Paper ET04

Transforming Clinical Trials: Present Innovations and Future Paths with AI and Agentic Tools

Sandeep Purwar, Summit Therapeutics, Palo Alto, USA

ABSTRACT

Artificial intelligence (AI) and large language models (LLMs) are rapidly reshaping clinical development. Current applications are already reducing repetitive workload and improving efficiency drastically in statistical programming, data management, and regulatory submissions. The next stage introduces **agentic AI**—autonomous systems capable of monitoring trial workflows, triggering actions, and partnering with humans to accelerate decision-making. This paper presents a balanced view of where AI systems are today, what is realistically achievable in the next 3–5 years, and a practical roadmap for organizations to build, validate, govern, and scale agentic tools safely and compliantly. Implications are discussed across efficiency, quality, regulatory expectations, and evolving roles in the clinical research workforce.

INTRODUCTION

The BioPharma industry is navigating increasing trial complexity, rising cost pressures, and demanding timelines. Trials are more global, designs are more adaptive, and the volume and variety of data—from EDC, labs, imaging, ePRO, wearables, and real-world data, continue to grow. AI adoption has accelerated across many sectors, and life sciences are now at a pivotal moment: automation is not only feasible but already delivering operational efficiency and improved data quality. At the same time, regulators, sponsors, and technology providers are still shaping expectations for validation, traceability, and acceptable use of AI in GxP environments.

To adopt AI responsibly, organizations need clarity on three questions (1) Where are we today? (2) What is realistically possible in the next 3–5 years? (3) How do we get from current tools to agentic systems safely and compliantly? This paper addresses these questions from the perspective of a practitioner with over 15 years in statistical programming and development, focusing on pragmatic steps rather than hype. It first reviews the current state of AI in clinical trials, then explores near-future agentic tools, and finally outlines an implementation path with governance considerations and a roadmap for adoption.

TODAY'S STATE OF AI IN CLINICAL TRIALS

AI is already embedded across multiple phases of the clinical trial lifecycle. Most current uses are assistive (providing suggestions or first drafts) or "co-pilot" (automating steps once triggered by a human), rather than fully autonomous systems. Key application areas for AI include study design, data management, statistical programming, and submissions, as described below.

Study Planning and Documentation

Draft protocol and analysis content: Generate first drafts of protocol sections, statistical analysis plans (SAPs), and clinical study report (CSR) sections.

Harmonize documents: Propose edits for consistency across related documents (e.g. ensure endpoints and population definitions align between protocol and SAP). Prepare trial materials: Help create annotated case report forms (CRFs) and quality checklists for trial documentation. In practice, AI provides a starting draft or suggestions, while human experts remain responsible for scientific correctness and regulatory alignment.

Data Management and Cleaning

Suggest and prioritize queries: Identify potential data discrepancies or outliers and prioritize data queries based on patterns in the data and historical trends. Detect anomalies: Flag anomalous or inconsistent data patterns across visits, sites, or related datasets (e.g. adverse events vs. concomitant medications vs. lab results).

These applications reduce manual data review burden and help shorten time to database lock, while still relying on data managers to make final decisions.

Statistical Programming and Analysis

Draft dataset specifications and code: Auto-generate initial SDTM and ADaM dataset specifications from protocol text and even draft portions of the corresponding transformation code.

Generate shells and checks: Produce first-pass table, listing, and figure (TLF) code templates for standard analyses, and suggest validation checks or derivation logic for complex endpoints. By handling boilerplate coding tasks, AI allows programmers to spend less time on repetitive programming and more on complex derivations, interpretation, and study-specific nuances

Regulatory Submissions and Medical Writing

Auto-populate documents: Pre-fill sections of submission documents (e.g. clinical summaries) from source data tables and prior study text.

Consistency checks: Check for cross-document consistency, ensuring endpoints, visit windows, and population definitions are aligned across protocols, SAPs, CSRs, and submission dossiers.

