's secure infrastructure and that breach notification protocols are in place.

In collaboration with the McCradden Lab | Funded by the Canadian Institutes of Health Research

