MadeAi

AI-Assisted JCA Preparation: PICO Anticipation and Evidence Planning



White Paper

Summary

The European Union (EU) Joint Clinical Assessment (JCA), mandated under Regulation (EU) 2021/2282, requires manufacturers to submit a complete clinical evidence dossier within 100 days of the official scope being issued by the Health Technology Assessment Coordination Group (HTACG). The official scope defines the populations, interventions, comparators, and outcomes (PICO) against which the product will be assessed. Because Member State variation in standard of care frequently produces multiple PICO sets, manufacturers who wait for the formal scope face severe time pressure for systematic literature reviews (SLRs) and indirect treatment comparisons (ITCs).

This white paper presents a human-in-the-lead artificial intelligence framework for anticipatory PICO scoping, developed by operationalising EU HTA Coordination Group guidance with structured HEOR and AI expert input. The framework pairs automated evidence processing with expert oversight and provenance-tagged outputs at every stage. A validation of the MadeAi-JCA platform against the first **three published JCA reports** (tovorafenib, lurbinectedin, and tarlatamab) demonstrated that, when supplied only with Population and Intervention, the platform captured 10 of 11 unique JCA-defined **comparators** (91 percent), all core effectiveness **outcomes** (overall survival, progression-free survival, objective response rate), **safety endpoints** and health-related quality of life instruments (100 percent), and produced fully source-traceable outputs in under one hour. Manual equivalent work typically requires three to five weeks of specialist effort.

The approach aligns with the HTACG AI principles of Accountability, Transparency, and Reporting published in July 2026. It does not replace expert judgement; it supplies a structured, auditable starting point that experts finalise. The evidence base for AI-assisted methods remains subject to recognised limitations in reproducibility and prompt dependency, reinforcing the necessity of human-in-the-lead design.

Industry Context and Problem Statement

The EU HTA Regulation entered into force for oncology products and advanced therapy medicinal products on 12 January 2025, with progressive extension to orphan medicines (2028) and all centrally authorised products (2030). The HTACG defines the assessment scope after EMA validation of the marketing authorisation application. Manufacturers must then deliver a dossier addressing every defined PICO within a fixed 100-day window.

Comparators are defined as the treatments most likely to be replaced by the new product and must reflect standard of care across participating Member States. Where practice varies, multiple comparators and therefore multiple PICO sets arise. Anticipating this scope requires synthesis of regulatory labels and EPARs, clinical guidelines (ESMO, ASCO, SIOPE, and national societies), national HTA decisions (IQWiG, HAS, NCPE, AIFA, NICE, and others), drug registries, and formularies across 27 Member States, and published systematic reviews and ITCs. Performed manually, this synthesis is a multi-week specialist HEOR task.

The first published JCA (tovorafenib, June 2026) illustrated the practical consequence of incomplete anticipation: evidence gaps appeared where Member State needs were not fully foreseen. Subsequent reports for lurbinectedin and tarlatamab confirmed the same pattern of multi-PICO complexity in extensive-stage small-cell lung cancer.

Current Challenges and Gaps

Three structural challenges dominate current practice. First, temporal compression. Evidence packages that previously took months must now be completed in 100 days after the scope is known. Starting only after the official scope is issued is operationally unfeasible for comprehensive SLRs and ITCs.

Second, information fragmentation. Standard of care and preferred comparators differ across Member States. No single published source consolidates this variation; the work requires repeated cross-referencing of heterogeneous national documents.

Third, auditability and reproducibility requirements. The July 2026 HTACG AI principles require manufacturers to retain full responsibility for accuracy (Accountability), to maintain audit trails of inputs, processes and outputs (Transparency), and to disclose AI-assisted work with sufficient detail for assessors to evaluate reliability (Reporting). Parallel methodological assessment of the broader AI evidence base underpinning related frameworks has shown that many foundational large-language-model evaluations lack full prompt disclosure, temperature reporting, repeated-run consistency testing, open code and data sharing, or pre-registration. These gaps underscore that purely automated or black-box approaches cannot yet meet regulatory-grade standards.

