Artificial Intelligence in Patient Clinician Decision Making

Posted on April 5, 2025

Introduction

Artificial intelligence is increasingly used to support patient–clinician decision-making by organizing data, identifying patterns, estimating risk, generating drafts, and retrieving evidence. Its appropriate role is decision support rather than unaccountable replacement of clinical judgment. A useful system can reduce cognitive and administrative burden, but a poorly governed system can automate bias, create false confidence, expose private information, and obscure responsibility.

This chapter explains how health professionals should evaluate and use AI in clinical work. It addresses informed consent, data quality, fairness, workflow, documentation, regulation, cybersecurity, monitoring, and the continuing importance of communication between patient and clinician. The central principle is that every consequential decision remains a human and institutional responsibility even when software contributes to the recommendation. (World Health Organization, 2021; National Institute of Standards and Technology, 2023)

What Counts as Clinical AI?

Clinical AI includes machine-learning prediction models, image-analysis systems, natural-language processing, digital triage, remote monitoring, automated documentation, and generative systems. Some tools are regulated medical devices, while others perform administrative or informational functions. The risk depends on intended use, the population, and the consequences of error rather than on the presence of a fashionable label.

A calculator using a fixed clinical formula may be software without machine learning. A model trained on thousands of records may change only when the developer issues an update, while another system may learn continuously. These differences affect validation, change control, and regulation.

Decision Support Rather Than Decision Authority

A prediction is not a diagnosis, and a ranked option is not a treatment order. AI estimates patterns from data and may be useful when those patterns are relevant to the individual patient. Clinicians must interpret output alongside symptoms, examination, history, values, and constraints.

The phrase “human in the loop” is insufficient when the human lacks time, information, training, or authority to disagree. Meaningful oversight requires understanding the system’s purpose and limitations, access to relevant evidence, and a workflow that permits review rather than automatic acceptance.

Potential Benefits

AI can support earlier recognition of deterioration, image interpretation, medication safety, scheduling, documentation, and matching patients to trials. It may identify subtle combinations that are difficult to notice in large records. Automated summaries can help clinicians review long histories, and translation or accessibility tools can improve communication.

Benefits should be demonstrated in the actual clinical environment. Improved accuracy on a benchmark does not guarantee better patient outcomes. The tool may create new tasks, alerts, or delays. Evaluation should include safety, equity, workload, patient experience, and total cost.

Clinical Prediction

Prediction models estimate outcomes such as readmission, sepsis, falls, deterioration, or treatment response. Their usefulness depends on discrimination, calibration, timing, and actionability. A model can rank patients accurately while consistently overestimating absolute risk, which may produce unnecessary intervention.

The prediction must arrive early enough for an effective action. Clinicians need guidance about what to do at each risk level and evidence that the response improves care. A high score without an intervention pathway becomes another alert.

Diagnostic Imaging

AI systems can identify or prioritize findings in radiology, pathology, dermatology, ophthalmology, and other image-intensive fields. They may function as a second reader, triage tool, measurement aid, or quality check.

Performance can differ by scanner, protocol, disease prevalence, image quality, and demographic group. A system validated at one hospital should not be assumed to work identically elsewhere. Clinicians remain responsible for integrating imaging with the rest of the clinical picture.

Generative AI

Generative AI can draft clinical notes, patient instructions, referral letters, discharge summaries, and answers to questions. It can also fabricate facts, misattribute sources, omit uncertainty, or produce unsafe recommendations. Fluent language should not be confused with verified knowledge.

Generated content requires review before use. High-risk tasks need approved sources, traceability, and stronger controls. Patient information should not be entered into public or unapproved systems. Organizations must decide whether vendor agreements permit data retention or model training.

Clinical Documentation

Ambient documentation systems can transcribe conversations and prepare draft notes. They may reduce typing and allow more eye contact, but errors can enter the record with an appearance of authority. The clinician should verify diagnoses, medications, history, consent, and plans.

Patients should know when audio or automated processing is used. The organization should define recording, storage, deletion, access, and vendor use. A generated note should preserve the patient's meaning rather than replace it with standard language that obscures uncertainty or disagreement.

Patient-Facing Systems

Chatbots and symptom tools can provide education, reminders, navigation, and self-management support. They should not imply that emergency or individualized care is available when it is not. Clear escalation is essential for chest pain, severe breathlessness, self-harm, pregnancy emergencies, and other urgent situations.

Users need information about limitations, privacy, and whether a clinician reviews the interaction. Systems should be tested for health literacy, language, disability access, and the risk of delaying care.

