PICO Provenance: Evaluating AI Traceability for EU Joint Clinical Assessments

HTA153

Rhythm Arora[1], Victoria L Molenkamp[2], Hanna-Liisa Vilu[3]

[1]Thermo Fisher Scientific, Hyderabad, India; [2]Thermo Fisher Scientific, London, UK; [3]Thermo Fisher Scientific, Tallinn, Estonia

Background

- The EU Joint Clinical Assessment (JCA) process requires the assessment scope to be translated into specific PICO questions that define the populations, comparators, and outcomes relevant to the clinical assessment.
- Once the final assessment scope is communicated, the health technology developer generally has 100 days to submit the JCA dossier. This is reduced to 60 days in specified accelerated or new-indication procedures, while JCA activities run in parallel with the EMA regulatory process.
- These compressed timelines, together with the potential for multiple country-driven PICO requirements, create a need to anticipate plausible populations, comparators and outcomes before final scope is received. This supports evidence planning and dossier readiness.
- AI-assisted PICO generation may support early preparation at scale. Validation should establish whether sources are retrievable and directly support each AI-generated PICO claim without material inference or overstatement (traceability and accuracy).[1-3]

Objectives

- To evaluate the traceability of AI-generated PICO predictions by determining whether the evidence for each predicted PICO was verifiable, accurately represented, and supported by the cited source.

Methods

- We analysed evidence across multiple European settings to assess whether country-specific differences affected the provenance of AI-generated PICO predictions.
- For each claim, we retrieved and reviewed the cited source to confirm that it existed and directly supported the claim. **Table 1** summarises the study configuration, evidence sources, validation approach, and traceability assessment framework.

Table 1. Study Configuration and Evidence Verification Approach

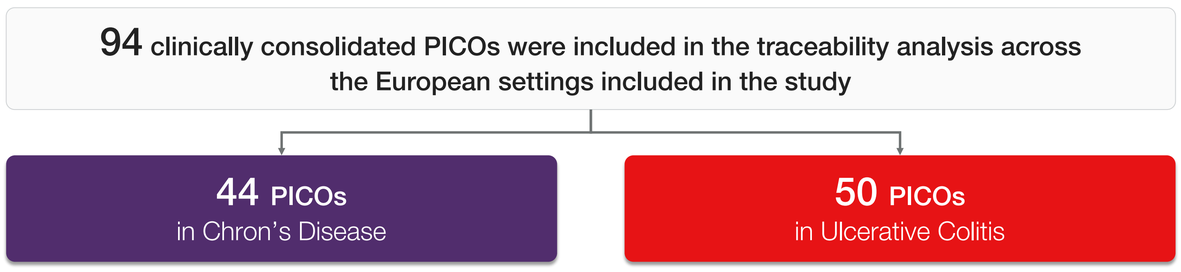
Domain	Study specification
AI tool	Custom GPT built on OpenAI GPT-5.6 sol
Supporting materials	Example PICO outputs and EU HTA/JCA methodological reference materials
Geographic scope	France, Germany, Italy, Poland, Spain, Sweden along with EU-level evidence
Intervention	Investigational biologic
Indications	Crohn's disease (CD) and ulcerative colitis (UC)
Evidence sources	European and national clinical guidelines; open-web regulatory sources; national HTA and payer evidence; relevant development programme sources
Validation approach	Each cited source was independently retrieved and compared with the AI-generated claim attributed to that source
Traceability scale	0 = unverifiable/fabricated; 1 = verifiable but inaccurate/overstated; 2 = verifiable and accurate

Abbreviations: AI = artificial intelligence; CD = Crohn's disease; EMA = European Medicines Agency; GPT = generative pre-trained transformer; HTA = health technology assessment; JCA = joint clinical assessment; UC = ulcerative colitis

Results

PICO Predictions Across Crohn's Disease and Ulcerative Colitis

Figure 1. Core PICO Patterns Differed by Indication



Crohn's Disease

- Comparator convergence/divergence: European and national evidence converged most consistently on infliximab, adalimumab, vedolizumab, and ustekinumab. Greater variation was seen for newer therapies, combination treatment and selected surgical comparators.
- Key populations included:
  - Moderate-to-severe disease
  - Disease refractory to conventional treatment or biologic-naïve patients
  - Prior anti-TNF or biologic failure
  - Steroid-dependent or steroid-refractory disease
  - Maintenance populations, with selected fistulising or perianal scenarios