Proposed Solution and Approach

A human-in-the-lead AI framework for anticipatory JCA PICO scoping by synthesising EU HTA Coordination Group procedural requirements with structured input from HEOR and AI experts. The framework comprises six stages:

1. AI-assisted construction of a baseline PICO from regulatory and clinical evidence sources.
2. Identification of potential Member State variation in populations, comparators, and outcomes.
3. Consolidation of anticipated requirements into a defensible set of PICOs with documented rationale.
4. Assessment of comparator relevance, including individualised comparators.
5. Identification and prioritisation of evidence gaps across the anticipated PICOs.
6. Generation of protocol-ready outputs for SLRs and ITCs.

At each stage, expert review is required, and all outputs are provenance-tagged to preserve auditability. The design deliberately positions AI as a structured starting-point generator rather than a decision-maker. The human expert retains final responsibility for every scope decision that enters the dossier.

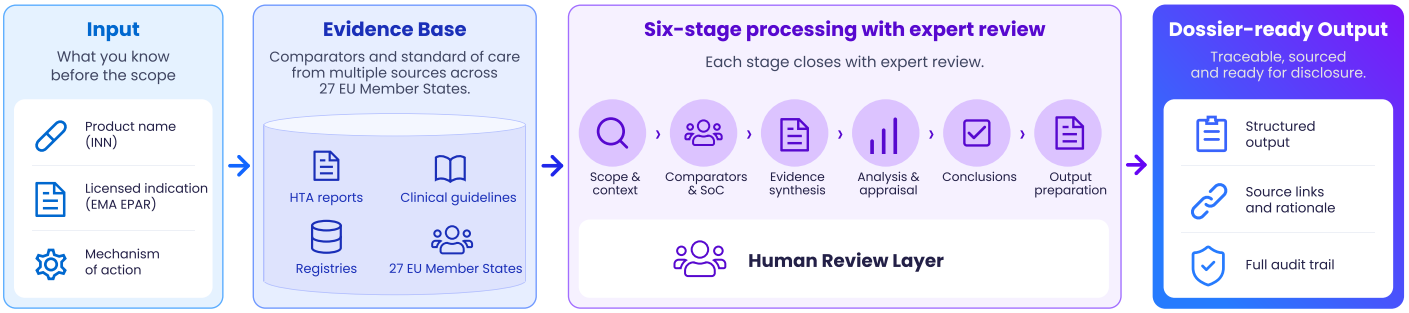
Core Features and Capabilities

MadeAi-JCA implements the framework. Given only the product name (INN), licensed indication (Population), and mechanism of action (Intervention), the platform:

- Identifies anticipated comparators and standard of care across all 27 EU Member States.
- Suggests relevant clinical effectiveness, safety, and HRQoL outcomes, including mechanism-specific signals (for example, cytokine release syndrome and ICANS for BiTE therapies; immune-mediated adverse events for checkpoint-inhibitor contexts).
- Generates anticipated PICO sets that reflect sub-population structure.
- Links every comparator, outcome, and PICO element to its exact source document with rationale.
- Produces reviewer-ready output in under one hour.

The platform therefore supports the transition from weeks of manual cross-referencing to a single automated run that covers the full Member State landscape while remaining fully traceable.

Architecture and Technical Analysis



AI-Assisted JCA Workflow: From Input to Dossier-Ready Evidence

The architecture is a pipeline with a narrow input, a wide evidence base, and a human review layer before anything reaches a dossier. The reference material describes inputs, sources, stages, review points, and traceability; it does not disclose model types, prompts, or infrastructure, and this paper does not speculate. Inputs are limited to what a manufacturer knows before the official scope; the expert review band sits between the AI stages and anything that enters a dossier.

The platform's architecture is designed to transform limited product information into structured, traceable JCA scoping outputs. The workflow begins with three inputs: the product name (INN), licensed indication from the EMA EPAR, and mechanism of action, without requiring trial data, evidence packages, or previous HTA decisions, as demonstrated by Bajpai et al. (2026).

A multi-source evidence layer then retrieves and synthesizes information on comparators and standards of care from HTA reports, clinical guidelines, and drug registries across all 27 EU Member States. This information feeds into a six-stage processing workflow, with expert review integrated at each stage to validate and refine the generated outputs. The architecture's emphasis on human oversight is further detailed in the research by Angeline Babitha et al. (2026).

