Assessing the Feasibility of Applying Natural Language Processing for Systematic Literature Reviews: A Case Study in Non-Small-Cell Lung Cancer

Nicola Rath[1], Harrison Walne[1], Emilie Louvet[1], William Dunlop[1], Bengt Liljas[2]

[1]AstraZeneca, Cambridge, UK; [2]AstraZeneca, Gaithersburg, MD, USA

MSR112

Introduction

- Many health technology assessment (HTA) organisations require systematic literature reviews (SLRs) as part of reimbursement submissions.[1]
- However, as the literature grows, SLRs are increasingly challenging to perform manually.
- In this study, we used natural language processing (NLP) – a rapidly advancing field of artificial intelligence (AI) – to conduct an SLR.
  - Based on its capacity to accurately filter large quantities of data, NLP has previously been successfully employed in disciplines, such as radiology (to identify injuries within reports and classify the latter), cardiology (to extract specific diagnoses from reports), and clinical trials (to identify specific cohorts), amongst others.[2–5]

Objective

- The objective of this study was to assess the feasibility of applying NLP to SLRs by comparing the results of this approach to a previous SLR for non-small-cell lung cancer (NSCLC) that was conducted entirely manually (i.e., a "fully human" approach).

Natural Language Processing and the Linguamatics I2E Application

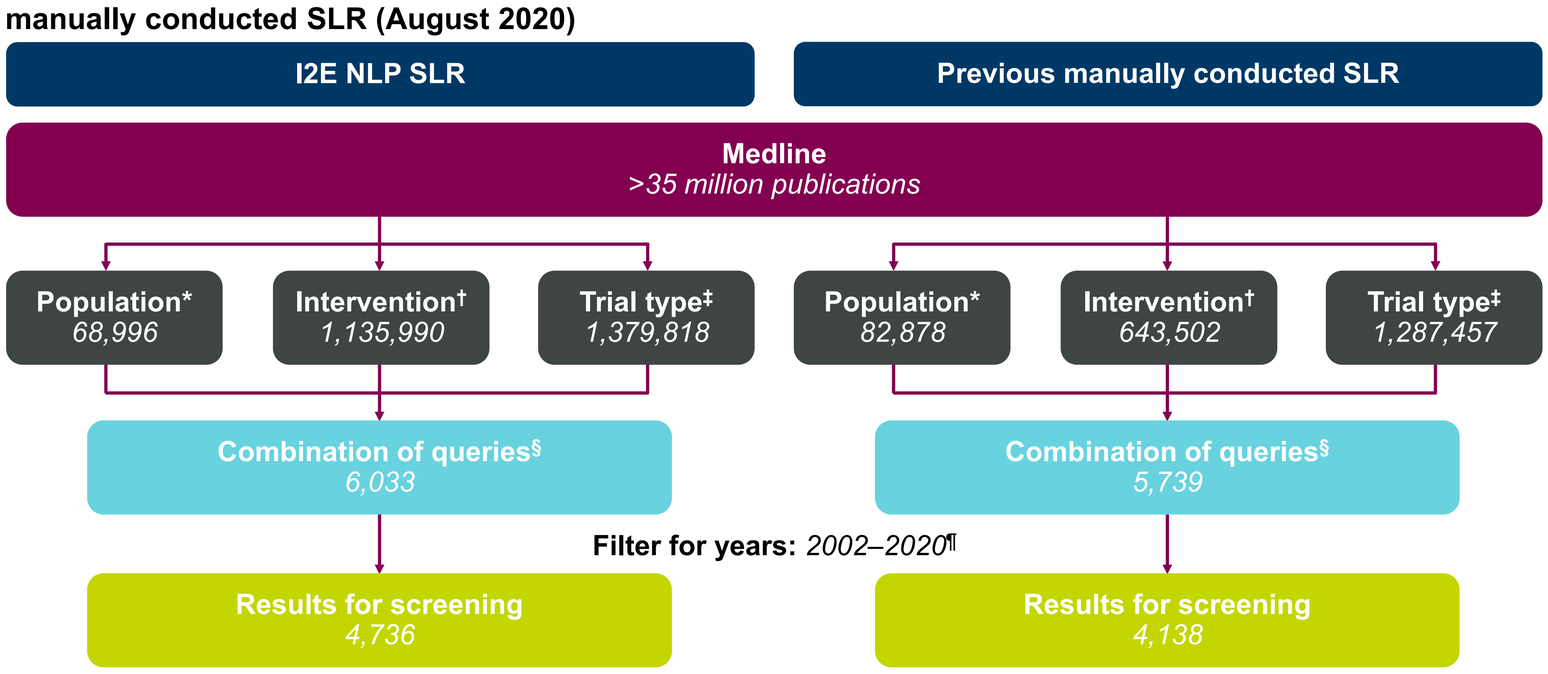
- NLP enables computers to comprehend, interpret, and generate natural language.[6,7]
- The NLP process uses a broad set of techniques and algorithms that, through deep understanding of the context of, and learnings from, a given text, enable it to perform tasks, such as:[6,7]
  - extract information
  - classify documents
  - answer questions
  - summarise text.
- A primary benefit of using NLP for an SLR is efficient information retrieval, which helps the user to quickly navigate content from external data sources and significantly reduce manual effort.
  - In addition, NLP improves the accuracy of information retrieval by reducing the risk of human error.[2]
- In this study, we used the I2E application from Linguamatics (an IQVIA company), which uses NLP technology and allows for the creation of transparent rule-based or keyword search queries for information retrieval.
- Such an approach is similar to that used for standard SLRs but with the added benefit of using interactive text mining and data extraction from unstructured text.
- Queries are often simple in nature and explainable, utilising certain keywords or patterns within unstructured text to aid transparency.
- Moreover, precision and recall can be adjusted by leveraging ontologies and broadening search criteria.
- The interface is also user-friendly, so non-experts can develop and understand their queries using a drag-and-drop interface.