Ulcerative Colitis

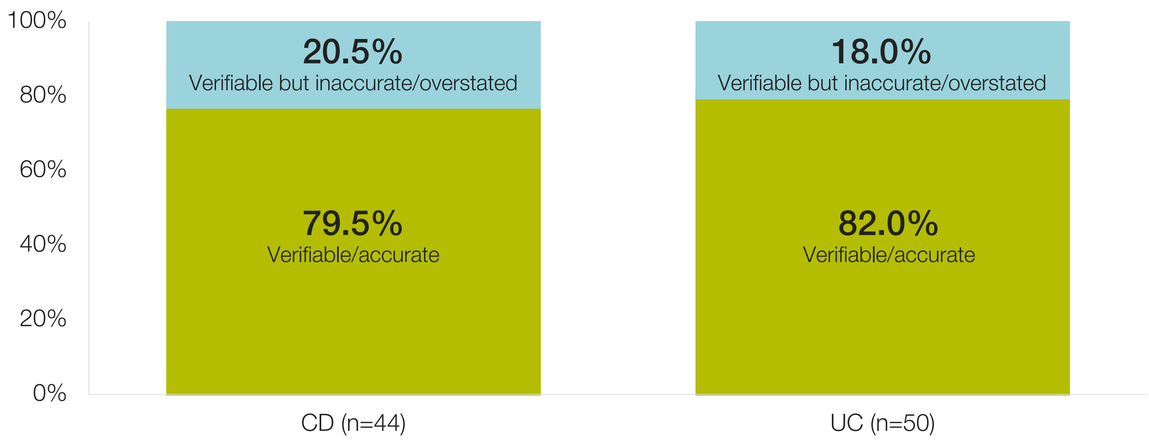
- Comparator convergence/divergence: European and national evidence broadly converged on anti-TNF therapies, vedolizumab and ustekinumab. Greater variation was seen for JAK inhibitors, S1P modulators, newer IL-23 therapies, and combination treatment.
- Key populations included:
  - Moderate-to-severe UC with failure or intolerance conventional treatment and no prior advanced therapy
  - Prior advanced-therapy failure
  - Failure or intolerance to 5-ASA, corticosteroid, or thiopurine; primary non-response, loss of response, or intolerance to anti-TNF therapy
  - Maintenance populations, with selected acute severe UC and late-line scenarios

Results (cont.)

Traceability of Predicted PICOs

- Overall, 76/94 (80.9%) predicted PICOs were supported by verifiable evidence that accurately supported the corresponding PICO claim; 18/94 (19.1%) were linked to verifiable evidence but involved inference or greater specificity than was directly supported by the source. No PICO was classified as unverifiable.

Figure 1. Evidence Traceability of AI-generated PICOs by Indication



Abbreviations: AI = artificial intelligence; CD = Crohn's disease; UC = ulcerative colitis

Table 2. Evidence Provenance and Traceability Findings by Source Type

Evidence source	Key findings
Clinical guidelines	Provided the most direct support for established comparators and treatment populations; limitations arose where guidance predated newer agents or lacked detailed sequencing
National HTA/payer evidence	Provided explicit country-specific positioning where available, but was the main area of residual uncertainty for newer-agent positioning and sequencing

Abbreviations: HTA = health technology assessment

- Lower traceability reflected the lack of source specificity and the need for inference, rather than failure to identify an underlying source.
- **Crohn's disease, 35/44 (79.5%):** PICOs were verifiable and accurately supported. Lower traceability was mainly associated with newer-agent positioning and selected fistulising/perianal or surgical scenarios.
- **Ulcerative colitis, 41/50 (82.0%):** PICOs were verifiable and accurately supported. Lower traceability was mainly associated with sequencing of newer agents and inferred country-specific comparator relationships.

Discussion

- Traceability was broadly consistent across CD and UC. It was strongest when predicted PICOs were anchored to explicit guideline recommendations and established treatment pathways.
- Reduced traceability was concentrated in the positioning of newer agents and country-specific sequencing, HTA, payer, or regulatory evidence often required interpretation and did not directly support the predicted comparator relationship.
- These findings suggest that AI-assisted PICO prediction can provide a useful starting point for JCA preparation. Source-level validation remains necessary to determine whether a retrievable source directly supports the AI-generated claim, particularly in rapidly evolving treatment landscapes.
- The AI-enabled approach allowed efficient evaluation of many PICO scenarios (44 CD; 50 UC) across multiple European settings, while retaining sufficient depth to examine population definitions, comparator selection, treatment sequencing, and evidence provenance.
- AI-generated PICO outputs still require human review and source verification, particularly when predictions depend on the positioning of newer agents, country-specific sequencing, or inferred comparator relationships. Reliability also dependent on the completeness, recency, and specificity of available evidence. AI-supported outputs should inform, rather than replace, expert assessment.

Conclusions

- AI-assisted PICO prediction enabled efficient assessment of a broad, multi-country PICO landscape. Traceability was strongest when predictions were supported by explicit statements in current, authoritative sources, including clinical guidelines and formal HTA, payer, or regulatory evidence.
- Reduced traceability was concentrated in the positioning of newer agents, country-specific sequencing, and inferred comparator relationships. A retrievable source does not necessarily directly support the AI-generated claim.
- Systematic source verification and expert review remain essential. AI can support and accelerate JCA preparation, while expert human assessment remains necessary.

References

1. European Commission. *Off J Eur Union*. 2024;2024/1381. **2.** Directorate-General for Health and Food Safety. *Guidance on the Scoping Process*. European Commission; 2024. **3.** Directorate-General for Health and Food Safety. *Procedural Guidance for JCA Medicinal Products*. European Commission; 2024.

Disclosures and Acknowledgements

All authors are employees of PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific, and acknowledge support with review and analysis by their colleagues. Thermo Fisher Scientific provided funding for this study. Editorial and graphic design support were provided by Carrie Collins-Cann and Karissa Calara of Thermo Fisher Scientific.

Contact: rhythm.arora@thermofisher.com

Presented at ISPOR-EU—The Professional Society for Health Economics and Outcomes Research—Europe | 8–11 November 2026 | Vienna, Austria

© 2026 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. This information is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others.