A traceability layer connects each output to its supporting source and rationale while recording inputs, processing steps, and review decisions. Together, these components create a structured, auditable workflow that supports evidence preparation and documentation aligned with HTACG guidance.

The architecture stops where published evidence stops: assessor-defined composites and interval thresholds come from expert input, so the review layer is structural, not optional.

Use Cases and Implementation Considerations

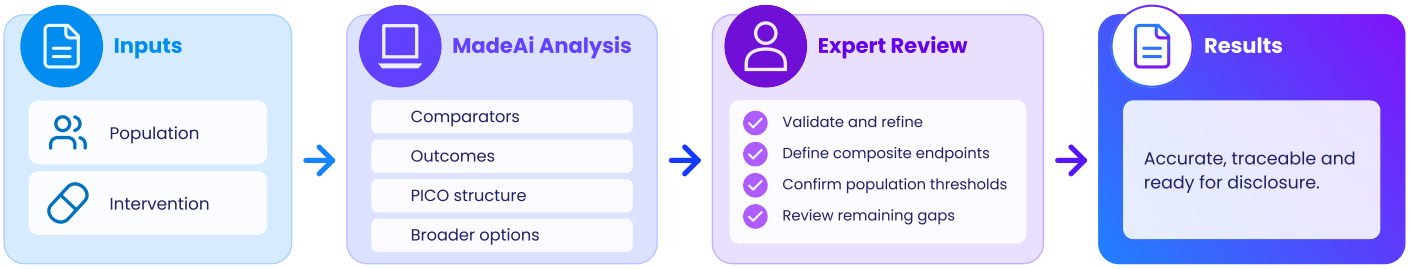
The primary use case is early JCA preparation for oncology and ATMP products now subject to the regulation, and subsequently for orphan and all centrally authorised products. Manufacturers can run the platform as soon as the licensed indication is known, obtain a structured anticipated scope, and begin SLR protocol development and evidence generation months before the official HTACG scope is issued.

Implementation requires:

- An HEOR analyst or clinical expert to operate the platform and perform the mandatory review at each decision point.
- Governance processes that ensure every AI-generated element is reviewed, edited if necessary, and formally owned by the manufacturer before inclusion in the dossier.
- Documentation of the AI assistance used, including performance characteristics from validation, so that the Reporting principle can be met.

The framework is intended to integrate with downstream SLR, data extraction and evidence-synthesis activities within a single platform environment, reducing tool, and vendor fragmentation.

Research Findings and Credible Industry Examples





MadeAi-JCA Validation: Key Findings and Performance Results

A retrospective accuracy validation used the first three published JCA reports as ground truth: tovorafenib (paediatric low-grade glioma with BRAF alteration), lurbinectedin (ES-SCLC first-line maintenance), and tarlatamab (ES-SCLC second-line and beyond). The platform received only Population and Intervention inputs.

Comparator Results:

- **Tovorafenib:** 6 of 7 JCA-defined unique comparators identified; the single miss (everolimus) was an edge-case individualised option restricted to SEGA/TSC, outside the BRAF-altered indication. Five additional clinically real options were surfaced.
- **Lurbinectedin:** 1 of 1 primary comparator identified; one additional valid alternative (durvalumab) surfaced.
- **Tarlatamab:** 3 of 3 core comparators identified; five additional real-world or trial-control options surfaced.
  Overall: 10 of 11 unique JCA comparators captured (91 percent). No clinically critical comparator was missed.

Outcome Results:

- Core effectiveness outcomes (OS, PFS, ORR) and duration of response identified for all three products (100 percent).
- Full safety bundle (grade ≥3 adverse events, serious adverse events, discontinuations) identified for all three; mechanism-specific toxicities (CRS/ICANS for tarlatamab; immune-mediated events for lurbinectedin) additionally surfaced.
- Disease-specific and generic HRQoL instruments correctly named for each indication (100 percent).
- Assessor-defined composite endpoints (symptomatic disease control, PFS2, complete response rate) were only partially captured, reflecting a boundary that requires human clinical definition rather than derivation from published sources.

PICO Structure:

- Primary PICO content and sub-population logic correctly identified in all three cases.
- Lurbinectedin matched the published single-PICO scope exactly.
- For the more complex multi-population cases (tovorafenib and tarlatamab), population labelling differed in terminology but remained clinically equivalent; final labelling requires subject-matter expert confirmation.

