DAILY ENGLISH READING

26 July 2026 - 30-minute active reading pack for Davin

TODAY'S QUESTION

Where can AI create real value in clinical trials - and what must be validated before it can be trusted?

WHY SELECTED

This topic sits directly at the intersection of Davin's work and the future of clinical research. The main review explains how AI may affect protocol design, recruitment, trial conduct, monitoring and data interpretation. The regulatory sources show that faster and more digital trials still require validation, governance, reliable data and accountable human oversight.

Difficulty: C1 - clinical research + AI governance

READING ORDER - 30 MINUTES

Min	Task	Focus
0-3	Preview	Read the title, sources and five questions. Predict which trial stage is easiest to automate.
3-15	Main article	Read the abstract, applications, implementation challenges and conclusion.
15-20	FDA	Identify what is already being tested in real-time clinical trials.
20-25	EMA	Note fit for purpose, risk-based and proportionate approaches.
25-30	Output	Answer the questions, then give a 90-second CRO/SP summary.

SOURCES - OPEN ACCESS

MAIN - European Journal of Clinical Investigation - 2026
The potential of artificial intelligence in clinical trials
PMC full text | PubMed abstract

SUPPORT 1 - U.S. FDA - 28 Apr 2026
FDA Announces Major Steps to Implement Real-Time Clinical Trials
FDA announcement

SUPPORT 2 - European Medicines Agency - updated Jul 2026
ICH E6 Good Clinical Practice - Scientific Guideline
EMA guideline page

BRIEF BACKGROUND

Clinical trials generate the evidence used to evaluate the safety and efficacy of therapies, but they remain expensive, slow and operationally complex. The main review describes potential AI applications across the trial lifecycle: protocol optimization, synthetic or external control approaches, adaptive designs, participant identification, recruitment, monitoring, endpoint assessment and data interpretation.

The most credible near-term opportunities are often high-volume, repetitive and review-heavy tasks. Natural language processing can extract eligibility information from unstructured clinical notes. Machine-learning systems can help identify recruitment risks, prioritize monitoring, detect unusual data patterns and support endpoint review.

The central limitation is that model performance depends on the data and operating context. Incomplete records, inconsistent definitions, duplicated observations, shifting populations and weak system integration can make outputs unreliable. A model can also degrade after deployment, which is why continuous validation and monitoring matter.

In April 2026, the FDA announced two proof-of-concept real-time clinical trials that report selected endpoints and data signals to the agency as they emerge. ICH E6(R3) promotes fit-for-purpose, risk-based and proportionate trial conduct, while Annex 2 addresses decentralized, pragmatic and real-world-data-enabled designs.

SP / DATA LENS

For statistical programming, trustworthy automation requires more than accurate-looking output. The full chain should remain reconstructable: source data, data standards, transformations, model or prompt version, code version, validation evidence, exceptions, reviewer decisions and final deliverables. Efficiency is valuable only when reproducibility and auditability remain intact.

KEY VOCABULARY

Term	中文	Meaning / use
trial lifecycle	临床试验全生命周期	All stages from protocol design to analysis and reporting.
protocol optimization	方案优化	Improving design, feasibility, criteria or endpoints.
participant recruitment	受试者招募	Finding, screening and enrolling eligible participants.
eligibility criteria	入排标准	Rules determining who may or may not enter a trial.
unstructured data	非结构化数据	Information such as free-text clinical notes that is not stored in fixed fields.
endpoint assessment	终点评估	Determining whether and when a predefined trial outcome occurred.
synthetic control arm	合成对照组	A model-based comparison group created from historical or external data.
adaptive design	适应性设计	A design allowing prespecified changes based on accumulating trial data.
data lineage	数据血缘	A traceable record of where data came from and how it changed.
fit for purpose	适合预定用途	Appropriate for the specific decision, use case and level of risk.
risk-proportionate	与风险相称的	Applying controls in proportion to the importance and risk of an activity.
continuous validation	持续验证	Repeatedly checking model performance after initial deployment.
human oversight	人工监督	Qualified people reviewing, challenging or overriding system outputs.
interoperability	互操作性	The ability of systems to exchange and use data consistently.

Term	中文	Meaning / use
auditability	可审计性	The ability to reconstruct and verify decisions, data and processes.

USEFUL PHRASES

1. **across the clinical trial lifecycle**
AI may create value across the clinical trial lifecycle.

2. **streamline a manual process**
The system is designed to streamline a manual screening process.

3. **extract information from unstructured notes**
NLP can extract eligibility information from unstructured notes.

4. **depend on the quality of the input data**
Model reliability depends on the quality of the input data.

5. **remain under human oversight**
High-impact decisions should remain under human oversight.

6. **be rigorously validated and continuously monitored**
The model must be rigorously validated and continuously monitored.

7. **adopt a risk-based and proportionate approach**
Sponsors should adopt a risk-based and proportionate approach.

8. **be fit for its intended purpose**
The tool must be fit for its intended purpose.

9. **preserve data lineage and auditability**
Automation should preserve data lineage and auditability.

10. **move from proof of concept to operational use**
The challenge is moving from proof of concept to operational use.

COMPREHENSION QUESTIONS

1. Which stages of the clinical trial lifecycle may benefit from AI?

2. Why is participant screening a suitable near-term use case for NLP and generative AI?

3. What data-quality or infrastructure problems can reduce the reliability of an AI system?

4. What is the difference between a proof of concept and a validated operational system?

5. How do fit-for-purpose and risk-proportionate principles affect AI use in trials?

RETELLING PROMPTS

30 seconds: Explain the problem, the proposed uses and the main caution.

45 seconds: Retell the topic as: design -> recruitment -> conduct -> data -> governance.

60 seconds: Explain one use case from the perspective of a CRO statistical programmer.

5-MINUTE SPEAKING / OUTPUT TASK

Your role: You are speaking in an internal CRO meeting about whether an AI tool should be introduced into a clinical programming workflow.

Minute 1 - Choose one use case

Choose one: protocol review, eligibility screening, data-quality checks, coding assistance or draft result summaries.

Minutes 2-3 - Give the business case

Explain the current manual process, the expected efficiency gain and the evidence required before deployment.

Minute 4 - Define the controls

Mention data lineage, validation datasets, version control, exception handling, human review and audit trails.

Minute 5 - Final 90-second take

Use this structure:

1. **The use case I would prioritize is...**
2. **It could streamline...**
3. **However, its performance depends on...**
4. **Before operational use, we would need...**
5. **My recommendation is to begin with...**

ONE SENTENCE TO KEEP

AI can reduce manual burden in clinical trials, but trustworthy deployment depends on validated performance, traceable data and accountable human oversight.