

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

S03: Four Miners and a Scribe

Huan He, PhD^{1*}

Irbaz Bin Riaz, MD, MS ^{2,4*}, Syed Arsalan Ahmed Naqvi, MBBS ³,
Rabbia Siddiqi, MBBS ³, Noureen Asghar, MBBS ³,
M.Hassan Murad, MD, MPH ⁴, and Hongfang Liu PhD [†]

¹ Department of AI and Informatics Research, Mayo Clinic, Rochester, MN, USA

² Department of Oncology, Mayo Clinic, Phoenix, AZ, USA

³ Dow University of Health Sciences, Karachi, Pakistan

⁴ Mayo Clinic Evidence Based Practice Center, Mayo Clinic, Rochester, MN, USA

* Those authors have contributed equally as co-first authors

He.Huan@mayo.edu, riaz.dr@mayo.edu

† Liu.Hongfang@mayo.edu

Disclosure

- I and my spouse / partner have no relevant relationships with commercial interests to disclose.

Learning Objectives

- How machine generated labels could be used in screening studies in systematic review and meta-analysis.
- User interface design for improving screening efficiency

Background - SRMA

- Systematic Review (SR) and Meta-Analysis (MA)
 - Evidence synthesis for a specific topic
 - SR is widely used by researchers across many fields
 - MA is used to get precise estimates of treatment effects
- SRMA in healthcare context
 - Clinical tests, public health interventions, social and environmental interventions
 - Outcomes, adverse events, quality of life, qualitative evidence syntheses
 - Methodological reviews, policy reviews, and economic evaluations
- ...

But conducting a SR is time-consuming. It takes, on average, 67 weeks¹ .

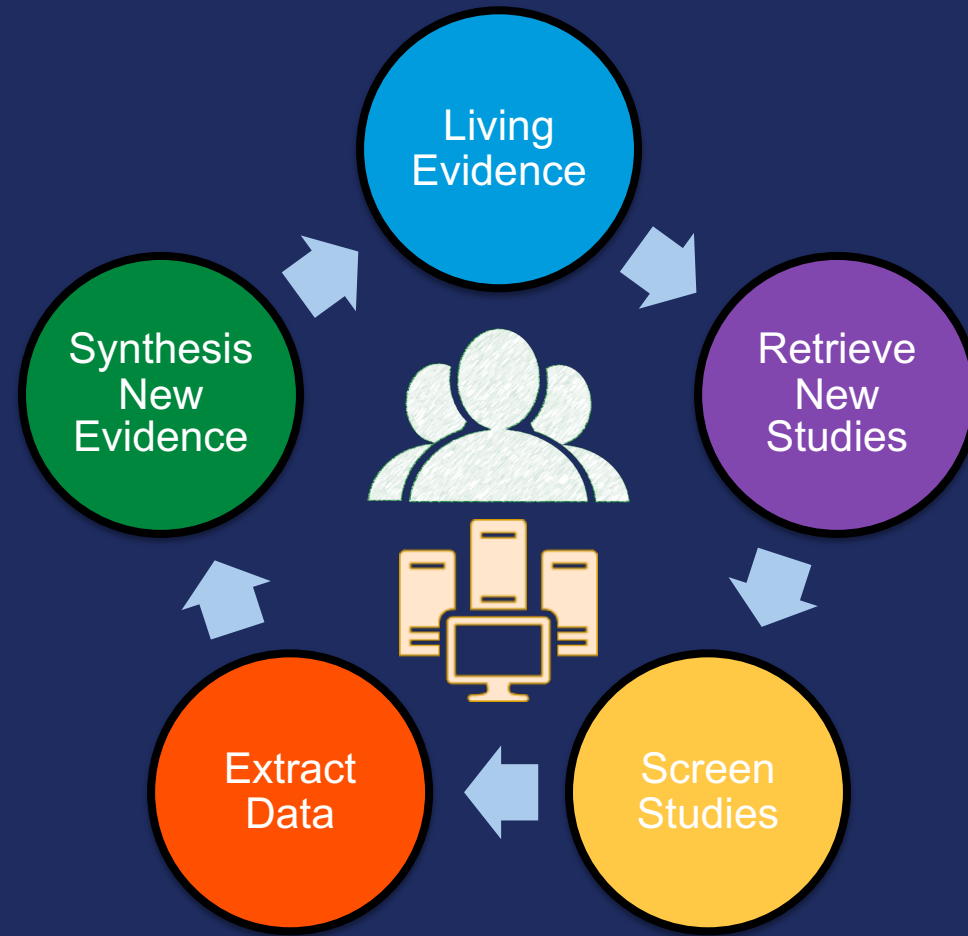
SRMA are quickly outdated if cannot keep pace with new evidence

¹ Marshall, I.J. and Wallace, B.C. (2019) Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. *Systematic Reviews*, **8**, 163.

