Individual PICO Review and Hierarchy in Systematic Literature Reviews: Why Guideline Adherence Matters and How REAL-SLRs Reduce Rework

Rhiannon Campden[1], Rozee Liu[1], Arup Pramanik[2], Anna Forsythe[1]

[1]Oncoscope-AI, Miami, FL, USA; [2]Independent consultant, Boston, MA, USA

CONCLUSIONS

- Failure to adhere to hierarchical PICOS review leads to inefficiencies, errors, and repeated effort following PICOS criteria changes. Common changes to PICOS criteria include redefining intervention or outcome terms, changing a comparator intervention, changing the population, adding or removing outcomes, or changing the included study designs[1]
- REAL-SLR systems that preserve independent PICOS-level decisions enable rapid, targeted updates while maintaining methodological rigor
- The REAL-SLR approach reduces time, cost, reviewer burden, and supports more agile evidence synthesis for HTA and market access decision making

ONCOSCOPE AI
REAL-TIME ONCOLOGY LIBRARY

BACKGROUND

- Systematic literature review (SLR) guidelines specify that records can be evaluated sequentially by Population, Intervention/Comparator, Outcomes, and Study design (PICOS)
- In practice, human reviewers often use pattern recognition and keyword-based judgments
- Available artificial intelligence (AI) tools that support SLRs are trained on human review and therefore replicate these shortcuts
- Based on recent review, changes in PICOS elements occur in over 50% of published SLRs with protocols submitted to PROSPERO[1]

OBJECTIVES

- To examine non-hierarchical PICOS evaluation and demonstrate how a Real-time AI-assisted Living Systematic Literature Review (REAL-SLR) approach that preserves individual PICOS decisions can reduce errors, rework, and time loss when review criteria change
- As an illustrative case, a change in prostate cancer PICOS screening criteria was used to demonstrate the flexibility in the model to adapt to PICOS updates

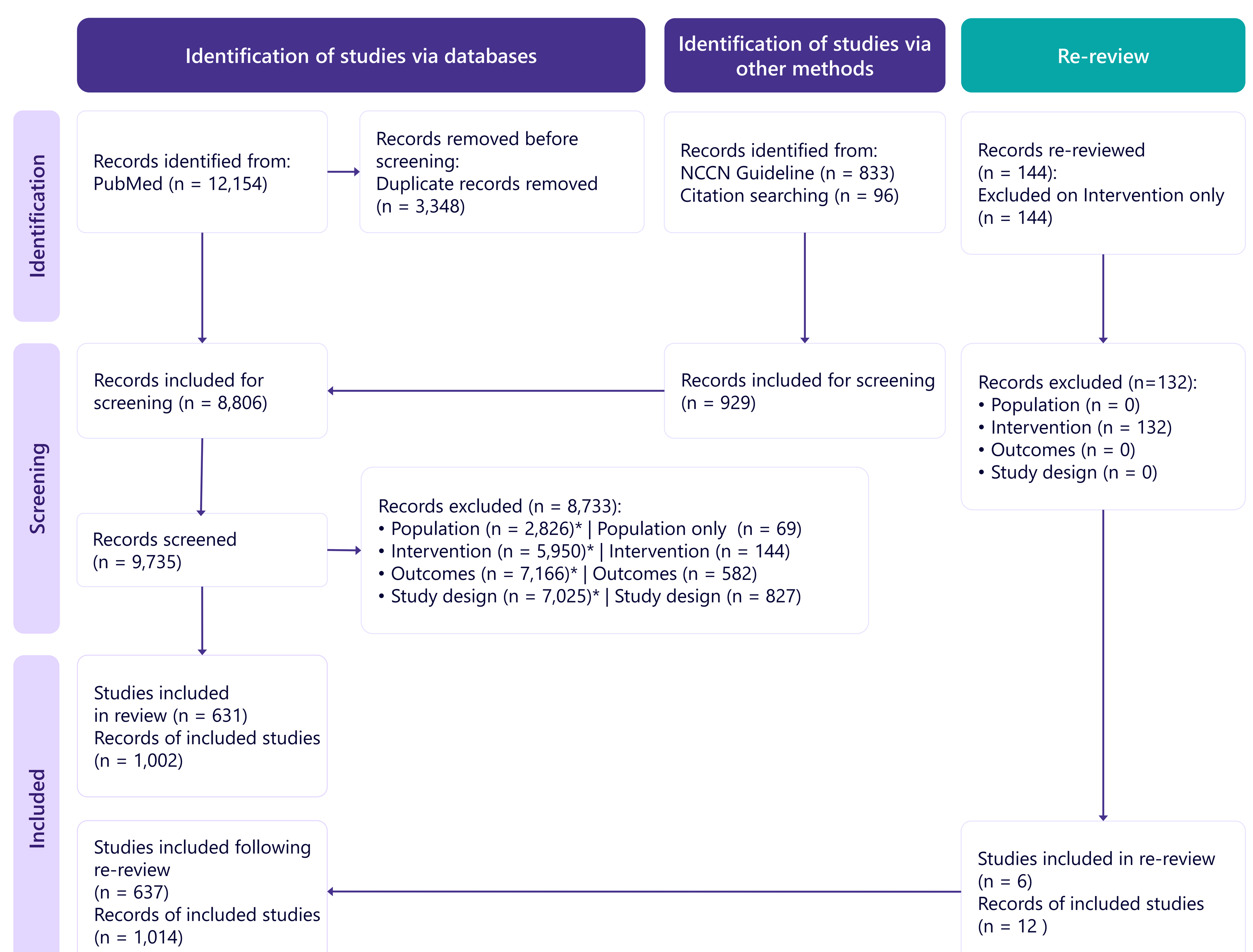
Table 1. PICOS criteria for the Prostate Cancer REAL-SLR