Methods

- A literature search was conducted in Medline in February 2023 using the I2E NLP application.
- Search strings, inclusion/exclusion criteria, and date restrictions were applied from a prior SLR for Stage III NSCLC, conducted manually in August 2020 in accordance with key HTA guidelines.
- The initial search parameters (queries) were based on population, intervention, and trial type.
- A combination of all three parameters was filtered by year, ranging from 2002 to 2020.
- The resulting publications were further narrowed down using NLP to specifically identify *cCRT [concurrent chemoradiotherapy] therapies* and *maintenance therapies* mentioned in the results sections of the SLR.
- This study assessed the extent to which the NLP-based search could replicate the original manually conducted search and reduce the extent of human efforts required.
- The sensitivity of the NLP-based search was adjusted to identify all of the studies in the original SLR.

Results

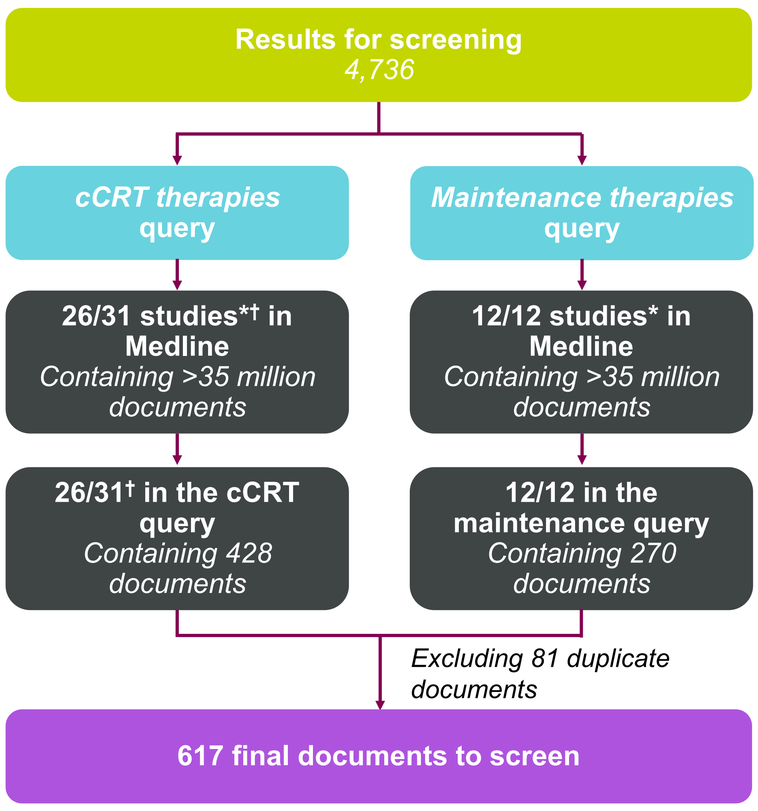
Figure 1. Comparison of document outputs from the I2E NLP SLR (February 2023) with the previous manually conducted SLR (August 2020)