Background - Living SRMA

To keep evidence up to date

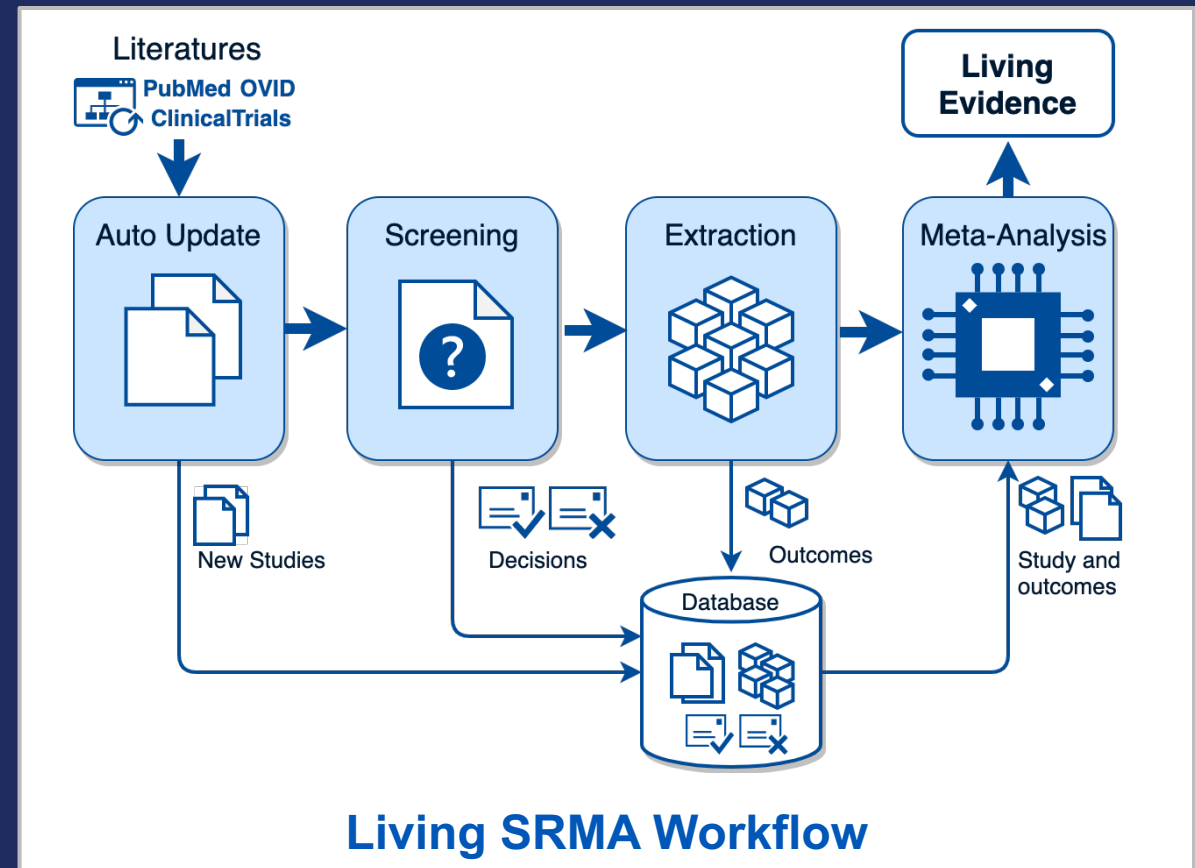
- Continually updated and incorporate new evidence when it becomes available
- Long-term iterative process
- Technical supports
 - Data sources and APIs (application programming interface)
 - Automation framework and tools
 - Machine learning (ML) and natural language processing (NLP)



Background - Challenges in Living SRMA

It is challenging to maintain a living SRMA in a rapidly evolving field

- Large size of studies
 - Many new studies arrive daily
- Intensive labor work
 - Expertise is always needed
- Complex workflow
 - Data update
 - **Study screening**
 - Data extraction
 - Meta-analysis
 - Publication



Task Analysis - Related tools and systems

Existing tools and systems

- Covidence
- RevMan
- Rayyan
- GRADEpro
- JBI-SUMARI
- EndNote
- Zotero

...