Element	Inclusion
Patient population	• Patients diagnosed with prostate cancer
Intervention and Comparators	• Pharmacological interventions, surgical procedures, radiotherapy, bone marrow and cell transplant and any other therapy (drug or cell) for prostate cancer • Palliative care and observation as comparators only • Active surveillance (added as a protocol update)
Outcomes measures	• Overall survival (OS) and mortality • Progression-free survival (PFS) • Other progression measures, such as time to progression (TTP) or time to treatment failure (TTF), or metastases free survival (MFS). • Response rate (including objective response [ORR], and other response) • Quality of life • Safety / toxicity (including adverse events [AEs] and discontinuations)
Study design	• Randomized controlled trials (RCT) • Single-arm and non-randomized trials • Externally controlled trials (ECT) • Pooled analyses of interventional studies
Restrictions	• English language

METHODS

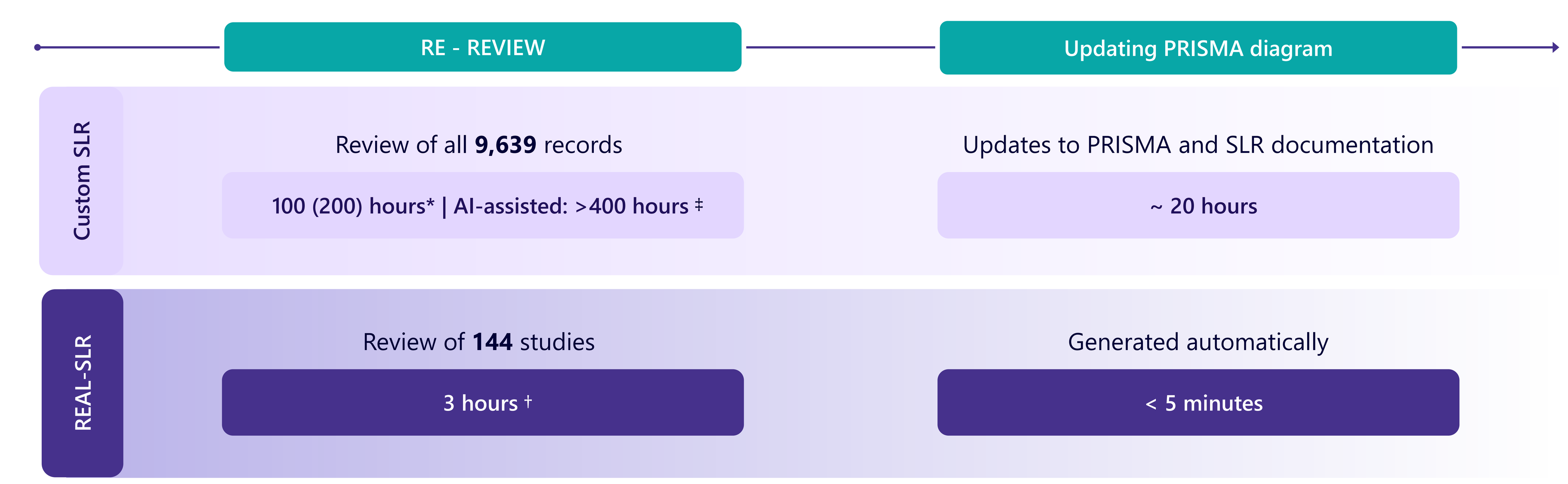
- Daily searches were conducted in PubMed starting September 29, 2025 using prostate cancer and clinical trial terms with observational studies excluded
- The database was supplemented with a review of NCCN and clinicaltrials.gov references
- The initial SLR scope included interventional studies of pharmacology interventions for prostate cancer
- The REAL-SLR agentic AI-model was trained to deliver individual PICOS decisions using prompt engineering and was validated against human review (See MSR18, MSR22)
- Studies were evaluated for inclusion/exclusion against each PICOS criterion (**Table 1**) individually, with decisions for each recorded
- Following a scoping review, active surveillance terms were added to the PICOS criteria in addition to pharmacological interventions for prostate cancer
- Records excluded by "intervention" were then rereviewed using the REAL-SLR agentic AI-model followed by human review
- The REAL-SLR approach was compared with traditional SLR workflows both with and without large language model (LLM) assistance