Shared Decision-Making

Shared decision-making combines evidence with the patient's goals, preferences, and circumstances. AI can help estimate individualized benefit and harm, but it cannot decide what trade-off matters most to the patient.

Clinicians should explain relevant uncertainty and how the tool contributed. The amount of technical detail should match the decision. Patients do not need an engineering lecture, but they should know when an automated recommendation materially influences care and whether alternatives exist.

Informed Consent

Not every background use of software requires a separate signature. Consent requirements depend on risk, law, institutional policy, and whether data or care are used in an unexpected way. Experimental or high-impact use may require explicit consent or research oversight.

Transparency should be meaningful. A generic notice that “AI may be used” does not explain purpose, data use, consequences, or alternatives. Patients should not lose necessary care merely because they ask questions about automation.

Data Quality

Models learn from data that may contain missing values, errors, inconsistent coding, historical inequity, and changes in practice. A model can reproduce documentation patterns rather than biological truth. Labels may be proxies, such as using cost to represent need.

Data quality assessment includes completeness, accuracy, representativeness, provenance, and relevance to intended use. Developers and hospitals should document how variables are created and which populations are underrepresented.

Bias and Fairness

Bias can enter through sampling, labels, measurement, access to care, clinical decisions, and deployment. A system may perform well overall while harming a subgroup. Fairness cannot be solved by removing race because other variables may act as proxies and because race can be relevant to detecting unequal performance.

Evaluation should report performance by clinically and socially relevant groups, with adequate sample sizes and privacy protection. When disparities appear, the response may involve data, thresholds, workflow, access, or abandoning the tool. Fairness is an ongoing governance responsibility.

Race and Clinical Algorithms

Race is a social and political category, not a simple biological measurement. Some historical algorithms adjusted results by race without sufficient justification, potentially affecting access to care. Each use of race requires a clear causal and clinical rationale.

Removing race from a model can also worsen equity if the system then ignores disparities produced by racism. The appropriate approach is transparent, evidence-based evaluation of variables and outcomes rather than automatic inclusion or exclusion.

Social Determinants of Health

Housing, income, transportation, food access, environment, and discrimination affect health. AI may identify social needs, but prediction does not create resources. Labeling a patient as high risk without offering support can increase surveillance without improving care.

Organizations should link screening or prediction with services and evaluate whether the tool improves access. Sensitive social data require clear purpose and protection.

Privacy

Health AI can process records, images, voice, location, genetics, and behavior. De-identification reduces but does not eliminate reidentification risk. Data should be minimized, protected, retained only as needed, and used according to law and reasonable patient expectation.

Secondary use for model development requires governance. Organizations should understand where data travel, which subcontractors have access, and whether information crosses jurisdictions. Patients need routes to exercise applicable rights.

Cybersecurity

AI expands the attack surface through data pipelines, application interfaces, vendors, and model files. Threats include ransomware, data poisoning, malicious inputs, credential theft, and extraction of sensitive information.

Security requires access control, encryption, logging, segmentation, patching, testing, backup, and incident response. A model should fail safely when inputs or services are unavailable. Clinical continuity plans are essential.

Explainability

Explainability can mean different things: understanding the model structure, identifying influential variables, presenting a patient-level reason, or giving a clinically meaningful justification. A technically detailed explanation may not be useful to a clinician or patient.

The required level depends on risk. High-stakes decisions need enough information to detect error, challenge output, and communicate reasons. Explanation should not be used to make an unreliable model appear trustworthy.

Uncertainty

AI output should communicate uncertainty rather than present false precision. Confidence scores, prediction intervals, or categories can help, but users need training in interpretation. A 20 percent risk does not mean the outcome will occur in one fifth of an individual.

Uncertainty also comes from changing populations, incomplete data, and disagreement among clinicians. Systems should allow “insufficient information” rather than forcing a recommendation in every case.

Automation Bias

Automation bias occurs when users over-rely on computer output or fail to notice errors because the system appears authoritative. The opposite problem is algorithm aversion, where users reject a useful system after seeing one error.

Interface and training influence both behaviors. The tool should not present recommendations in a way that discourages review. Monitoring should examine when clinicians override the model and whether overrides improve outcomes.

Alert Fatigue

Excessive alerts can reduce attention and increase routine dismissal. Adding an AI alert to an already crowded system may worsen safety. The threshold should reflect disease prevalence, consequences, and available response.