Limitations:

- Lack of fully automation
- Steep learning curve
- Massive manual efforts
- Costs of integration

Task Analysis - Improve the screening efficiency

For truly living SRMA, we must first automate the most laborious step:

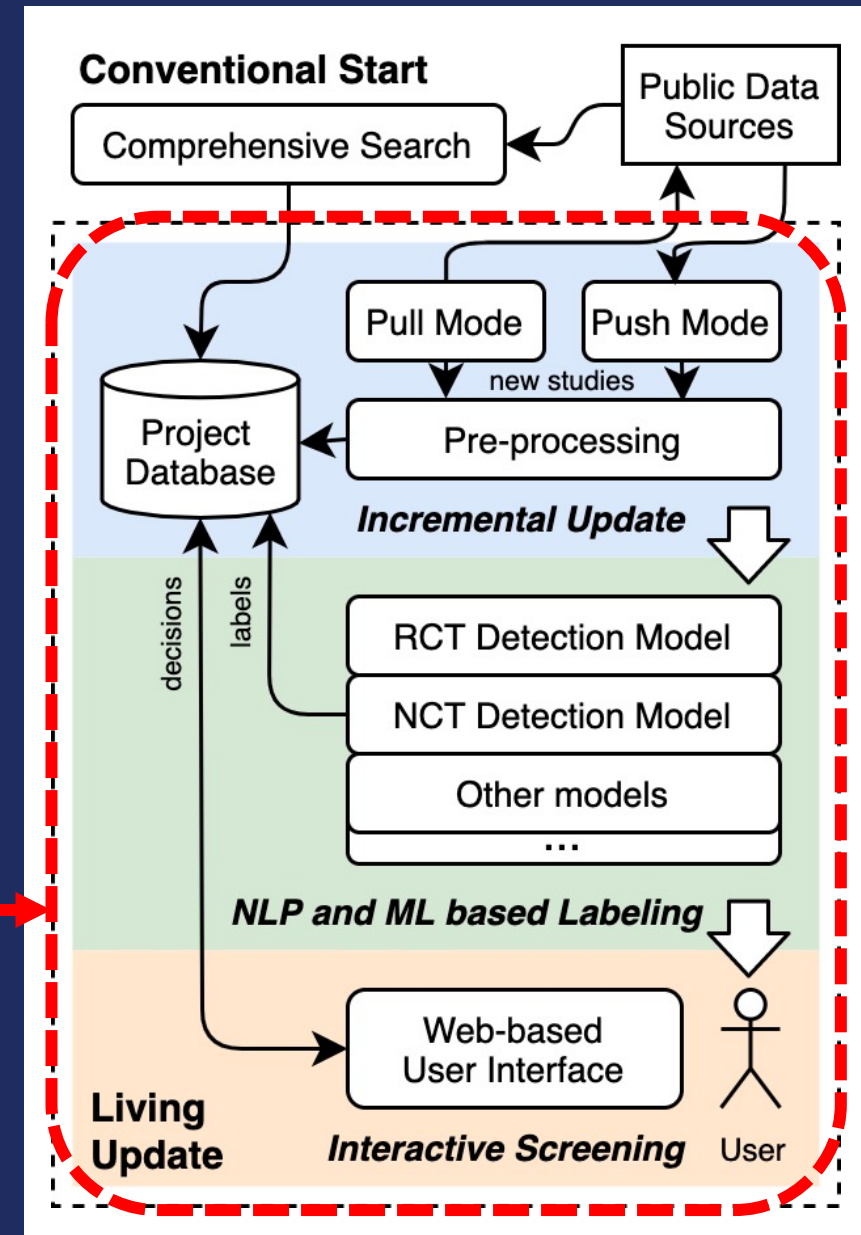
Screening thousands of citations to identify few relevant studies

Two tasks for us to improve the screening efficiency:

- Reduce the number of studies
 - The number of studies that needs human intervention
 - The number of studies in each processing batch
- Reduce labor workload
 - The cognitive load of screening
 - The physical efforts of user operation