Draft responses: Assist with early drafting of responses to regulatory agency questions by retrieving relevant context from submission documents. In all cases, robust human review and quality control remain mandatory before any AI-generated content is finalized.

SUMMARY OF CURRENT BENEFITS AND LIMITATIONS

Early deployments of AI in clinical development have demonstrated several benefits: reduced time spent on repetitive tasks, improved first-pass quality of drafts and checks, faster access to preliminary analyses, and the ability to reallocate human effort to higher-value activities. For example, some sponsors report faster time to database lock and increased programmer focus on scientific interpretation rather than rote programming. However, limitations are also evident: LLMs can “hallucinate” (produce inaccurate content), results may lack consistency, and integrating AI into validated workflows poses technical and compliance challenges. Regulatory expectations are evolving, and there is a clear need for validation and governance frameworks to manage risk. Today’s reality is that AI is a useful assistant and co-pilot under human control and review – it augments productivity but does not replace human oversight.

AI MATURITY: FROM ASSISTIVE TO AGENTIC TOOLS

A simple maturity model helps frame where we are and where we are going in terms of AI capabilities in trials. We can categorize tools as Assistive AI, Co-Pilot AI, or Agentic AI, representing increasing levels of autonomy and proactivity:

Assistive AI (Reactive AI) : Responds to user prompts; drafts content or suggests queries and checks. All execution and decision-making are done by humans based on the AI’s suggestions.

Co-Pilot AI (Orchestrated AI): Can trigger predefined workflows (e.g. automatically rerun analysis outputs when new data are available) and orchestrate multiple steps once initiated by a human. Decisions remain human-owned, but the AI speeds up execution and improves consistency.

Agentic AI (Proactive AI Agents): Monitors data and processes continuously, and initiates tasks when certain conditions are met (for example, detecting a safety signal or a data threshold being crossed). These systems operate within defined boundaries and escalate to humans as needed, acting as autonomous partners in the workflow.

This paper focuses on the transition from today’s assistive/co-pilot systems to emerging agentic systems over the next 3–5 years. Following figure 1 illustrates three stages of AI capability representing increasing autonomy and proactivity.