All outputs were generated in under one hour with 100 percent source traceability. The same manual exercise is reported to require three to five weeks of specialist HEOR effort, corresponding to an acceleration of more than 200 times.

These findings align with the human-in-the-lead design: the platform provides a complete, structured starting point; the expert finalises population thresholds, composite endpoint definitions, and any remaining gaps.

Quantifiable Benefits

Documented performance metrics from the validation are:

- **Scope output time:** under 1 hour versus 3–5 weeks manually (greater than 200-fold acceleration).
- **Comparator completeness:** 91 percent of unique JCA-defined comparators (10 of 11).
- **Core outcome coverage:** 100 percent for OS, PFS, ORR, safety endpoints, and HRQoL.
- **Source traceability:** 100 percent of outputs linked to exact source documents.
- **Coverage of 27 EU Member States** in a single automated run.

These figures are derived solely from the three-product validation against published JCA reports and should be interpreted within that limited sample.

Limitations



AI-Assisted JCA: Limitations, Assumptions, and Key Risks

Limitations explicitly acknowledged in the reference materials include:

- Sample size of three products, all in oncology. Generalisability to other therapeutic areas remains untested.
- The ground truth is the final published JCA scope, which incorporates manufacturer input and expert consultation after the initial scope. This is a higher bar than pure pre-scope anticipation from published sources alone.
- Population labelling and assessor-defined composite endpoints require human finalisation; the platform does not claim to replicate these exactly.
- The broader AI evidence base still exhibits gaps in prompt disclosure, temperature reporting, consistency testing and open sharing, as characterised in related methodological assessments. These gaps reinforce the necessity of human

Assumptions underlying the approach are that published regulatory, guideline and HTA sources contain sufficient signal to anticipate the majority of clinically critical comparators and outcomes, and that a trained HEOR reviewer can efficiently filter additional surfaced items.

Primary risks include failing to anticipate rare, individualised comparators or newly defined composite endpoints, and over-reliance on AI output without expert review. Mitigation is built into the design: mandatory human review at every stage, full source provenance, explicit disclosure of AI assistance, and positioning of the platform as a starting point rather than a finished scope. Manufacturers remain fully accountable for the final dossier content.

Security, Compliance, and Ethical Considerations

The framework and platform are designed to meet the three HTACG AI principles of July 2026:

- **Accountability:** the human expert team retains full responsibility; AI produces a structured draft only.
- **Transparency:** every output is source-linked with rationale; full audit trails of inputs, processes, and outputs are maintained.
- **Reporting:** the validation study itself supplies empirical performance evidence that can be disclosed in the dossier.

No claims are made regarding data security architecture or specific regulatory certifications beyond the stated alignment with HTACG principles. Ethical use requires that AI assistance is never presented as a substitute for clinical or methodological judgement and that limitations are disclosed transparently to assessors.

Actionable Recommendations and Conclusion

Manufacturers preparing for JCA should:

1. Initiate anticipatory PICO scoping as soon as the licensed indication is known, using a human-in-the-lead AI workflow that supplies source-traceable drafts.
2. Mandate expert review and formal ownership of every AI-generated element before it enters evidence-generation plans or the dossier.
3. Document AI use, including validation performance characteristics and known limitations, to satisfy the Reporting principle.
4. Prioritise completeness of clinically critical comparators and outcomes over classical precision metrics, given the asymmetric cost of false negatives.
5. Integrate anticipatory scoping with downstream SLR and ITC protocol development to maximise the value of the 100-day window.
6. Monitor the evolving HTACG guidance and continue to demand higher standards of reproducibility and auditability from all AI tools used in regulatory evidence generation.

The combination of automated evidence synthesis across 27 Member States with mandatory human oversight at predefined decision points offers a practical route to JCA readiness. Validation against the first three published reports shows that a carefully designed human-in-the-lead system can deliver reviewer-ready, fully traceable anticipated scopes in under one hour while meeting the accountability, transparency, and reporting expectations now required by the HTACG. Future work should extend validation to additional therapeutic areas and continue to strengthen the methodological maturity of the underlying AI evidence base.

MadeAi

1 Broadway, CIC, Cambridge, MA 02142
+1 8006749760 | info@madeai.com
www.madeai.com