Methods - Pipeline Overview

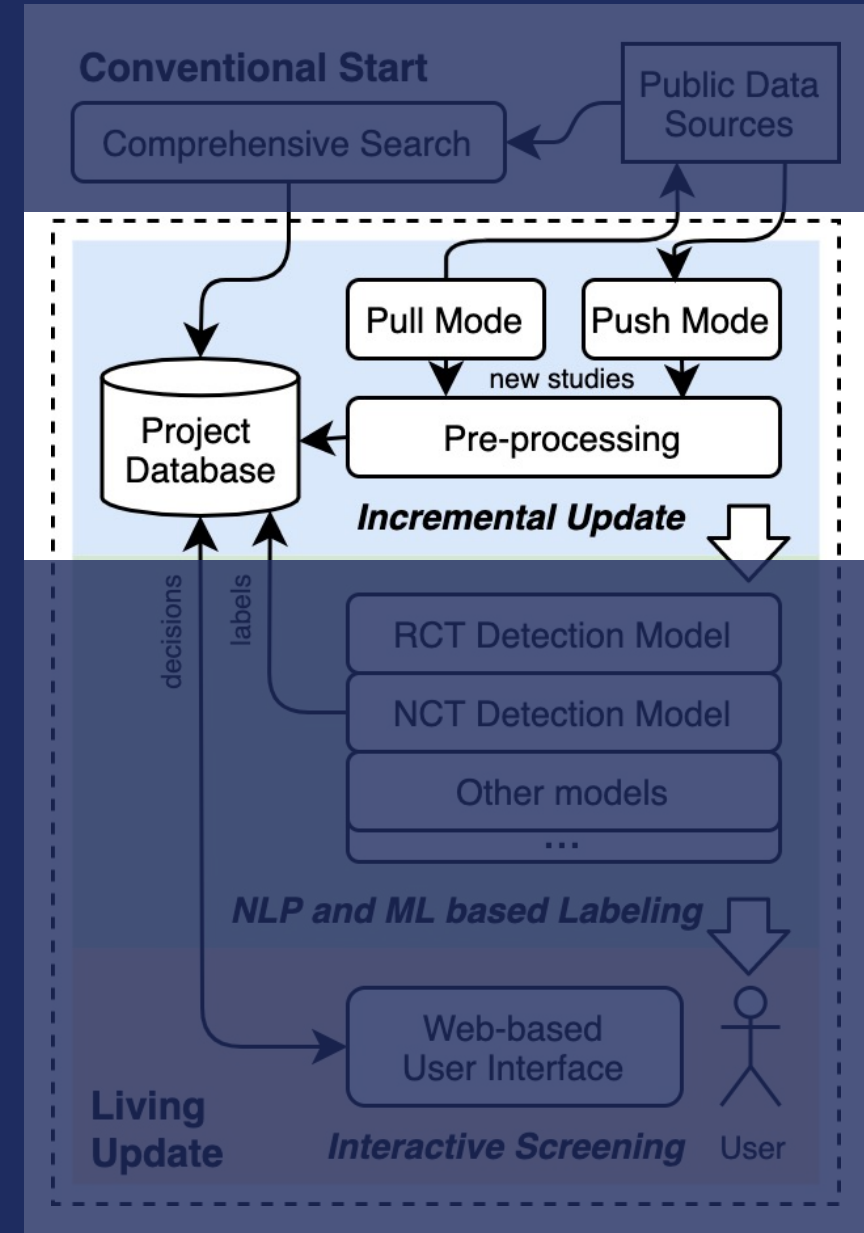
- Conventional Start
 - Define search strategies for a given clinical question for SR
 - Search with customized strategies in multiple databases (e.g., PubMed, EMBASE, etc.)
 - Initialize screening with UI or import results on existing SR projects
- Living Update
 - Hybrid approach for small batch update on coming new studies
 - *Incremental update*
 - *NLP and ML based labeling*
 - *Interactive screening UI*



Methods - Incremental Update

Retrieve new studies from data sources with automated programs

- Pull Mode
 - Search: PubMed
 - Manual import: EndNote XML and others
- Push Mode
 - Email alert: subscribe pre-defined emails
 - API push
- Pre-processing
 - Duplicate study detection
 - Attribute fixing
 - Project routing

