

DAILY ENGLISH READING

7 August 2026 · 30-minute active reading pack for Davin

TODAY'S QUESTION

Where should LLMs stop and deterministic validation begin in TFL quality control?

WHY SELECTED

PROVE directly addresses TFL QC: LLMs interpret varied reporting language, while programmed validators retain responsibility for arithmetic and logical correctness. This is a strong design pattern for regulated statistical-programming automation.

Difficulty: C1 · TFL QC + LLM semantics + deterministic validation

READING ORDER · 30 MINUTES

Min	Task	Focus
0-3	Preview	Predict which QC problems need semantics rather than arithmetic.
3-15	Main article	Read framework, validators, simulation and results.
15-20	CDISC CORE	Review machine-executable conformance rules.
20-25	FDA/EMA	Focus on context of use and lifecycle management.
25-30	Output	Design one hybrid TFL QC workflow.

SOURCES · OPEN ACCESS

MAIN · arXiv preprint · 30 Jul 2026

PROVE: Bridging Probabilistic LLMs and Deterministic Statistical Validation

[Open preprint](#)

SUPPORT 1 · CDISC

CDISC Open Rules Engine (CORE)

[Official page](#)

SUPPORT 2 · FDA / EMA

Guiding Principles of Good AI Practice in Drug Development

[Official page](#)

BRIEF BACKGROUND

PROVE builds source-linked semantic records from clinical reporting outputs so related tables, listings and document context can be compared consistently.

All arithmetic and logical decisions remain in programmed validators. LLMs are optional and used for language interpretation, label mapping and retrieval.

The simulation used ten synthetic oncology reporting packages spanning raw data, SDTM, ADaM and TFL outputs. Each altered package contained 15 known injected discrepancies.

With exact labels, all automated variants performed perfectly within implemented rule classes. Under label variation, LLM-assisted semantic matching achieved mean overall recall of 0.993 and F1 of 0.996.

The gain came from recognizing wording variation, not from doing better mathematics. CDISC CORE reinforces the same principle by making conformance rules machine-executable and repeatable.

KEY VOCABULARY

Term	中文	Meaning / use
deterministic validator	确定性校验器	Programmed logic that produces the same result for the same input.
semantic matching	语义匹配	Recognizing equivalent meaning despite different wording.
cross-output consistency	跨输出一致性	Agreement among related tables, listings, figures or versions.
reporting graph	报告语义图	A structured representation linking values, concepts and source evidence.
traceability	可追溯性	The ability to connect a finding back to its source and rule.
discrepancy injection	差异注入	Deliberately inserting known errors to test a validation system.
label variation	标签变体	Different wording used for the same clinical or statistical concept.
denominator consistency	分母一致性	Agreement in population denominators across related outputs.
safety hierarchy	安全性层级关系	Count relationships among SOC, PT and related safety summaries.
within-library recall	规则库内召回率	Detection rate among discrepancies covered by implemented rules.
false positive	假阳性	A reported issue where no true discrepancy exists.
canonical label	标准语义标签	The normalized label used by a validation system.
sponsor-defined rule	申办方自定义规则	A controlled executable rule specific to a sponsor or study.
evidence-centered review	证据中心化审核	Review output that links each finding to supporting evidence.
machine-executable rule	机器可执行规则	A formal rule that software can run reproducibly.

USEFUL PHRASES

1. reserve numerical decisions for programmed validators

The workflow reserves numerical decisions for programmed validators.

2. use LLMs for interpretation rather than arithmetic

LLMs are used for interpretation rather than arithmetic.

3. map varied labels to canonical concepts

The model maps varied labels to canonical concepts.

4. link every finding back to source evidence

Every finding is linked back to source evidence.

5. check consistency across related outputs

The engine checks consistency across related outputs.

6. recompute percentages from counts and denominators

The validator recomputes percentages from counts and denominators.

7. remain robust to wording variation

The system should remain robust to wording variation.

8. treat natural-language rules as review support

Natural-language rules are treated as review support.

9. keep the final correctness decision reproducible

The final correctness decision remains reproducible.

10. validate against known injected discrepancies

The framework was validated against known injected discrepancies.

COMPREHENSION QUESTIONS

1. Why separate language interpretation from numerical verification?
2. What is the semantic reporting graph for?
3. Why did label variation reduce lexical matching performance?
4. What does the recall improvement reveal about the LLM's contribution?
5. How does CDISC CORE support deterministic validation?

RETELLING PROMPTS

30 seconds: Problem -> architecture -> result.

45 seconds: Ingest -> normalize -> validate -> evidence -> review.

60 seconds: Apply the pattern to AE, efficacy or disposition QC.

5-MINUTE SPEAKING / OUTPUT TASK

Your role: Design a hybrid QC assistant for one TFL family.

Minute 1: Choose the TFL family.

Minutes 2-3: Separate semantic interpretation from deterministic calculations and cross-output checks.

Minute 4: Add executable rules, evidence links, regression tests and LLM-off mode.

Minute 5: State your acceptance criteria.

ONE SENTENCE TO KEEP

In clinical reporting QC, LLMs should help the system understand what an output means, while reproducible code decides whether the output is correct.