

Artificial Intelligence in Patient Clinician Decision Making

Posted on April 5, 2025

View this essay online:

<https://academic-master.com/artificial-intelligence-in-patient-clinician-decision-making/>

Introduction

Artificial intelligence is becoming part of clinical decision-making through tools that estimate risk, interpret images, summarize records, support diagnosis, recommend treatments, and help patients understand options. The more important question is whether AI can improve the quality of decisions made jointly by patients and healthcare professionals without weakening clinical judgment, informed consent, trust, or patient autonomy. Shared decision-making requires more than selecting the statistically most likely option. A good clinical decision must take account of evidence, the patient's goals, uncertainty, competing risks, practical constraints, and the clinician's knowledge of the individual situation. Recent research suggests that AI-enabled decision aids can support this process by providing personalized predictions and clearer information, but patients and clinicians also identify concerns about bias, accuracy, transparency, over-treatment, and the possibility that algorithms may become influential without being sufficiently understood (Hassan et al., 2024). The most defensible role for AI is therefore decision support rather than decision authority: a tool that strengthens the patient-clinician conversation while leaving responsibility, interpretation, and value-sensitive choices within a human relationship.

AI in Clinical Decisions

Clinical decisions often require combining large amounts of information under time pressure. A clinician may need to consider laboratory values, imaging, prior diagnoses, medications, family history, disease prevalence, treatment guidelines, and patient preferences at the same time. AI systems can identify patterns across these data and produce estimates that would be difficult to calculate manually. In radiology and pathology, machine-learning models may highlight abnormalities or prioritize studies for review. In chronic disease management, prediction models can estimate the likelihood of deterioration or readmission. In treatment planning, decision-support tools can compare a patient's characteristics with patterns observed in large datasets. Generative AI can also summarize records or explain complex information in more accessible language. These capabilities can reduce cognitive burden, support consistency, and draw attention to information that might otherwise be missed. Systematic reviews of decision aids show that well-designed tools can improve knowledge, support clearer understanding of options, and help patients participate more actively in decisions

(Park et al., 2024). AI adds another layer by making those aids more individualized, although the quality of the output still depends on the quality and relevance of the data on which the system was developed.

The most promising applications are those in which AI contributes information without pretending to replace clinical reasoning. A risk calculator, for example, may estimate a probability of complication, but the clinician must still determine whether the estimate applies to the patient in front of them and how it should influence care. A model trained in one hospital population may perform differently in another. A patient may also value outcomes differently from the assumptions built into a model. Hassan et al. (2024) found that patients often viewed AI-enabled decision aids as understandable and useful, and some reported greater ownership of decisions and better adherence to recommended care. Clinicians, however, raised questions about whether information was current and whether the tools could contribute to inappropriate over- or under-treatment. This difference in perspective demonstrates why clinical AI should not be treated as an automatic answer generator. Its strongest contribution is to make relevant evidence more visible and personalized so that clinicians and patients can discuss what that evidence means.

Patient Autonomy

Shared decision-making becomes especially important when several medically reasonable options exist. Cancer treatment, surgery, long-term medication, screening, and chronic disease management may involve trade-offs among survival, side effects, convenience, cost, recovery time, and quality of life. AI can support these conversations by presenting individualized risk estimates or by organizing information in ways that make alternatives easier to compare. A 2025 systematic review of AI in collaborative care planning concluded that AI can contribute to shared decision-making by improving personalization and access to decision-relevant information, while also emphasizing the importance of transparency and human oversight (Di Palma et al., 2025). Used appropriately, an AI tool could help a clinician explain how likely a complication is for a particular patient or show how predicted outcomes change under different treatment choices. The technology should not decide which outcome the patient ought to value. That distinction preserves the ethical role of informed consent: the patient needs understandable information, an opportunity to ask questions, and freedom to make a choice that reflects personal priorities rather than simply accept the recommendation with the highest predicted score.

