

DAILY ENGLISH READING

28 August 2026 - 30-minute active reading pack

TODAY'S QUESTION

If AI can do the analysis, what still makes the result statistically defensible?

Difficulty: C1 | statistical warrant | AI | CDISC 360i | regulated analytics

WHY SELECTED

AI can automate programming, model fitting, simulation, visualization and synthesis.

The new paper argues that automation cannot remove the logical conditions connecting data to a scientific claim. For SP work, the scarce skill shifts toward defining the target, validating the evidence chain, quantifying uncertainty and preserving accountability.

READING ORDER

0-3 min Preview: define statistical warrant in one sentence.

3-14 min Main: target, observation regime, assumptions, uncertainty and validation.

14-20 min CDISC 360i: machine-readable SAPs, Analysis Concepts and traceability.

20-25 min FDA/EMA: context of use, risk, governance and lifecycle management.

25-30 min Output: build a warrant card for one TFL number.

SOURCES

Main: Statistics in the Age of AI - arXiv, submitted 20 Aug 2026

<https://arxiv.org/abs/2608.20541>

Support: CDISC 360i - Analysis Concepts and end-to-end automation

<https://www.cdisc.org/standards/cdisc-360i>

Support: FDA / EMA - Guiding Principles of Good AI Practice in Drug Development

BRIEF BACKGROUND

Statistical warrant connects a target claim to data through design and assumptions.

It includes procedure, uncertainty, validation criterion, loss, governance and accountability.

Identifiability is a hard boundary: an algorithm cannot recover information absent from the observation regime without new data or assumptions.

Analytical abundance creates selection uncertainty because many workflows are now cheap.

SP hierarchy: correct code < correct implementation < correct question < defensible warrant.

KEY VOCABULARY

statistical warrant	统计论证依据	full chain that justifies a claim
observation regime	观测机制	how data were generated or collected
identifiability	可识别性	whether the target can be recovered
target quantity	目标量	precise estimand prediction or decision target
evidential meaning	证据意义	what data can legitimately support
uncertainty assessment	不确定性评估	quantifying uncertainty around results
validation criterion	验证标准	rule for acceptable performance
loss structure	损失结构	costs and consequences of errors
analytical abundance	分析方案过剩	many plausible analytical paths
model selection uncertainty	模型选择不确定性	uncertainty from choosing among workflows
provenance	来源与处理谱系	record of data origin and transformations
deployment validation	部署验证	testing in the real use environment
governance	治理	oversight controls ownership and rules
accountability	问责 / 责任归属	clear responsibility for choices and outcomes
consequential decision	高影响决策	decision with meaningful consequences

USEFUL PHRASES

1. data do not speak for themselves
2. make the target explicit before choosing the method
3. separate description, prediction, inference, and decision
4. state the observation regime and assumptions
5. quantify uncertainty around the full analytical system
6. treat provenance as part of the evidence
7. validate the system in its deployment environment
8. do not confuse automation with identification
9. make responsibility attributable
10. ask what claim the analysis is actually allowed to support

COMPREHENSION QUESTIONS

1. What does statistical warrant add beyond correct code and a valid model?
2. Why can AI not solve a target that is not identifiable from the available data?
3. Why does analytical abundance create uncertainty of its own?
4. How do CDISC 360i Analysis Concepts strengthen the warrant behind a result?
5. Which SP responsibilities must remain attributable to people or organizations?

RETELLING PROMPTS

30 sec: define statistical warrant and name four components.

45 sec: target -> data generation -> assumptions -> method -> uncertainty -> validation.

60 sec: explain why less manual coding can make reasoning and governance more valuable.

5-MINUTE SPEAKING / OUTPUT TASK

Minute 1: choose one TFL result.

Minutes 2-3: state target, population, observation regime, assumptions and uncertainty.

Minute 4: add AI context-of-use, versioning, deterministic validation and reviewer controls.

Minute 5: defend why the result is worth believing.

ONE SENTENCE TO KEEP

When analytical execution becomes cheap, the scarce professional skill is not producing an answer; it is defending the evidence chain that makes the answer worth believing.