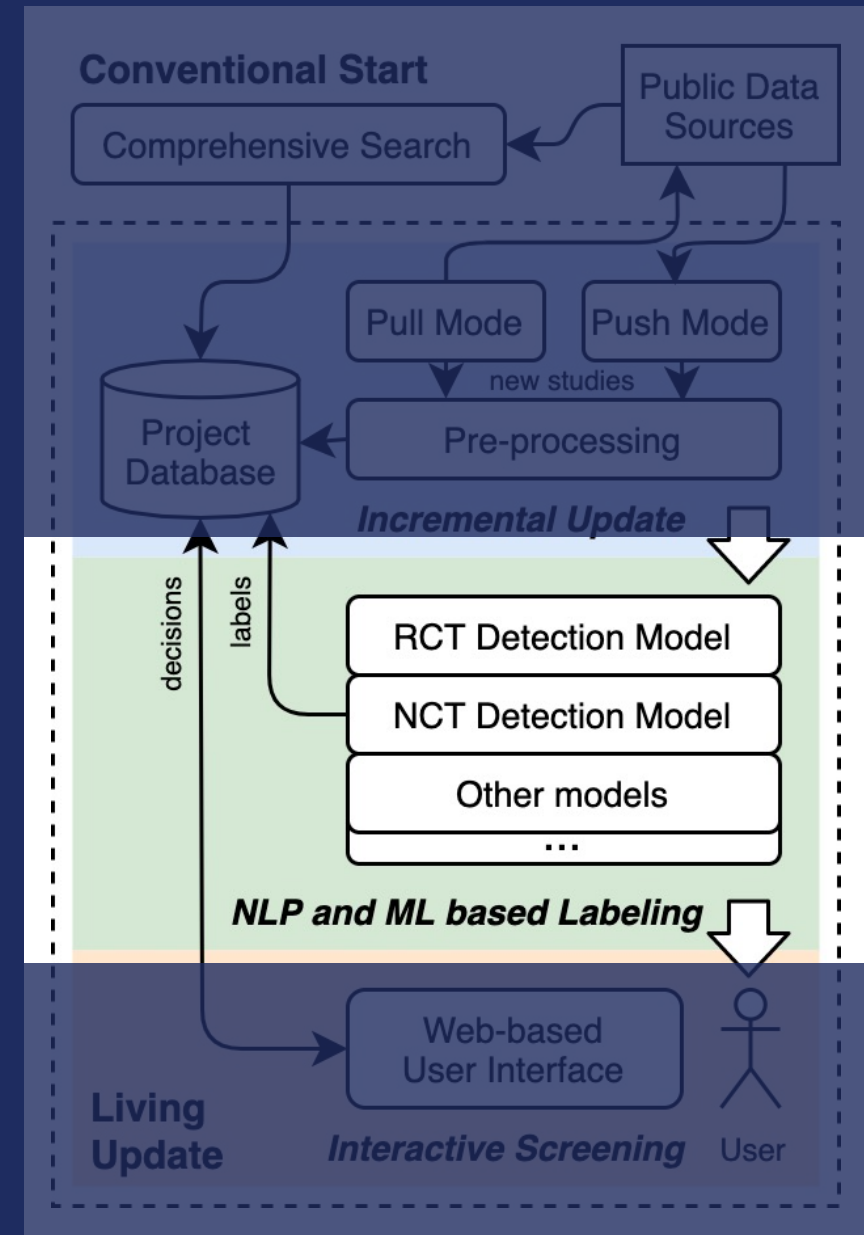


Methods - Labeling

Add additional information to studies with NLP and ML based methods

- Randomized Controlled Trial (RCT)
 - RobotSearch *
- Clinical trial number
- PMID / DOI number
- Inclusion / exclusion criteria keywords
- User customized tags
- Publication date

...