Patient autonomy can be strengthened or weakened depending on how AI is designed and introduced. If a tool gives patients access to clearer information and helps them formulate questions before an appointment, it may reduce dependence on hurried explanations and improve participation. If the output is presented as an unquestionable recommendation, however, it can create a new form of authority that patients feel unable to challenge. Mohammadnejad et al. (2026) identified transparency, dialogue, human-centered design, and an appropriate balance between human and algorithmic authority as central issues in protecting autonomy. The risk is not limited to obvious

coercion. Subtle framing can influence decisions by emphasizing certain outcomes, hiding uncertainty, or presenting one option as algorithmically preferred. Clinicians therefore need to explain what an AI tool does, what information it uses, what its limitations are, and how much weight should be placed on the output. The patient should understand that a prediction is an estimate rather than a guarantee and that personal goals remain part of the decision.

Trust and Accountability

Trust is essential because clinicians are unlikely to use systems they regard as unreliable, and patients may be harmed if either clinicians or users trust a system too much. Jones, Thornton, and Wyatt (2023) argue that discussions of clinical AI need to distinguish trust from trustworthiness. A clinician may trust a tool because it is convenient or familiar even when there is insufficient evidence that it performs well in the relevant population. Conversely, a highly validated system may be underused because its recommendations are difficult to interpret or because staff do not understand how it fits into existing responsibilities. Trustworthiness therefore depends on evidence, governance, validation, transparency, and accountability rather than confidence alone. Clinicians need information about model performance, limitations, intended use, and circumstances in which human review is especially important. They also need clarity about responsibility when AI contributes to an adverse outcome. If a system recommends a course of action, professional responsibility does not disappear simply because the recommendation came from software.

Explainability should be understood practically rather than as a demand that every complex model reveal every mathematical step to every patient. The level of explanation should match the decision. A clinician deciding whether to rely on a high-risk diagnostic model may need evidence about validation, subgroup performance, false-positive and false-negative rates, and known failure modes. A patient may need a clear explanation of what the prediction means and which personal factors influenced it. The U.S. Food and Drug Administration's 2025 draft guidance for AI-enabled medical devices emphasizes lifecycle risk management, documentation, transparency, performance monitoring, and appropriate controls rather than assuming that a model is safe once it has been released (FDA, 2025). This lifecycle approach is important because clinical environments change. Disease prevalence, treatment practices, equipment, documentation patterns, and patient populations can shift over time, causing model performance to deteriorate. Trustworthy clinical AI therefore requires continuing evaluation rather than one-time approval.

Bias and Safety

AI systems learn from data created by healthcare systems, and those data reflect both medical reality and existing inequalities. If certain groups are underrepresented in training data, or if historical treatment patterns contain bias, an algorithm may reproduce or amplify those problems. A model can also appear accurate overall while performing poorly for smaller demographic or clinical subgroups.

This is especially concerning when AI influences diagnosis, triage, access to treatment, or predictions of future risk. Bias is not solved simply by removing sensitive variables such as race or sex, because other variables may act as proxies and because some clinical differences genuinely matter.

Developers and health systems need subgroup testing, representative data, clear documentation, independent evaluation, and mechanisms for reporting problems after deployment. The FDA's evolving approach to AI-enabled devices reflects the need for risk controls across design, validation, modification, and post-market use rather than focusing only on a single performance number (FDA, 2025).

Safety also depends on avoiding automation bias, the tendency to accept a computerized recommendation even when conflicting evidence is available. Clinicians may be particularly vulnerable when workloads are high or when a system has performed well in the past. The opposite problem can also occur: repeated false alarms can lead to alert fatigue and cause useful warnings to be ignored. Clinical AI should therefore be integrated into workflows in ways that make disagreement possible and encourage review when a recommendation does not fit the clinical picture. Human oversight is not merely a ceremonial final click. It requires sufficient time, expertise, and authority to question the system. This is one reason patient-clinician decision-making should remain relational. A clinician can recognize details that may not be represented in structured data, while a patient can provide information about symptoms, fears, responsibilities, treatment burden, and values that a model cannot reliably infer. AI is most useful when it expands the information available to that relationship rather than narrowing the decision to what is easiest to quantify.

