Artificial Intelligence Tools for PICO Prediction: A New Reality or a Future Dream?

HTA108

Jaiswal H[1], Rolska-Wójcik P[2], Delaitre-Bonnin C[3]

[1]Evidera Ltd., Part of PPD, Thermo Fisher Scientific, Mumbai, India; [2]Evidera Ltd., Part of PPD, Thermo Fisher Scientific, Warsaw, Poland; [3]Evidera Ltd., Part of PPD, Thermo Fisher Scientific, Paris, France

Background

- The new European Regulation on Health Technology Assessment (EU HTAR) will apply starting January 2025. For Joint Clinical Assessment (JCA), a large number of Populations, Interventions, Comparators, and Outcomes (PICOs) may be expected.

Objectives

- This study aims to conduct an efficient and comprehensive PICO prediction exercise for an approved oncology Product X by manual PICO scoping and testing the applicability of publicly available generative artificial intelligence (AI) Tools A and B.

Methods

AI Tools and Manual Scoping

- The PICO prediction exercise was conducted through assessment of product profile, analysis of European and local guidelines at the time of EMA registration, and competitive landscape assessment. Analysis was executed for Europe, focusing on France, Germany, and Italy. This approach was followed for generative AI tools and the manual PICO scoping exercise in addition to specific tasks mentioned below. A list of PICOs obtained from AI tools and manual scoping have been compared (**Figure 1**).

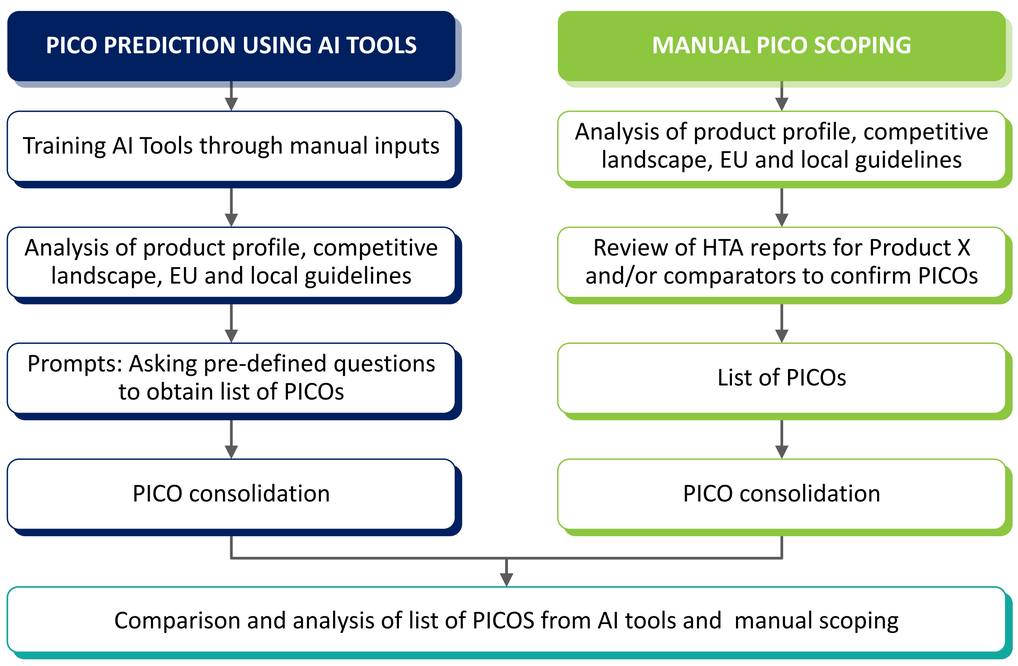
AI Tools

- Training prompts about PICO framework and PICO prediction were developed.
- Pre-defined questions were asked to check if one of the pre-trained AI tools could be used for PICO prediction.
- Several prompts were tested for the PICO prediction exercise. Key points asked during the exercise were:
  - To share list of PICOs.
  - To identify relevant subpopulations and comparators.
  - To specify geographic scope (Region: Europe; Countries: France, Germany, Italy).
  - To identify relevant European treatment guidelines.
  - To identify country-specific guidelines.
  - To attach country-specific guidelines as a portable document format (PDF) file to support PICO prediction.