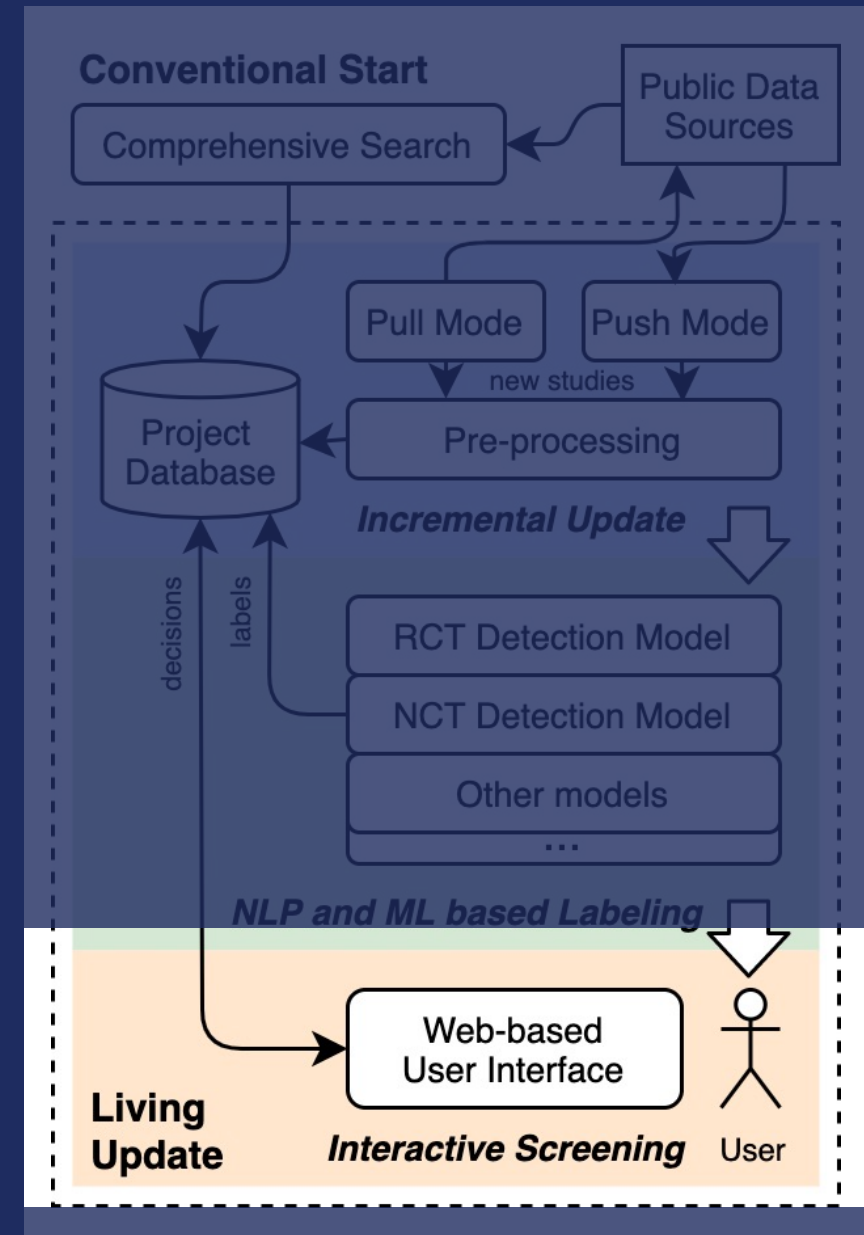


* Marshall I, Storr AN, Kuiper J, Thomas J, Wallace BC. Machine Learning for Identifying Randomized Controlled Trials: an evaluation and practitioner's guide. Res Syn Meth. 2018. <https://doi.org/10.1002/jrsm.1287>

Methods - Interactive UI

Assist users in the screening process with interactive and intuitive interface

- Web-based UI
 - Multi-user collaboration / assessment
 - Optimized workflow for screening
- User centered design
 - Focus on the key steps of screening
 - Task-driven interface
- Human-Computer Interaction principles
 - Reducing the cognitive overhead
 - Low physical effort
 - Flexibility in use



Methods - Interactive UI

Renal Cell Carcinoma
Screener Overview
🔔
👤

All References

Unscreened 1073 0

Decided 66 23

Further Check

Full Text Review 3 0

Included References

Included 46 10 4

- Metastatic Re... 35 11

- First-line Tre... 31 15

- Adjuvant Tyr... 9 37

Excluded References

By Title 12 10 1

By Abstract 4 3 0

By Full Text 4 0 0

Reference List | Included in SR [Inclusion / Exclusion Criteria](#)

Filters: Show RCT Only Show Check Later Only **Agreement:** ALL Agreed Disagreed Unconfirmed **Sort by:** Decision Reason

Tag Filters: Reset Other MA

Show 10 entries Search:

#	Date	Labels	Title	
1142	2021-10-03	RCT	Adjuvant Pembrolizumab after Nephrectomy in Renal-Cell Carcinoma.	2021-10-03 Included in SR Through Abstract
1141	2021-09-11	RCT	Adjuvant Pazopanib Versus Placebo After Nephrectomy in Patients With Localized or Locally Advanced Renal Cell Carcinoma: Final Overall Survival Analysis of the Phase 3 PROTECT Trial.	2021-09-11 Included in SR Through Abstract
1140	2021-09-11	RCT	Adjuvant sunitinib or sorafenib for high-risk, non-metastatic renal-cell carcinoma (ECOG-ACRIN E2805): a double-blind, placebo-controlled, randomised, phase 3 trial.	2021-09-11 Included in SR Through Abstract
1139	2021-09-11	RCT	Axitinib versus placebo as an adjuvant treatment of renal cell carcinoma: results from the phase III, randomized ATLAS trial.	2021-09-11 Included in SR Through Abstract
1138	2021-09-11	RCT	Adjuvant sunitinib in patients with high-risk renal cell carcinoma: safety, therapy management, and patient-reported outcomes in the S-TRAC trial.	2021-09-11 Included in SR Through Abstract
1137	2021-09-11	RCT	Adjuvant Sorafenib for Renal Cell Carcinoma at Intermediate or High Risk of Relapse: Results From the SORCE Randomized Phase III Intergroup Trial.	2021-09-11 Included in SR Through Abstract
	2021-09-11	RCT	Adjuvant Sunitinib in High-Risk Renal-Cell Carcinoma after Nephrectomy.	2021-09-11 Included in SR Through Abstract
	2021-09-11	RCT	Adjuvant Sunitinib for High-risk Renal Cell Carcinoma After Nephrectomy: Subgroup Analyses and Updated Overall Survival Results.	2021-09-11 Included in SR Through Abstract
	2021-09-11	RCT	Randomized Phase III Trial of Adjuvant Pazopanib Versus Placebo After Nephrectomy in Patients With Localized or Locally Advanced Renal Cell Carcinoma.	2021-09-11 Included in SR Through Abstract
	2021-07-12	RCT	A phase 3 study (COSMIC-313) of cabozantinib in combination with nivolumab and ipilimumab in patients with previously untreated advanced renal cell carcinoma of intermediate or poor risk.	2021-08-17 Included in SR Through Full Text Review

10 of 46 entries Previous 1 2 3 4 5 Next

Study categories

Categorized by screening stages and topics with numbers indicating the screening progress

Methods - Interactive UI

The screenshot shows a web interface for a clinical trial reference list. The main content is a table of references, with a blue callout box highlighting the filter controls. The callout box contains the text: "Dynamic filters" and "Filters for displaying those studies of interests for making screening decisions, such as labels, tags, conditions".

Dynamic filters
Filters for displaying those studies of interests for making screening decisions, such as labels, tags, conditions

Reference List | Included in SR Inclusion / Exclusion Criteria

Filters: Show **RCT** Only Show **Check Later** Only **Agreement:** ALL Agreed Disagreed Unconfirmed **Sort by:** Decision Reason

Tag Filters: Other MA

Show 10 entries

#	Date	Labels	Title	
1142	2021-10-03	RCT	Adjuvant Pembrolizumab after Nephrectomy in Renal-Ce	2021-10-03 Included in SR ? Through Abstract
1141	2021-09-11	RCT	Adjuvant Pazopanib Versus Placebo After Nephrectomy i	2021-09-11 Included in SR ? Through Abstract
1140	2021-09-11	RCT	Adjuvant sunitinib or sorafenib for high-risk, non-metast	2021-09-11 Included in SR ? Through Abstract
1139	2021-09-11	RCT	Axitinib versus placebo as an adjuvant treatment of renal cell carcinoma: results from the phase III, randomized ATLAS trial.	2021-09-11 Included in SR ? Through Abstract
1138	2021-09-11	RCT	Adjuvant sunitinib in patients with high-risk renal cell carcinoma: safety, therapy management, and patient-reported outcomes in the S-TRAC trial.	2021-09-11 Included in SR ? Through Abstract
1137	2021-09-11	RCT	Adjuvant Sorafenib for Renal Cell Carcinoma at Intermediate or High Risk of Relapse: Results From the SORCE Randomized Phase III Intergroup Trial.	2021-09-11 Included in SR ? Through Abstract
1136	2021-09-11	RCT	Adjuvant Sunitinib in High-Risk Renal-Cell Carcinoma after Nephrectomy.	2021-09-11 Included in SR ? Through Abstract
1135	2021-09-11	RCT	Adjuvant Sunitinib for High-risk Renal Cell Carcinoma After Nephrectomy: Subgroup Analyses and Updated Overall Survival Results.	2021-09-11 Included in SR ? Through Abstract
1134	2021-09-11	RCT	Randomized Phase III Trial of Adjuvant Pazopanib Versus Placebo After Nephrectomy in Patients With Localized or Locally Advanced Renal Cell Carcinoma.	2021-09-11 Included in SR ? Through Abstract
422	2021-07-12	RCT	A phase 3 study (COSMIC-313) of cabozantinib in combination with nivolumab and ipilimumab in patients with previously untreated advanced renal cell carcinoma of intermediate or poor risk.	2021-08-17 Included in SR ✓ Through Full Text Review

Showing 1 to 10 of 46 entries

Previous 1 2 3 4 5 Next

Screener Tools

- Update Original/Followup
- Export Reference List
- Show PRISMA

Further Check

- Full Text Review 3 0

Included References

- Included 46 10 4
- Metastatic Re... 35 11
- First-line Tre... 31 15
- Adjuvant Tyr... 9 37

