## Automated Data Extraction Using Artificial Intelligence to Accelerate Systematic Literature Reviews in Rheumatoid Arthritis

von Hein, M<sup>\*</sup>, Talbot-Watt, N<sup>\*</sup>, Chen CH<sup>\*\*</sup>, Lundie G<sup>\*</sup> \* Gilead Sciences, London, UK, \*\* Virtual Science Al



- The systematic literature review (SLR) is considered the "gold standard" in evidence-based research, sitting at the top of the evidence-hierarchy pyramid.
- With the developments and significant improvements in in artificial intelligence (AI) in recent years, researchers have begun exploring the application of AI to enhance the SLR process.
- The hope is that AI will improve "the speed, rigour, transparency, and repeatability of SLRs", thus making it easier for research to summarise evidence in any given field.<sup>1</sup>
- Methods in which AI can be used to improve the SLR process span across several stages, such as



• To test the application of a custom-build AI to automatically extract data from publications to accelerate the SLR process in RA.





- search strategy and development of in- and exclusion criteria, title and abstract screening, data extraction as well as data synthesis and data abstraction.
- As one of the world's leading Health Technology Assessment (HTA) bodies, the National Institute for Care and Excellence (NICE) recently released a position statement on the use of AI in the evidence generation process, acknowledging potential benefits for this process.<sup>2</sup>
- While the use of AI for the SLR process has been assessed across multiple disease areas, no study has been published looking at rheumatoid arthritis (RA) to the knowledge of the authors.
- be included in the SLR.
- The data extraction output of the AI was compared against a human-curated data extraction grid from the 2018 SLR, using the Jaccard similarity coefficient, whereby the human-curated data extraction was assumed to be perfectly accurate.
- Additionally, the accuracy rate of the AI was evaluated by a human reviewer, whereby the reviewer checked for meaning rather than exact matches against the data extraction grid.







- The human review identified minor errors in the conventional 2018 SLR data extraction grid, questioning the perfect accuracy assumption of the data sample against which the accuracy rates were calculated.
- There was a mismatch in the total number of records identified in the conventional SLR from 2018, which identified 992 records, compared to the updated SLR from 2024, which identified 1043 records, despite using the exact same search strategy and databases.
- Furthermore, these results may be affected by training bias, textual ambiguity and variations in data formatting and presentation.
- Training an AI to automate the data extraction process for SLRs in RA is feasible but needs further work to improve accuracy rates.
- Human-supervised quality control of data extraction results remains an important aspect to ensure quality, transparency and validity of the SLR outputs.
- Al solutions to support SLR development remain fragmented, with various tools from different providers being required to cover different aspects of the SLR process, such as search strategy formulation, record screening, data extraction or data synthesis.
- Further research is required to consolidate such fragmented tools into one universal solution that researchers can use to accelerate their SLR process.

1 Atkinson, Cameron F. "Cheap, quick, and rigorous: Artificial intelligence and the systematic literature review." Social Science Computer Review 42.2 (2024): 376-393.

2 NICE 2024. Use of AI in evidence generation: NICE position statement. Available from: <u>https://www.nice.org.uk/about/what-we-do/our-research-work/use-of-ai-in-evidence-generation--nice-position-statement</u>

3 Niwattanakul, Suphakit, et al. "Using of Jaccard coefficient for keywords similarity." Proceedings of the international multiconference of engineers and computer scientists. Vol. 1. No. 6. 2013.