Manual PICO Scoping

- Health technology assessment (HTA) reports for Product X were also reviewed to confirm PICOs.

Figure 1. Methodology



Abbreviations: AI = artificial intelligence; HTA = health technology assessment; PICO = Population, Intervention, Comparator, and Outcome

Results

Manual PICO Scoping

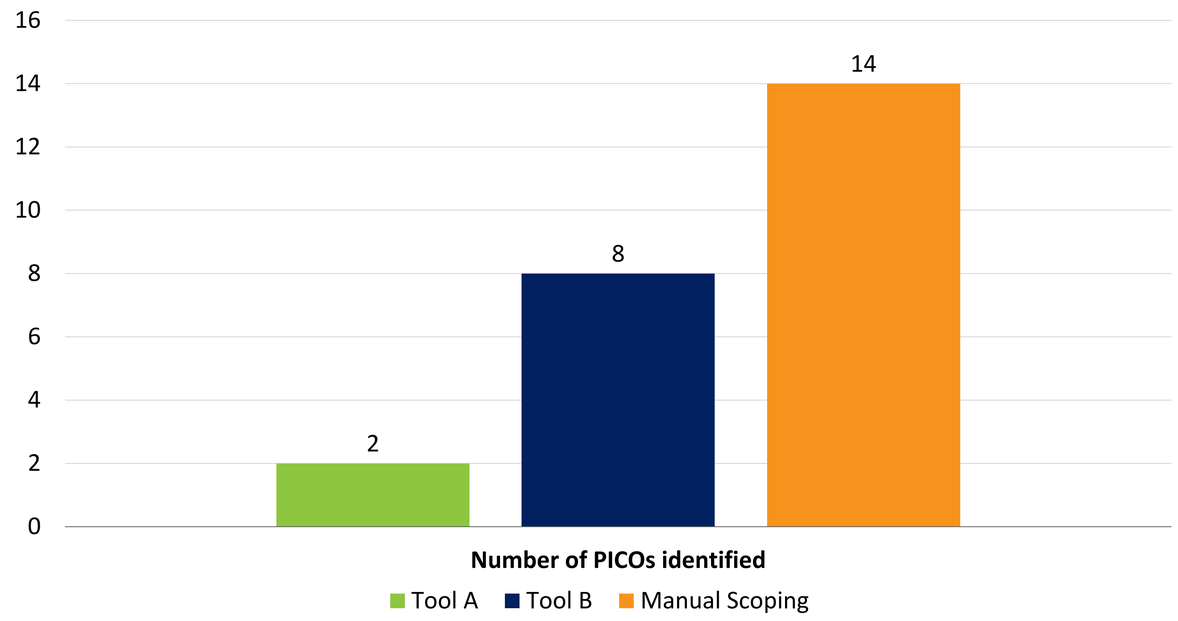
- Manual scoping resulted in 14 consolidated PICOs (removed duplicates only) based on European and local guidelines.
- A thorough selection and review of sources resulted in a comprehensive PICO prediction and consolidation exercise.

AI Tools

- Tools A and B were trained on PICO framework and PICO prediction.
- AI tools often used European guidelines for country-specific requirements, resulting in an inconsistent number of PICOs.
- Identification of correct local guidelines, their translation, and identification of comparators were challenging.
- Testing HTA reports added another layer of complexity.

Results (cont.)

Figure 2. Number of PICOs Identified by AI Tools vs. Manual Scoping



Abbreviations: PICO = Population, Intervention, Comparator, and Outcome

Tool A

- Tool A identified two PICOs and had difficulties identifying relevant comparators.
- Identified subpopulations were based on line of therapy. However, Tool A was unable to identify comparators, stating: “Not applicable since it’s a monotherapy.”
- Tool A was restricted to attaching only one PDF for support, limiting the research.