*Query is a combination of indexed and free text search terms for NSLC, including proximity operators to combine non-small-cell type with lung and tumour search terms.
†Query of exploding indexed search terms for therapeutic interventions and comparators, including generic and brand names.
‡Query is a combination of indexed and free text search terms for clinical trials, including clinical trial characteristics.
§Query is a combination of all aforementioned queries.
¶Query is a combination of all aforementioned queries, that were subsequently fitted to the date range.

- The Medline literature search was reproduced with acceptable accuracy using the I2E NLP application by applying equivalent search terms and ontologies used in the previous manually conducted, fully human search (**Figure 1**).
  - Here, the search replicated the broad (i.e., exhaustive) query of the manually conducted SLR by using the I2E NLP platform, which uses synonyms and ontologies to enable easier setup of literature search queries.
- Following the initial search, the I2E NLP application was used to allow more specific queries to be used, thereby further narrowing down the results and reducing the number of irrelevant documents (**Figure 2**).
  - Here, the I2E NLP was applied to narrow down the number of hits by using semantic/syntactic analysis of text in addition to ontologies (e.g., finding keywords in a specific context or looking for a specific sentence structure).
  - The documents were narrowed down for manual screening by 87% (from 4,736 to 617), while identifying all 38 unique studies that had been identified in the original manually conducted SLR.

Figure 2. I2E NLP enabled narrowing of search results



*Compared to the previous manually conducted SLR.
†Five conference abstracts were not part of the Medline library.

- No references were lost between the I2E NLP search and manually conducted search.
  - The five discrepancies in the *cCRT therapies* query were conference abstracts, which were not part of the Medline library.
- Further narrowing of the cCRT NLP query to 118 documents led to a loss of references.
  - One reference was excluded because it did not mention any drugs in the abstract.
  - Four references were excluded because they did not mention radiotherapy in the abstract.

Conclusions

- This study provides support that NLP can be used to substantially reduce SLR screening turnaround and human workload with promising accuracy.
  - Here, the number of documents selected for manual screening was decreased by 87%, while still identifying all 38 unique studies from the original manually conducted SLR.
- Human supervision is still expected to be required with the current technology for SLRs, particularly for the final suitability screening of narrowed down documents after applying NLP, analysis, and report writing.
- One limitation of this study was that the results of the human search were known before the NLP search was applied.
  - A next step would be to compare the manual versus NLP approach, which would involve rerunning the existing NLP approach in parallel to a manually conducted search, for a literature search update (i.e., any literature after August 2020).
- Data extraction is another area where NLP, or AI more generally, could assist the SLR process.
- Given these developments, HTA organisations should consider reviewing their policies regarding the acceptance of AI for SLRs.
  - To help maintain the transparency of NLP when engaging with HTA organisations, the search terms, general approach, and platform used could be reported.

References

1. Fernandez MM, et al. *Value Health*, 2018;**21**:S108.
2. Olthof AW, et al. *Comput Methods Programs Biomed*, 2021;**208**:106304.
3. Tan WK, et al. *Acad Radiol*, 2018;**25**:1422–32.
4. Johnson SB, et al. *AMIA Jt Summits Transl Sci Proc*, 2018;**25**:104–12.
5. Chen L, et al. *J Am Med Inform Assoc*, 2019:**26**:1218–26.
6. Hirschberg J & Manning CD. *Science*, 2015;**17**:261–6.
7. Zheng Y, et al. *JMIR Diabetes*, 2022;**7**:e34681.

Poster

Please scan this Quick Response (QR) code with your smartphone camera or app to obtain a copy of these materials. Alternatively, please click on the link below.
https://bit.ly/3QbaMCL

Acknowledgements
Medical writing support, under the direction of the authors, was provided by Werner Gerber, PhD, (Fourways, South Africa) of Ashfield MedComms, an Inizio company. The study was funded by AstraZeneca.

Presented at ISPOR Europe; Copenhagen, Denmark; 12–15 November 2023