DAILY ENGLISH READING · 23 SEPTEMBER 2026

When AI enters a clinical trial, what exactly are we validating?

C1 · 30 minutes · AI × clinical trials × validation

Why selected

A new open-access EClinicalMedicine review distinguishes AI-as-intervention from AI-for-trial-operations. The distinction changes what should be prespecified, how change control works, and what evidence is needed. For statistical programming, it raises a practical question: when does an AI productivity tool become part of the evidentiary chain?

Reading order

0–3 min · Define the two AI roles.

3–13 min · Read trial-lifecycle applications and evidence.

13–19 min · Focus on prespecification, change control, transportability and oversight.

19–24 min · Connect E6(R3) risk management with E9(R1) estimands.

24–25 min · Classify SAS log review, ADaM derivation and endpoint classification.

25–30 min · Build a TFL-agent validation matrix.

Brief background

Useful phrases

1. the validation target depends on the role AI plays in the trial.

2. an operational tool can still become inferentially important.

3. prespecification limits hidden flexibility after outcomes are known.

4. change control should cover both model versions and surrounding workflow logic.

5. performance must be evaluated in the population and setting where the system will be used.

6. risk-based oversight should focus on critical-to-quality factors.

7. a high-accuracy component may still create unacceptable downstream risk.

8. the same AI system may require different evidence under different intended uses.

9. traceability should preserve inputs, versions, decisions, and human approvals.

10. automation should be judged by consequence as well as technical performance.

Comprehension questions

• What is the central difference between AI-as-intervention and AI-for-trial-operations?

• Why can operational AI still become inferentially important?

• How does E6(R3)'s risk-proportionate approach help scale validation effort?

• Why is correct SAS insufficient if it implements the wrong estimand?

• What evidence makes an AI-generated TFL workflow auditable?

Retelling prompts

30 sec · Explain the two AI roles without using the words treatment or operations.

45 sec · Explain why an endpoint classifier deserves more governance than an email assistant.

60 sec · Explain intended use → decision impact → deterministic controls → human review.

5-minute speaking/output task

Build a validation matrix for an AI-assisted TFL agent. Define intended use, three result-changing failure modes, deterministic controls, audit evidence, and the mandatory human-review boundary.

One sentence to keep

The right question is not whether AI is accurate in general, but whether its evidence, controls, and oversight are proportionate to the decision it can change.

Sources

https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370(26)00449-9/fulltext

https://www.ema.europa.eu/en/ich-e6-good-clinical-practice-scientific-guideline

https://www.fda.gov/regulatory-information/search-fda-guidance-documents/e9r1-statistical-principles-clinical-trials-addendum-estimands-and-sensitivity-analysis-clinical