Tool B

- Tool B resulted in only eight PICOs based on provided guidelines.
- Results for subpopulation were inconsistent. Some responses were based on line of therapy whereas others were based on age, comorbidities, or underlying conditions.
- For comparators, Tool B sometimes mentioned patient group instead of a medicinal product. Tool B also grouped a class of drugs as a single comparator.
- Tool B temporarily restricted access to the tool’s latest version.

Comparison and Analysis

- Manual scoping resulted in a more comprehensive PICO prediction and consolidation exercise, whereas varying results from generative AI Tools A and B highlighted their limitations (**Figure 2**).

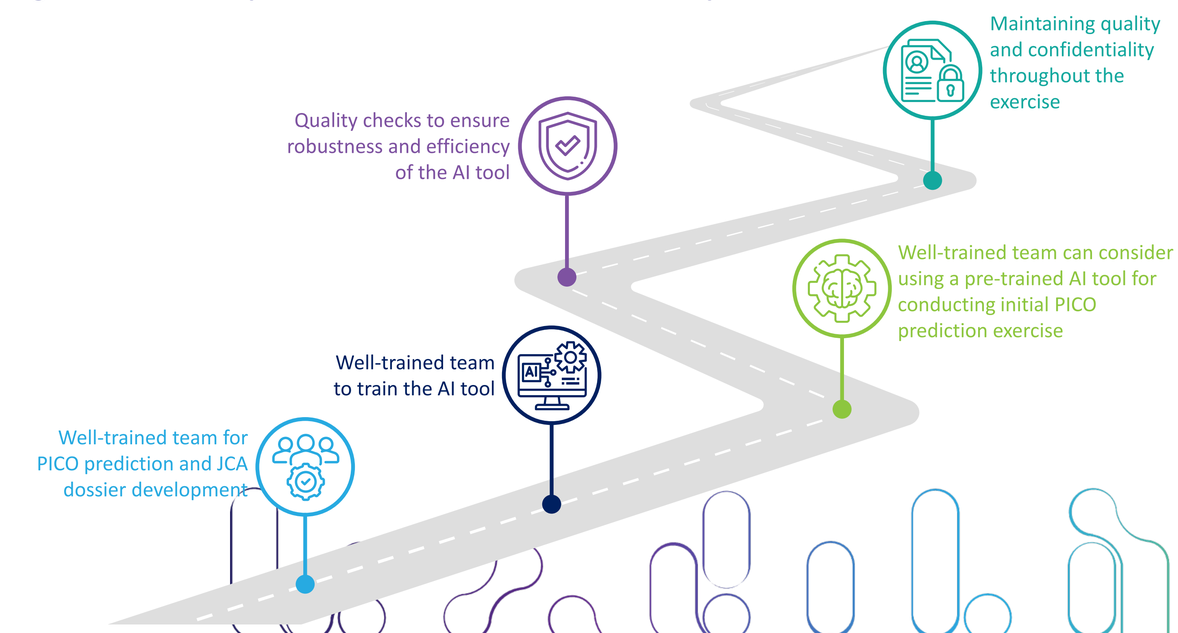
Discussion

- The importance of a robust PICO prediction exercise is exponentially increasing with the introduction of the new EU HTAR. While manual scoping can result in a robust PICO prediction and consolidation exercise, several limitations were noted for generative AI tools.
- Generative AI tools often lack the understanding of HTA/JCA requirements, which makes it difficult to identify correct subpopulations and comparators.
- Identification of correct local guidelines, HTA reports, and their translations can also be challenging.

Conclusions

- Lack of well-trained and robust AI tools still makes human involvement necessary for an essential step such as PICO prediction.
- There is a need for a well-trained team that understands JCA requirements, European guidelines, local guidelines, and preferred PICOs by HTA bodies.
- Maintaining quality and confidentiality remains critical for such exercises.

Figure 3. Roadmap for Future Research and Development



Abbreviations: AI = artificial intelligence; JCA = Joint Clinical Assessment; PICO = Population, Intervention, Comparator, and Outcome

Presented at the ISPOR Europe Conference ● 17–20 November 2024 ● Barcelona, Spain

Funding provided by Evidera Ltd., a business unit of PPD, part of Thermo Fisher Scientific