AI in Clinical Development: Beyond the Discovery Phase

On January 21-22, 2026, the *Duke Clinical Research Institute* convened a cross-disciplinary Think Tank of experts from academia, health systems, regulatory agencies, funders, industry sponsors, clinical research organizations, technology developers, and patient-facing organizations to examine the current state, limitations, and future role of artificial intelligence (AI) in Phase III and IV research, and post-market evidence generation.

Late-stage clinical research is facing strain, including rising costs, limited workforce capacity, increasing clinician burden, variability and expense associated with adjudication and core laboratories, and growing mistrust among patients and clinicians. At the same time, expectations for post-market safety surveillance, real-world evidence generation, and movement towards a continuous evidence generation and regulatory pathway have expanded.

This Think Tank focused on identifying where AI can revolutionize evidence generation through increased efficiency, reduced costs, improved access to and harmonization of data, as well as enable new data insights and trial endpoints. Consideration was also given to limitations and risks of adoption, and the institutional, regulatory, and operational challenges that currently constrain implementation. Across discussions, a consistent conclusion was that AI has demonstrated promise for both improving efficiencies and enabling new data sources and data analytics and potentially new endpoints within late-stage clinical research. However, implementation barriers, rather than technical limitations, now represent the dominant obstacle to effective AI adoption. Slow and fragmented governance processes, misaligned incentives and funding mechanisms, regulatory inspection anxiety, and insufficient operational readiness were repeatedly cited as more limiting than AI performance itself. A fundamental tension between the need for collaboration, through shared data, standards, and infrastructure, to enable AI at scale, and the need to protect proprietary assets that sustain competitive advantage and investment incentives was also highlighted.

KEY TAKEAWAYS AND THEMES:

Operational efficiency increases have been successful, but not at scale: The following applications were viewed as incremental but immediately actionable, with the greatest potential to relieve workforce strain without altering trial endpoints or causal inference and were consistently characterized as low-risk operational tools. These approaches have been applied across both Phase III and Phase IV research; in Phase III trials, they primarily support execution efficiency without affecting evidentiary standards, whereas in Phase IV and post-market studies they can be deployed more flexibly and at greater scale. Examples include:

- Large Language Model (LLM)-assisted Electronic Health Record (EHR) screening to support site-level patient identification and recruitment
- AI-supported protocol feasibility assessment and site optimization to reduce screen failure and startup delays
- Agentic AI to coordinate routine trial workflows
- Document drafting and summarization to reduce administrative burden
- Adjudication of selected endpoints
- Analysis of high-dimensional, multimodal data
- As continuous data streams from wearables and other always-on biosensors expand beyond feasible human review capacity, AI-enabled first-pass analyses represent a practical, low-risk approach to managing data and identifying early safety or efficacy signals

Methodological innovation: These approaches hold promise but require careful methodological justification and regulatory engagement (higher risk, selective use):

- Exploration of synthetic control arms and digital twins in carefully defined settings
- Potential to enable more continuous, adaptive evidence generation and eliminate traditional distinctions between Phase III and Phase IV evidence generation

Governance and oversight remain central determinants of adoption: Existing FDA guidance supports risk-based validation and governance of AI, emphasizing that validation should be proportional to context of use, rather than applying uniform standards across all applications. Failure to distinguish between low-, intermediate-, and high-risk cases was cited as a key driver of overvalidation, delayed implementation, and prolonged pilot phases, particularly for low-risk operational tools.

A practical risk-based spectrum emerged:

- **Low-risk applications** (e.g., document drafting, summarization, workflow routing), where lightweight assurance is appropriate.
- **Intermediate-risk applications** that inform trial operations or analyses, requiring focused validation of critical-to-quality elements.
- **High-risk applications** that directly affect endpoint determination, causal inference, or patient-facing decisions, which require rigorous analytic and clinical validation.

While human oversight remains necessary, reviewing every AI output is not feasible at scale. Scalable oversight models include risk-tiered escalation, uncertainty-based triage, and the use of supervisory AI agents that operate alongside primary models. These supervisory agents can be used synergistically to monitor model drift, assess output quality, and trigger escalation when predefined thresholds are exceeded. They may also bring additional complexity and system fragility and require focused research themselves.

Data limitations remain a foundational constraint on AI use in late-stage research: AI performance in Phase III-IV research is fundamentally limited by the availability, quality, and fitness of underlying data, rather than algorithmic capability alone. Key challenges center on three areas: 1) enabling data portability and interoperability across health systems to support multi-institutional use; 2) unlocking unstructured data at scale through LLM-enabled approaches; and 3) integrating data beyond the electronic medical record (EMR) and traditional healthcare system structures (e.g., wearables). Without addressing these areas, AI tools will be artificially limited from achieving the maximal possible benefit, regardless of model sophistication.

Education and alignment across the ecosystem to garner trust and establish guidelines: Effective AI adoption requires trust grounded in clear evidentiary standards with two trust domains; 1) trust among clinical and operational users, which depends on clear understanding of AI purpose, risk, and governance (e.g., appropriate differentiation between research operations, quality improvement, and clinical care to avoid over-regulation and unnecessary delays); and 2) patient trust, which depends on transparency, safety, equity, and alignment with accepted standards of evidence, particularly as patients increasingly encounter AI-driven tools outside institutional oversight. Coordinated education, early engagement, and shared expectations across sponsors, regulators, clinicians, and health systems are required to ensure AI use remains evidence-based and trustworthy.

ACTIONABLE ITEMS

- **Develop risk-based validation guidance for AI in Phase III-IV research**
  Clarify expectations for operational, analytical, and evidence-generating AI applications.

- **Define institutional governance boundaries**
  Provide guidance distinguishing research, operational, and clinical governance, and clarify appropriate supervision and monitoring structures for AI-enabled systems to reduce overreach and delays.

- **Leverage public-private partnerships to standardize multi-institution AI governance**
  Bring together health systems, regulators, sponsors, and technology developers to document how AI governance has been successfully implemented in practice and produce reusable guidance through multi-stakeholder collaboration.

- **Align AI development in clinical research with clinically meaningful goals and sustainable incentives**
  Align AI applications with clinically meaningful activities where reimbursement pathways already exist, clarify cost ownership and reimbursement mechanisms for AI infrastructure, and maintain dedicated capacity to innovate in areas without clear reimbursement pathways through alternative incentive and investment models.

- **Establish post-market learning environments**
  Support Phase IV pilot approaches to enable shared learning, inspection readiness, and standard-setting with verified, trusted, unbiased data sources.

- **Promote shared data standards and harmonization**
  Encourage AI-enabled interoperable approaches to data extraction, curation, and analysis.

- **Advance scalable human oversight models**
  Develop practical frameworks for escalation, monitoring, and management of AI-enabled systems.

For more information, please visit
https://dcri.org/insights-and-news/insights/dcri-think-tanks.