Organizations should measure alert volume, acceptance, timeliness, and outcomes. Low-value alerts should be removed or redesigned.

Workflow Integration

A model can be accurate yet fail because it arrives at the wrong time, to the wrong person, or without resources for action. Workflow design identifies who receives output, what action follows, and who documents the decision.

Frontline staff should participate in design and testing. Integration should reduce duplicate work and preserve communication. A pilot should examine normal operations and exceptional situations.

Human Factors

Human-factors engineering studies how people interact with systems under real conditions. Screen layout, terminology, timing, interruption, workload, and team roles affect performance.

Usability testing should include the intended users and representative scenarios. Training cannot compensate indefinitely for poor design. The safest control may be to redesign the interface or remove a feature.

Clinical Validation

Validation should proceed through stages. Technical validation tests whether the system performs the intended computation. Clinical validation examines whether output corresponds to clinically meaningful outcomes. Prospective validation evaluates performance in actual workflow.

External validation is essential before broad deployment. Randomized or quasi-experimental studies may be appropriate for evaluating effect on care, while observational monitoring can identify rare harm. Validation should compare the AI-supported pathway with current practice.

Clinical Trials of AI

Trials involving AI require reporting of intended use, model version, input data, workflow, human interaction, errors, and changes during the study. CONSORT-AI and SPIRIT-AI extensions provide guidance for trials and protocols involving AI interventions. (Liu et al., 2020; Cruz Rivera et al., 2020)

A trial should evaluate patient or process outcomes relevant to the claim. Improved model accuracy alone is not equivalent to improved health. Investigators should report subgroup effects and failures.

Regulation

Regulatory treatment depends on jurisdiction and function. In the United States, the Food and Drug Administration regulates certain software as medical devices and has issued guidance concerning clinical decision-support software. Not every tool is regulated in the same way. (U.S. Food and Drug Administration, 2022)

Organizations remain responsible for procurement and local use even when a product has regulatory authorization. Authorization does not guarantee benefit in every population or workflow. Changes to a model may require review.

Adaptive Models

An adaptive model changes after deployment. This creates a challenge because the validated system may not remain identical. Predetermined change-control plans can define which modifications are permitted and how they will be tested.

Many healthcare organizations may choose locked models for stability. Continuous learning should not occur invisibly in a high-risk environment. Every update needs versioning, monitoring, and rollback capability.

Procurement

Procurement should evaluate the clinical problem, evidence, data requirements, integration, security, accessibility, cost, and vendor stability. Demonstrations using ideal data are insufficient.

Contracts should address performance, audit rights, incident reporting, model changes, data ownership, subcontractors, indemnity, support, termination, and export of data. The organization needs a plan if the vendor ends the service.

Vendor Claims

Marketing may report high accuracy without the prevalence, comparator, threshold, or confidence interval. A strong purchasing process requests peer-reviewed or independently reviewable evidence and asks whether the study population resembles local patients.

Claims of reducing workload or cost should include implementation and review time. Institutions should conduct their own pilot and not rely only on vendor-selected testimonials.

Model Documentation

Documentation should describe intended use, users, inputs, outputs, training data, validation, limitations, known failure modes, and update history. Model cards or similar documents can support review, though sensitive intellectual property may require controlled access.

Clinicians need concise operational information, while governance teams need technical detail. Documentation should be updated when the system changes.

Monitoring After Deployment

Performance can drift as clinical practice, disease prevalence, equipment, coding, or population changes. Monitoring should include calibration, discrimination, missingness, subgroup performance, overrides, alerts, errors, and patient outcomes.

Monitoring requires thresholds and action. A dashboard that shows decline without authority to pause the tool is inadequate. Organizations need owners, review schedules, and incident procedures.

Incident Reporting

Staff should be able to report suspected AI errors without blame. Incidents may involve wrong recommendations, delayed care, privacy breaches, or interface confusion. Investigation should examine the complete sociotechnical system rather than assuming user error.

Serious events may require disclosure to regulators, vendors, patients, or safety bodies according to law and policy. Lessons should be shared while protecting privacy.

Audit Trails

Systems should record model version, inputs, output, time, user, and relevant actions. Audit trails support investigation and accountability. They should not become indiscriminate surveillance of clinicians.

Access to logs should be controlled, retention defined, and interpretations cautious. A recorded override does not prove that a clinician acted improperly.

Documentation of Clinical Decisions

Clinicians should document the reasoning appropriate to the clinical decision, including significant AI input when relevant. Copying a model score without interpretation is insufficient.