Human Oversight

Healthcare organizations should evaluate AI tools according to the clinical problem they are meant to solve rather than adopting them simply because AI is available. Implementation should begin with a defined use case, suitable evidence of performance, testing in the local setting, and a clear plan for who reviews the output and what happens when the system fails. Patients and frontline clinicians should be involved in design and evaluation because technically accurate tools can still be unusable or disruptive. Training should explain not only how to operate the system but when not to rely on it. Organizations also need policies for privacy, data security, documentation, informed consent when appropriate, and ongoing monitoring. Measures of success should extend beyond algorithmic accuracy to include patient outcomes, workflow effects, equity, clinician burden, patient understanding, and whether the technology improves the quality of shared decisions.

The long-term goal should be a form of augmented clinical intelligence in which computational systems and human expertise contribute different strengths. AI can search, compare, calculate, summarize, and detect patterns at a scale that humans cannot match. Clinicians contribute contextual interpretation, ethical responsibility, communication, physical examination, practical judgment, and knowledge of the patient as a person. Patients contribute their lived experience, preferences, tolerance for risk, cultural values, and goals. A decision is strongest when these forms of

knowledge are combined rather than placed in competition. The rapid development of generative AI makes this principle even more important because fluent language can make uncertain or incorrect output appear convincing. Human review, evidence standards, and patient participation must therefore become more important as AI becomes more capable, not less.

Conclusion

Artificial intelligence can improve patient-clinician decision-making by organizing complex information, generating personalized predictions, supporting diagnosis, and making treatment choices easier to compare. AI also introduces risks involving bias, opacity, over-reliance, privacy, liability, and patient autonomy. These risks make it inappropriate to treat an algorithm as an independent clinical authority. The strongest model is one in which AI provides evidence and decision support while clinicians remain responsible for interpretation and patients remain active participants in choices about their own care. Trust should be based on demonstrated trustworthiness, including validation, transparency, subgroup performance, lifecycle monitoring, and clear accountability. As clinical AI continues to develop, success should not be measured simply by whether a model can equal or exceed human performance on a technical task. The more meaningful test is whether the technology helps patients and clinicians make safer, better informed, more equitable, and more personally appropriate decisions together.

References

- Di Palma, G., Scendoni, R., De Benedictis, A., et al. (2025). Leveraging artificial intelligence for collaborative care planning: Innovations and impacts in shared decision-making—A systematic review. *Open Medicine*, 20(1), 20251232. <https://doi.org/10.1515/med-2025-1232>
- Hassan, N., Slight, R., Bimpong, K., Bates, D. W., Weiland, D., Vellinga, A., Morgan, G., & Slight, S. P. (2024). Systematic review to understand users' perspectives on AI-enabled decision aids to inform shared decision making. *npj Digital Medicine*, 7, 332. <https://doi.org/10.1038/s41746-024-01326-y>
- Jones, C., Thornton, J., & Wyatt, J. C. (2023). Artificial intelligence and clinical decision support: Clinicians' perspectives on trust, trustworthiness, and liability. *Medical Law Review*, 31(4), 501–520. <https://doi.org/10.1093/medlaw/fwad013>
- Mohammadnejad, S., Raiesifar, A., Bazmi, S., & Ghonodi, F. (2026). Ethical challenges to patient autonomy in the era of artificial intelligence: A systematic review. *BMC Medical Informatics and Decision Making*, 26, 305. <https://doi.org/10.1186/s12911-026-03614-x>
- Park, M., Doan, T. T.-T., Jung, J., Giap, T.-T.-T., & Kim, J. (2024). Decision aids for promoting shared decision-making: A review of systematic reviews. *Nursing & Health Sciences*, 26(1), e13071. <https://doi.org/10.1111/nhs.13071>

U.S. Food and Drug Administration. (2025). *Artificial intelligence-enabled device software functions: Lifecycle management and marketing submission recommendations*. Draft guidance for industry and FDA staff.