Figure 1. PRISMA diagram for the Prostate REAL-SLR following addition of active surveillance intervention terms to PICOS criteria



*Reasons for exclusion were not mutually exclusive

Figure 2. Workflow estimates to update PICO criteria in a Traditional SLR vs. REAL-SLR



*estimates based on single (dual) review at a speed of 100 abstracts/hour
†estimates based on REAL-SLR AI screening times of 0.5 seconds/variable followed by human review of 100 abstracts/hour
‡AI-assisted review based on published AI-screening estimates[2-6]

RESULTS

- The original SLR contained 1,014 records as of October 27, 2025 which were screened on individual PICOS terms with decisions retained for each PICOS criteria, independent of each other. Of those, 144 were excluded based on intervention alone (**Figure 1**)
- Following the introduction of "active surveillance" as an included intervention term, the 144 studies were rescreened and 132 were excluded
- After re-screening, 12 new records on active surveillance were included for 6 new studies
- Using REAL-SLR, automated updates to the evidence set, the PRISMA flow diagram, and reports were generated, and even with 100% human quality control, the total effort required was 3 hours
- The REAL-SLR method was compared with workflows of a traditional SLR, where reasons for PICOS decisions are not documented or retained
- Using traditional SLR methodology, all 9,639 records would need to be rescreened, resulting in 100 hours of single review or 200 hours of dual review
- Based on published LLM reporting 40-89% efficiencies in screening,1-5 screening times could be reduced to >40 hours by using an LLM as a first reviewer; however, the REAL-SLR model still reduced overall efforts by conserving individual PICOS decisions (**Figure 2**)

REFERENCES

1. Siebert M, Caquelin L, Madera M, Acosta-Dighero R, Naudet F, Roqué M. Assessing the magnitude of changes from protocol to publication-a survey on Cochrane and non-Cochrane Systematic Reviews. *PeerJ.* 2023;11:e16016. Published 2023 Oct 2. doi:10.7717/peerj.16016
2. Hamel C, Kelly SE, Thavorn K, Rice DB, Wells GA, Hutton B. An evaluation of DistillerSR's machine learning-based prioritization tool for title/abstract screening - impact on reviewer-relevant outcomes. *BMC Med Res Methodol.* 2020;20(1):256. doi:10.1186/s12874-020-01129-1
3. Gates A, Johnson C, Hartling L. Technology-assisted title and abstract screening for systematic reviews: a retrospective evaluation of the Abstrackr machine learning tool. *Syst Rev.* 2018;7(1):45. doi:10.1186/s13643-018-0707-8
4. Chai KEK, Lines RLJ, Gucciardi DF, Ng L. Research Screener: a machine learning tool to semi-automate abstract screening for systematic reviews. *Syst Rev.* 2021;10(1):93. doi:10.1186/s13643-021-01635-3
5. van de Schoot R, de Bruin J, Schram R, et al. An open source machine learning framework for efficient and transparent systematic reviews. *Nat Mach Intell.* 2021;3:125-133. doi:10.1038/s42256-020-00287-7
6. Przybyła P, Brockmeier AJ, Kontonatsios G, et al. Prioritising references for systematic reviews with RobotAnalyst: A user study. *Res Synth Methods.* 2018;9(3):470-488. doi:10.1002/jrsm.1311

Presented at ISPOR 2026 | May 17-20, 2026 | Philadelphia, PA, USA | Poster Session 5

See more Oncoscope AI publications