Documentation should identify disagreement and uncertainty honestly. The record should not imply that the software made the decision because responsibility remains with authorized professionals and the organization.

Liability and Accountability

Liability depends on law, contracts, professional standards, and facts. Potentially responsible parties include clinicians, hospitals, vendors, and developers. A disclaimer cannot eliminate all duties.

Governance should define who approves deployment, monitors performance, responds to incidents, and informs patients. Accountability must exist before harm occurs.

Ethics Committees and Governance Boards

High-impact AI should be reviewed by a multidisciplinary group including clinicians, patients, nurses, data scientists, security, privacy, ethics, legal, and operations. The group should have authority to approve, restrict, or stop use.

Review should consider benefit, harm, equity, autonomy, transparency, and alternatives. Governance should continue after launch rather than functioning as a one-time permission.

Patient and Public Participation

Patients can identify concerns that technical teams overlook, including stigmatizing language, confusing explanations, or unacceptable data use. Participation should include diverse communities and compensate expertise.

Consultation should influence decisions rather than merely endorse a completed plan. Public reporting can improve trust when it includes limitations and incidents, not only success stories.

Clinician Training

Clinicians need practical AI literacy: intended use, inputs, output interpretation, common failure modes, bias, privacy, and escalation. Training should use realistic cases and include the ability to challenge the model.

Training must be updated with system changes. It should not shift responsibility for flawed technology onto individual users.

Patient Communication

Communication about AI should be proportional to its role. A clinician might explain that software helped compare the image with patterns, while the final interpretation considered the patient’s

history and examination.

Patients should be invited to ask questions. If an alternative pathway without AI is reasonably available, its implications can be discussed. Communication should avoid both hype and fear.

Health Literacy

Risk estimates and technical explanations can confuse patients. Visual aids, absolute risks, plain language, and teach-back can improve understanding. The clinician should explain what the result changes and what remains uncertain.

Translation should be performed through validated language support, not assumed accurate because a generative system produces fluent text. Clinical interpreters remain important.

Accessibility

AI interfaces should support screen readers, keyboard access, captions, color alternatives, adjustable text, and different communication needs. Patient-facing systems should not require high digital skill or the newest device.

Accessibility should be tested with disabled users. A tool that improves efficiency for clinicians while excluding patients is not successful.

Low-Resource Settings

AI may expand expertise in settings with limited specialists, but it can also create dependency on internet, vendors, and foreign data. Tools should be validated locally and designed for available equipment, language, and referral pathways.

When a model identifies a condition for which no treatment is accessible, the ethical value requires careful assessment. Investment should not divert resources from basic services with stronger evidence.

Global Equity

Many models are developed using data from high-income settings. Exporting them without validation may worsen inequity. Local institutions should participate in research, governance, and benefit sharing.

Data extraction from lower-resource communities without reciprocal capacity or care improvement is

ethically problematic. Global standards should support transparency and local control.

Research Ethics

AI research using clinical data may qualify for waivers or secondary-use pathways, but legal permission does not settle every ethical question. Researchers should minimize risk, protect privacy, and consider community expectations.

Prospective studies need clear protocols, oversight, adverse-event reporting, and plans for incidental findings. Research systems should not enter routine care without appropriate transition and governance.

Reproducibility

AI research can be difficult to reproduce because data and code are proprietary or preprocessing is unclear. Researchers should report methods, cohorts, missing data, thresholds, and evaluation plans in sufficient detail.

Privacy may limit data sharing, but secure environments, synthetic data, model sharing, and independent validation can improve trust. Reproducibility is especially important for high-impact claims.

Publication Bias

Successful systems are more likely to be published than failures. Vendor-sponsored studies may select favorable endpoints. Journals and institutions should report negative and null results so others do not repeat harm.

Clinical registries and protocol publication can reduce selective reporting. Claims should be compared with study design and conflicts of interest.

Environmental Impact

Training and operating large models consume energy, water, and hardware. Healthcare organizations should consider whether a model’s benefit justifies its environmental cost and whether smaller systems can perform the task.

Environmental effects are part of public health and should enter procurement and lifecycle assessment. Efficiency can reduce both cost and impact.

Administrative Uses

AI can support coding, scheduling, inventory, and prior-authorization work. Administrative automation may free time but can also create denials or hidden barriers. The same fairness and appeal principles apply when an algorithm affects access.

Administrative systems should not be considered low risk merely because they do not diagnose. A scheduling model can worsen access for patients with disability or transport constraints.