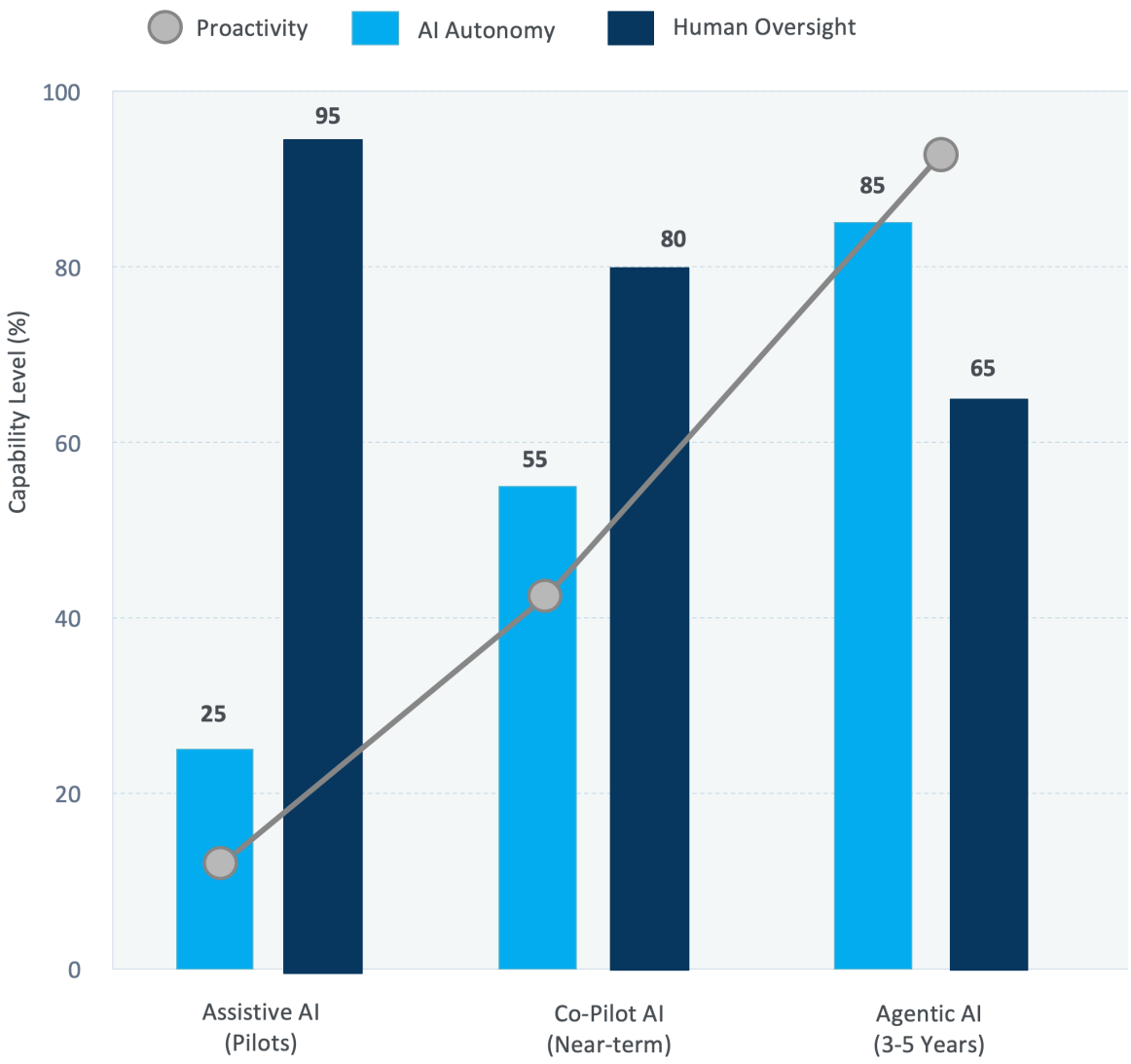

Figure 1. AI maturity progression from assistive to agentic capabilities in clinical development.

NEAR-FUTURE TRANSFORMATION: AGENTIC AI IN 3–5 YEARS

Within the next few years, early agentic AI systems will begin transforming clinical trial operations. These systems are proactive (detect needs and initiate actions), operate with bounded autonomy (within predefined rules), run continuously (real-time monitoring), and provide decision support while humans retain final authority.

In practice, agentic AI may monitor trial data continuously, auto-execute QC checks, create and re-run SDTM/ADaM/TLF pipelines, forecast enrollment and operational risks, run digital twin simulations, and assemble draft reports as data accumulate.

The impact is significant: trial execution shifts from batch to continuous, time to insight shrinks from weeks to hours, workforce roles evolve from hands-on execution to supervision and interpretation, and quality becomes proactive rather than reactive. Importantly, agentic AI does not replace scientific accountability—it amplifies human reach and enables management by exception.

To realize these advancements, organizations should start small and scale. A 24-month roadmap typically begins with low-risk assistive pilots, moves to event-driven co-pilot automations, and finally introduces continuous monitoring and auto-assembly capabilities. Throughout adoption, governance, measurement, and iteration are essential.

PROPOSED AGENTIC AI FRAMEWORK FOR CLINICAL DEVELOPMENT IN-LINE WITH FDA'S 10 GUIDING PRINCIPALS

The FDA's guiding principles for AI in drug development emphasize human-centric design, risk-based controls, adherence to standards, clear context of use, multidisciplinary involvement, strong data governance, sound engineering, performance validation, lifecycle monitoring, and transparent communication. These align closely with a practical framework for deploying agentic AI in clinical development, where autonomous systems must remain safe, traceable, and human-supervised.

