

# A Hybrid Approach to Semi-Automate the Screening Process for Living Systematic Reviews and Meta-Analysis

Huan He, PhD <sup>1\*</sup>, Irbaz Bin Riaz, MD, MS <sup>2,4\*</sup>, Syed Arsalan Ahmed Naqvi, MBBS <sup>3</sup>  
Rabbia Siddiqi, MBBS<sup>3</sup>, Noureen Asghar, MBBS <sup>3</sup>

M. Hassan Murad, MD, MPH <sup>4</sup>, Hongfang Liu, PhD <sup>1</sup>

<sup>1</sup>Department of AI and Informatics Research, Mayo Clinic, Rochester, MN, USA

<sup>2</sup>Department of Oncology, Mayo Clinic, Phoenix, AZ, USA

<sup>3</sup>Dow University of Health Sciences, Karachi, Pakistan

<sup>4</sup>Mayo Clinic Evidence Based Practice Center, Mayo Clinic, Rochester, MN, USA

## Abstract

*Systematic Reviews (SRs) and meta-analyses (MAs) are tools to synthesize evidence for important clinical topics. They are frequently used by decision makers as they provide precise estimates of effect for clinically important benefits and harms outcomes with associated certainty of evidence. However, when the field is rapidly evolving, laborious process of systematic reviews cannot keep pace with new evidence and SRMAs are quickly outdated. Thus, it is important to keep the systematic reviews “living”. For truly living systematic reviews (LSRs), the most laborious step of screening thousands of citations to identify few relevant studies must be automated. Thus, we propose a hybrid approach that integrates interactive web-based user interface and multiple natural language processing and machine learning based techniques to screen publication to maintain LSRs.*

## Introduction and Background

Systematic review is a synthesis technology that is widely used by researchers across many specialties <sup>1</sup>. Developing an SR generally requires screening of thousands of citations to identify a small number of studies which meet the inclusion criteria. To maintain an LSR, this becomes a long-term iterative process where new studies must be screened at a regular interval to identify new citations. Several machine learning (ML) and natural language processing (NLP) based methods have been proposed to help with the screening process <sup>2</sup>. Those methods could provide effective support and reduce the workload in one screening instance.

However, keeping a SR living requires repetitive screening, for example on a weekly basis, and thus applying these models necessitate users to learn not only the computer programming languages to use the packages, but also the modules and workflow to processing the massive data. Thus, maintaining LSRs continues to be an intensely laborious manual process. To semi-automate the screening process, we present a hybrid approach that integrates interactive web-based user interface and multiple NLP and ML based techniques to iteratively screen incoming citations and thus making the screening process efficient for LSRs.