Insurance and Utilization Management

Payers may use algorithms to identify fraud, manage care, or review coverage. Automated decisions can affect whether patients receive treatment. People need notice, understandable reasons, human review, and correction of inaccurate data.

Clinical recommendations should not be distorted by financial optimization that is hidden from the patient. Conflicts and objectives must be transparent.

Public Health

AI can support surveillance, outbreak detection, resource allocation, and communication. Public-health data may include people who did not directly choose participation, so proportionality and governance are important.

Models can reproduce testing and access patterns rather than true disease distribution. Community engagement and transparent use are essential.

Emergency Care

Emergency settings involve time pressure and incomplete data. AI may support triage or image prioritization, but false reassurance or overtriage can harm patients. Systems should be tested under crowding, downtime, and atypical presentations.

Clinicians need immediate access to the basis and limitations of output. The tool should not delay stabilizing treatment.

Primary Care

Primary care involves undifferentiated symptoms, prevention, chronic disease, and long relationships. AI can help summarize records and identify overdue care, but excessive prompts can disrupt conversation.

The system should support continuity and patient priorities rather than convert visits into checklist completion. Social context and multimorbidity often make single-disease recommendations inappropriate.

Specialty Care

Specialty systems may perform well on narrow tasks with high-quality data. Integration with broader care remains necessary. A cancer model must consider comorbidity, patient values, and treatment availability.

Specialists should avoid assuming that technical complexity guarantees clinical usefulness. Comparative evidence and external validation remain essential.

Mental Health

AI can support screening, appointment navigation, and between-visit monitoring. Mental-health data are highly sensitive, and language models may respond inadequately to crisis or delusion.

Patient-facing systems need clear crisis escalation, clinician oversight where promised, and careful claims. They should not present simulated empathy as a substitute for care when risk is high.

Pediatrics

Children differ physiologically and developmentally from adults, and models trained on adults may be unsafe. Consent, assent, parental access, and long-term data use require special attention.

Interfaces and explanations should be age appropriate. Systems should not permanently label children based on unstable developmental patterns.

Older Adults

Older adults may have multimorbidity, polypharmacy, sensory impairment, and goals that differ from standard disease targets. AI should avoid recommendations based on trials that excluded frail

patients.

Digital tools need alternatives for people without devices or confidence. Family involvement should respect the patient’s autonomy and privacy.

Genomics

AI can help interpret genetic variants and integrate genomic data, but many findings remain uncertain. Genetic information has implications for relatives and can create anxiety or discrimination concerns.

Results require qualified interpretation and appropriate counseling. Algorithms should not overstate pathogenicity or ancestry categories.

Precision Medicine

Precision medicine aims to tailor prevention or treatment using individual characteristics. AI can integrate complex data, but the term should not imply perfect prediction.

Highly personalized models may be difficult to validate and can overfit. The benefits must be compared with simpler clinical rules and population-level interventions.

Remote Monitoring

Wearables and home devices can detect trends and support chronic care. They also generate false alarms, burden patients, and exclude those without technology. Data quality can be affected by device placement, skin characteristics, movement, and adherence.

Monitoring programs need response capacity. Collecting continuous data without a team to act creates false reassurance and liability.

Robotics

Robotic systems can assist surgery, rehabilitation, transport, and pharmacy operations. AI may support perception or planning, but physical systems introduce injury and cybersecurity risk.

Training, maintenance, emergency stop, and accountability are essential. Robotic assistance should be evaluated through patient outcomes, not marketing claims about precision alone.

Medication Management

AI can identify interactions, suggest doses, and detect adherence patterns. Medication decisions depend on kidney function, age, pregnancy, comorbidity, and patient preference.

Recommendations should use current formularies and guidelines and be reviewed by qualified professionals. Alert fatigue is a major concern.

Clinical Guidelines

AI can retrieve and summarize guidelines, but guidelines vary and may be outdated or conflict. A system should identify source, date, population, and strength of evidence.

Guidelines support rather than replace judgment. Clinicians should document why a recommendation was adapted when the patient differs from the studied population.

Multimorbidity

Many clinical models focus on one disease, while patients have several conditions and treatments. Recommendations can conflict, increasing burden or harm.

Decision support should identify interactions and prioritize patient goals. More recommendations are not necessarily better care.

End-of-Life Care

Prediction of mortality or deterioration can support timely conversations, but it can also influence care through self-fulfilling expectations. Scores should not determine whether a person is offered treatment without individualized review.