Human-Centric and Context-Aware:
Agentic systems must support—not replace—clinical decision-makers. Each agent requires a clearly defined context of use, including scope, boundaries, and intended benefits, with plain-language communication of limitations and appropriate human oversight.

Risk-Based Governance and Standards:
A risk-based approach determines validation depth, sign-off requirements, and change controls. Cross-functional participation (clinical, statistics, data management, QA, regulatory, and AI engineering) ensures alignment with GxP, cybersecurity, and ethical standards throughout the lifecycle.

Data, Engineering, and Documentation:
Agentic AI depends on standardized, well-governed data, documented data lineage, and good engineering practices. Models must use fit-for-purpose data with attention to explainability, robustness, and traceability to meet inspection and compliance expectations.

Bounded Autonomy and Human Collaboration:
Agents should operate with bounded autonomy (observe → act → escalate) and clearly defined human–AI collaboration modes (e.g., human-in-the-loop). Evaluation must consider system-level performance, including how humans interpret, approve, or override outputs.

Validation, Monitoring, and Lifecycle Management:
Formal risk-proportionate validation is required prior to deployment, followed by ongoing monitoring for drift, safety, and compliance risks. Change control, scheduled reviews, and CAPA workflows maintain reliability across the agent lifecycle.

Transparency, Auditability, and Continuous Learning:
Agent behavior must be explainable, traceable, and auditable, with logs and versioning that allow reconstruction during inspections. Clear communication of updates, performance characteristics, and limitations supports trust and regulatory readiness.

In summary, aligning the FDA's AI principles with a structured agentic AI framework ensures that autonomous capabilities develop within safe, compliant, and human-supervised boundaries. When appropriately governed and validated, agentic AI can enhance evidence generation, reduce operational latency, and expand human oversight capacity in clinical development.

CONCLUSION

AI has already begun transforming clinical development by reducing repetitive workload and accelerating analytical workflows, yet these gains represent only the first phase of a broader shift. As the industry moves toward agentic AI—autonomous systems that observe, act, and escalate within defined boundaries—the operating model of clinical trials will change from episodic, batch-driven processes to continuous, data-driven cycles. This shift promises shorter time-to-insight, more proactive quality management, and increased scientific agility.

However, the transition must be responsibly managed. Agentic systems must be validated, traceable, human-supervised, and aligned with FDA's emerging principles for context of use, risk-based controls, standards adherence, data governance, and lifecycle monitoring. When organizations invest in these foundations, they create the conditions for safe and compliant deployment of autonomous capabilities in GxP environments.

The path forward is pragmatic rather than disruptive: start with assistive use cases, mature into orchestrated co-pilot automations, and introduce agentic behaviors only when data, governance, and validation are ready. This staged approach balances innovation with regulatory expectations and ensures that human expertise remains central. By embracing this evolution, BioPharma organizations can expand oversight capacity, accelerate evidence generation, and ultimately improve patient access to new therapies.

REFERENCES

- U.S. Food and Drug Administration. *Artificial Intelligence and Machine Learning in Drug Development (Discussion Paper).* 2023.
- European Medicines Agency. *Reflection Paper on the Use of Artificial Intelligence in the Medicinal Product Lifecycle.* 2024.
- Clinical Data Interchange Standards Consortium (CDISC). *Study Data Tabulation Model Implementation Guide (SDTMIG) Version 3.4.* 2021.
- Nature Medicine. *Large Language Models in Clinical Research: Opportunities and Challenges.* 2023.
- McKinsey & Company. *The Economic Potential of Generative AI.* 2023.
- Goldman Sachs. *The AI Growth Boom and Labor Productivity Outlook.* 2023.
- Deloitte Life Sciences. *Digitizing the Clinical Trial Value Chain.* 2022.
- Microsoft Research. *Autonomous AI Agents for Enterprise Workflows.* 2024.
- FDA: *Guiding Principles of Good AI Practice in Drug Development. 2026*