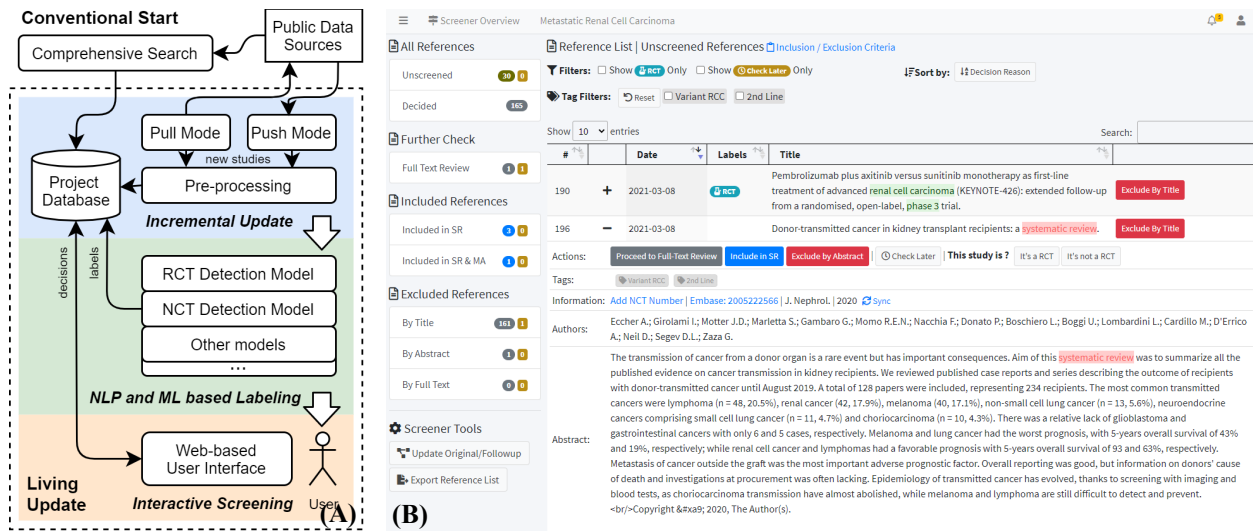
## Method

Studies or citations included in LSR are first identified using a conventional search strategy followed by an automated search strategy. Conventional search strategy is executed in multiple databases (e.g., PubMed, EMBASE, etc.) with customized keywords to retrieve the most recent literatures for a specific clinical question. At the first stage, our platform allows users to either complete the initial screening process utilizing the user-friendly interface or the results from the projects where screening has been completed can be simply uploaded (Fig.1 (B)). The main output of the first stage is the decisions for every record (i.e., included or excluded), which could be used as the input for the next stage. Since our team has conducted systematic reviews on several clinical research projects and accumulated relevant datasets, we could proceed directly to the next stage based on these datasets for each project. At the second stage, the search strategy developed in the first stage is re-used periodically based on the speed of influx for new evidence (e.g., weekly, bi-weekly, or monthly) to retrieve new records and expand the project dataset continuously.

In each small batch update, the pipeline works on the new emerging studies based on several different techniques in three steps (Fig.1(A)). We developed two modes to get new records, namely pull mode and push mode. In the pull mode, our system actively retrieves new records for each project by sending the pre-defined query requests to data sources based on the same search strategy designed in the first stage to make sure all the new coming studies meet the same criteria. The pull mode could be performed automatically at regular intervals or manually per user request. Once

---

\* Those authors have contributed equally as co-first authors



**Figure 1.** (A) The pipeline of our proposed approach. (B) The screenshot of our interactive web-based user interface, in which list the unscreened studies for screening. Users could search studies, sort list, add tags, and make decision in one place.

the new studies are pulled from the data source, our system will remove the duplicates. Then, all of the new studies are saved in the project database and set to an unscreened status. In contrast to the pull mode, the push mode passively receives new records from the data sources by using the email subscription services provided by the data sources (e.g., Ovid AutoAlert). Our system checks the email inbox regularly and generates a new study list by parsing new emails' attachments. Then, similar to the pull mode, each study will be further processed, and the new studies will be saved.

In the second NLP and ML based labeling step, various models based on NLP and ML methods are used to add labels to new studies for further screening. For example, the randomized controlled trial detect model based on support vector machine and convolutional neural network is integrated to detect whether a study is an RCT or not<sup>3</sup>; the national clinical trial (NCT) number detect model based on regular expression extracts the NCT number from study abstract; publication type model detects whether a study is original or follow-up. Those automatically generated labels will be stored in the project database as part of the meta-data and indexed by integrated search engine.

In the third interactive screening step, we present the labeled new studies through an interactive web-based user interface to assist users screening (Fig. 1(B)). In this web-based user interface, all studies and their meta-data are listed by the study status (e.g., unscreened, excluded, included, etc.). At this stage, we facilitate the screening process by an interactive design to reduce workload and improve screening efficiency. For example, the labels generated in the second step for study type (RCT or not) are displayed intuitively with each study to help users identify the study type. Furthermore, the search engine allows for filtering studies by text and the color-encoded inclusion/exclusion keywords highlighting eases the screening process. After the third step, the screened new studies have been saved in the project database with the decisions. The meta-data from already screened citations will be used to train and fine-tune the NLP model which will automate the screening process for LSR in future.

## Discussion

We iteratively designed and developed this screening system while maintaining living systematic review projects for first line treatment of metastatic kidney cancer, cancer associated thrombosis, and toxicity of immune checkpoint inhibitors. Qualitative feedback from experienced systematic reviewers suggests that our approach decreases the workload and increases efficiency for creating LSRs. The next steps include ongoing enhancements to the visual interface, implementing fine-tuned NLP models for each LSR based on labeled meta-data generated by initial screening of each project, and formal user testing.

## References

1. Bannach-Brown A, Przybyła P, Thomas J, et al. Machine learning algorithms for systematic review: reducing workload in a preclinical review of animal studies and reducing human screening error. *Systematic Reviews*. 2019;8(1):23.
2. van de Schoot R, de Bruin J, Schram R, et al. An open source machine learning framework for efficient and transparent systematic reviews. *Nature Machine Intelligence*. 2021;3(2):125-133.
3. I. J. Marshall, A. Noel-Storr, J. Kuiper, J. Thomas, and B. C. Wallace, "Machine learning for identifying Randomized Controlled Trials: An evaluation and practitioner's guide," *Research Synthesis Methods*, vol. 9, no. 4, pp. 602–614, 2018.