Communication must be compassionate and recognize uncertainty. The patient’s values and advance directives remain central.

Resource Allocation

AI may be used to prioritize beds, tests, or outreach. Allocation criteria should be ethically justified and transparent. Historical utilization can reflect unequal access and is a poor proxy for need.

Governance should examine who benefits and who is delayed. Appeals and emergency discretion are

necessary.

Learning Health Systems

A learning health system uses routine data to improve care continuously. AI can support this cycle, but the boundary between quality improvement and research may require oversight.

Learning should be transparent, secure, and connected to measurable changes. Patients and staff should receive feedback about how data improve care.

Implementation Science

Implementation science examines adoption, fidelity, context, sustainability, and outcomes. AI projects often fail because attention focuses on the model rather than organizational change.

Frameworks can identify leadership, resources, workflow, user beliefs, and external policy. Successful implementation includes the ability to stop ineffective technology.

Decommissioning

Systems should be retired when they are unsupported, unsafe, redundant, or no longer useful. Decommissioning includes data export, record retention, communication, and removal from workflow.

Organizations should avoid “zombie” models that continue influencing care after monitoring ends. Exit planning begins during procurement.

A Practical Evaluation Checklist

Before deployment, an organization should ask:

- What clinical problem is being solved?
- Is AI better than a simpler alternative?
- Who is the intended user and population?
- What evidence supports safety and benefit?
- How does performance vary across groups?
- What data are required and where do they go?
- How will output enter workflow?
- Who can override or stop the system?
- How will patients be informed?
- How will performance and incidents be monitored?

- What is the plan for updates and retirement?

Case Example: Deterioration Alert

Consider a model predicting deterioration on a hospital ward. The team should validate it locally, define the threshold, identify who receives the alert, and establish a response. Nurses and physicians should help design the workflow.

Evaluation includes timeliness, false alarms, workload, escalation, intensive-care transfer, mortality, and subgroup effects. If alerts are ignored or resources are unavailable, the project has not succeeded even when the model is statistically accurate.

Case Example: Draft Patient Message

A generative system may draft a message explaining laboratory results. The clinician verifies the facts, tone, urgency, and next steps. The system should use approved information and avoid diagnosing from one result.

The patient should receive contact information and understand whether the message was automated. Sensitive or serious findings may require direct conversation rather than a portal draft.

Case Example: Trial Matching

Natural-language processing can compare patient records with clinical-trial eligibility. It may identify opportunities that staff miss. The output remains a preliminary match because criteria can be ambiguous and records incomplete.

Research staff should verify eligibility and explain that matching does not guarantee enrollment or benefit. Privacy and data-sharing rules apply.

Future Directions

Multimodal systems will increasingly combine text, image, signal, and genomic data. Smaller local models and privacy-preserving methods may reduce some risks. Regulatory and reporting standards will continue to develop.

The most important innovation may be better evaluation rather than larger models. Healthcare needs systems that demonstrate benefit, integrate safely, and remain accountable.

Conclusion

Artificial intelligence can support patient–clinician decision-making through prediction, image analysis, documentation, evidence retrieval, monitoring, and communication. Its value depends on data quality, validation, workflow, fairness, privacy, cybersecurity, transparency, and human oversight.

AI should not be treated as an independent decision-maker. Patients and clinicians remain participants in a relationship that includes values, context, uncertainty, and responsibility. Organizations must govern the entire lifecycle from procurement through retirement and be prepared to pause or remove systems that do not improve care.

The appropriate question is not whether AI is intelligent enough to replace medicine. It is whether a specific system, used by trained people in a defined setting, improves outcomes without creating unacceptable harm or inequity. (World Health Organization, 2021; National Institute of Standards and Technology, 2023)

References

World Health Organization. (2021). *Ethics and governance of artificial intelligence for health*.

National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*.

U.S. Food and Drug Administration. (2022). *Clinical decision support software: Guidance for industry and Food and Drug Administration staff*.

Liu, X., Cruz Rivera, S., Moher, D., Calvert, M. J., & Denniston, A. K. (2020). Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension. *Nature Medicine, 26*, 1364–1374.

Cruz Rivera, S., Liu, X., Chan, A.-W., Denniston, A. K., & Calvert, M. J. (2020). Guidelines for clinical trial protocols for interventions involving artificial intelligence: The SPIRIT-AI extension. *Nature Medicine, 26*, 1351–1363.

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine, 380*, 1347–1358.

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine, 25*, 44–56.