

Does New Medical Evidence Change AI Treatment Answers?

A Longitudinal Study of Evidence Citation, Medicine Naming and Selected Treatment Framing Around EASD 2026

Ozue C. Ezeife · AlphaCitation Research · October 2026

Founder, AlphaCitation

Dubai, United Arab Emirates

Published by AlphaCitation Research, the research programme of AlphaCitation, a digital product of Akunudo LLC-FZ
Executive brief · v1.0.2 · October 2026 (post-publication note PN1 added 15 October 2026 and corrected 16 October 2026; see the paper's version history)

Bottom line

- 1. The largest movement sat in answers citing a local web page, not the congress.** Claude Haiku 4.5's naming of Wegovy or semaglutide in UAE answers rose about 17 percentage points around EASD 2026, and none of its UAE answers cited congress material. After the congress, 94% of its scored answers citing a local commercial blog post displaying an August 2026 date named the medicine, against 50% of the rest. Post-registration and descriptive (D24): an association, not proof of cause.
- 2. The tracked trial readouts were not verified in AI answers.** Ten to twelve days after presentation, the inspection procedure verified none of seven tracked readouts. In the registered markets, congress-derived press releases or media were cited in 6 of 2,408 post-congress answers; seven of 2,408 under the either-route definition (one text-only answer found after the freeze).
- 3. Single readings mislead.** Before the congress, observed naming rates varied across cycles; this study did not determine whether contemporaneous evidence or service changes contributed to that variation. Judge change against a multi-week, per-engine range.
- 4. Many sampled type 1 glycaemic-use mentions omitted indication-status wording.** In a post hoc, same-family model recode of sampled Saudi-arm namings, 30/44 before-window and 16/18 after-window glycaemic-use mentions lacked qualifying indication-status wording under the study's rules. Saudi product-specific labeling was not independently established; these counts identify a communication issue for further regulatory review.
- 5. Scope.** Four configured developer API endpoints with web search; observation from 22 August to 11 October 2026, including a short after-congress follow-up; fixed market-framed question slots with documented text changes. Findings concern sampled answers and citations, not consumer applications, prescribing or sales.

Question and design

When new scientific evidence enters the public domain, what changes in AI-generated treatment guidance, how can those changes be measured, and what evidence permits them to be attributed to the scientific event? Four endpoints (gpt-4o-mini, Claude Haiku 4.5, Gemini 2.5 Flash, Perplexity sonar) answered questions about semaglutide (Wegovy) in the UAE, empagliflozin (Jardiance) in Singapore and teplizumab (Tzield) in Saudi Arabia, an arm added after registration: 64 of 67 cycles admitted. The protocol was committed before collection of after-window cycles, after the congress presentations and inspection of a during-window UAE cycle. Some computational interpretations were specified after preliminary after-window counts were visible; additional robustness and source-tracing analyses were undertaken after the results. These timings limit the confirmatory interpretation of the movement verdicts.

Findings

Claude's rise. Claude named Wegovy or semaglutide in 256 of 378 scored post-congress answers (67.7%; Wegovy itself in 254), 0.4 points above its corrected pre-congress range: a registered movement, borderline. The pooled movement verdict is sensitive to the band-rule interpretation. A post hoc cycle-level analysis estimates a Claude increase of 16.8 percentage points (95% interval 7.2 to 26.4) under independent-cycle and bridge assumptions; it supports a descriptive increase within this design. The page first appeared in Claude's resolved citations on 29 September, then in 152 of its 420 post-congress answers. All 152 were scored: 143 named the medicine, against 113 of the 226 scored answers not citing it. From 29 September, 152 of 164 citing answers shared wording with one of two tested page sentences. A model or retrieval update could also explain it.

Perplexity's step. Perplexity sonar named the medicine in 17 of 54 scored answers on 6 October and 40 of 54 on 7 October, as its whole source mix changed (the share of cited links classified as "Other" rose 40.3 percentage points; many prominent newly cited sources were market-research sites). Pooled, it stayed 0.074 points inside its range: a registered "no movement" that another reading of the rule reverses (post-registration check). A post-publication check found that Perplexity's responses changed in structure on the same day in every question set examined, including the

11 of a monitoring panel's 12 product categories that track no diabetes or weight-management medicine (on each day from 7 October at least 98% of stored answers carried exactly 10 search results, against at most 2% on each day from 20 September to 6 October, when answers carried about 20), with no change in AlphaCitation's request on the record available (per-request parameters were not stored); the cause is not established (paper, Section 6.2, note PN1).

The congress route. Among verified E1-qualifying sources, none was a congress page or a journal publication of a readout-set trial; most qualifying material concerned a switching study or indirect comparisons.

Variation. After a post hoc adjustment for run counts, four of eight examined engine–market combinations (UAE Gemini and gpt-4o-mini, Saudi Perplexity and Gemini) varied clearly more between days than within one; UAE Claude did not. Saudi Perplexity's reading is confounded by the 7 October change in its responses (paper, note PN1).

What to do

- **Communications and digital.** Check crawlers can read your release pages; Lilly's returned HTTP 403 to our fetcher, and sponsor releases were often cited as syndicated copies.
- **Medical and regulatory review.** Put the trial name and local regulatory status in each congress release's first paragraph. Two answers stated a congress result without naming the trial.
- **Insights and brand.** Report brand-only naming beside brand-or-molecule naming.
- **Medical affairs.** Monitor competitors' comparative claims and their sources: non-randomised material may reach answers first (inference).
- **Medical information.** Retest missing indication-status wording in Saudi type 1 answers (exploratory, post-registration).

Uncertainties

After the congress, 9.9–31.5% of new-URL citations per engine could not be read: unverified, not negative. Re-run with the corrected count, the attribution test stays uninformative (Gemini UAE: 2/2 qualifying named vs 204/376; Fisher p 0.503, minimum attainable 0.206); no verdict changes. Coding and drafting were model-assisted (Claude models, a family including an engine studied) under the author's oversight. Blinded human validation, as selected quality checks: the author agreed with 17/17 codes in the ordinary sampled strata and 30/33 codes in deliberately risk-enriched strata; these are quality-check results for the selected items, not estimates of overall coding accuracy or independent clinical validation. The efficacy endpoint failed its audit and was quarantined. Gemini results are reported in aggregate; verbatim outputs and grounding links are not redistributed.

Read further

[Paper](#) · [Methodology](#) · [Evidence](#) · frozen record 6d1b45d3

© 2026 AlphaCitation (Akunudo LLC-FZ) · CC BY 4.0

Governance. Paper No. 01, v1.0, has not undergone independent external peer review. Medicines are named generic first, whatever their results, by a decision taken after the results; it is not pre-registered and is not neutral in perception. Naming implies no endorsement, sponsorship or commercially meaningful effect. Findings describe the engines' answers, not the medicines.

Funding and competing interests. This study was independently conducted and funded by AlphaCitation. No pharmaceutical manufacturer commissioned, sponsored, or influenced the research. The author is the founder of AlphaCitation, which develops and offers AI information intelligence services to life-sciences organizations. He previously held commercial leadership positions at Novo Nordisk, whose medicine Wegovy is included in the study.

Correspondence. Scholarly correspondence, methodological critiques and correction requests may be directed to analyst@alphacitation.com. Substantive corrections will be evaluated and documented in the publication's version history.

Request an AlphaCitation briefing.