Excluded References

- By Title 12 10 1
- By Abstract 4 3 0
- By Full Text 4 0 0

Renal Cell Carcinoma Screener Overview

All References

Unscreened 1073 0

Decided 66 23

Methods - Interactive UI

The screenshot shows a web interface for managing clinical trial references. The main section is titled 'Reference List | Included in SR' and includes filters for 'Unscreened' (1073) and 'Decided' (66/23). A table lists 46 entries, with the first 10 visible. A blue callout box labeled 'Study list' highlights that studies matching category, filters, and other criteria are listed with basic information and decision/operation options.

Study list
Studies that match the category, filters, and other criteria are listed with basic information and decision / operations

#	Date	Labels	Title	
1142	2021-10-03	RCT	Adjuvant Pembrolizumab after Nephrectomy in Renal-Cell Carcinoma.	2021-10-03 Included in SR ? Through Abstract
1141	2021-09-11	RCT	Adjuvant Pazopanib Versus Placebo After Nephrectomy in Patients With Localized or Locally Advanced Renal Cell Carcinoma: Final Overall Survival Analysis of the Phase 3 PROTECT Trial.	2021-09-11 Included in SR ? Through Abstract
1140	2021-09-11	RCT	Adjuvant sunitinib or sorafenib for high-risk, non-metastatic renal-cell carcinoma (ECOG-ACRIN E2805): a double-blind, placebo-controlled, randomised, phase 3 trial.	2021-09-11 Included in SR ? Through Abstract
1139	2021-09-11	RCT	Axitinib versus placebo as an adjuvant treatment of renal cell carcinoma: results from the phase III, randomized ATLAS trial.	2021-09-11 Included in SR ? Through Abstract
1138	2021-09-11	RCT	Adjuvant sunitinib in patients with high-risk renal cell carcinoma: safety, therapy management, and patient-reported outcomes in the S-TRAC trial.	2021-09-11 Included in SR ? Through Abstract
1137	2021-09-11	RCT	Adjuvant Sorafenib for Renal Cell Carcinoma at Intermediate or High Risk of Relapse: Results From the SORCE Randomized Phase III Intergroup Trial.	2021-09-11 Included in SR ? Through Abstract
1136	2021-09-11	RCT	Adjuvant Sunitinib in High-Risk Renal-Cell Carcinoma after Nephrectomy: Results From the Phase III, Randomized, Placebo-Controlled, Double-Blind, Parallel-Group, Multicenter, Prospective Trial.	2021-09-11 Included in SR ? Through Abstract
1135	2021-09-11	RCT	Adjuvant Sunitinib for High-risk Renal Cell Carcinoma After Nephrectomy: Results From the Phase III, Randomized, Placebo-Controlled, Double-Blind, Parallel-Group, Multicenter, Prospective Trial.	2021-09-11 Included in SR ? Through Abstract
1134	2021-09-11	RCT	Randomized Phase III Trial of Adjuvant Pazopanib Versus Placebo After Nephrectomy in Patients With Localized or Locally Advanced Renal Cell Carcinoma.	2021-09-11 Included in SR ? Through Abstract
422	2021-07-12	RCT	A phase 3 study (COSMIC-313) of cabozantinib in combination with sunitinib versus sunitinib alone in previously untreated advanced renal cell carcinoma of intermediate or poor risk.	2021-08-17 Included in SR ✓ Through Full Text Review

Case Study - Screening Studies

Renal Cell Carcinoma Screeener Overview

All References Reference List | Unscreened References Inclusion / Exclusion Criteria

Unscreened 1073 0 Decided 66 23

Filters: Show RCT Only Show Check Later Only

Tag Filters: Other MA

Sort by: Decision Reason

Further Check Show 10 entries Search:

#	Date	Labels	Title	
1133	N/DOI: 10.1046/j...	2021-09-01	Genetic changes in stage pT2N0 prostate cancer studied by comparative genomic hybridization	Exclude By Title
1132	PMID: 15153336	2021-09-01	High-resolution analysis of gene copy number alterations in human prostate cancer using CGH on cDNA microarrays: impact of copy number on gene expression	Exclude By Title
1131	PMID: 16037637	2021-09-01	NMD microarray analysis for rapid genome-wide screen of mutated genes in cancer	Exclude By Title
1130	N/DOI: 10.1093/c...	2021-09-01	A distinct ERCC1 haplotype is associated with mRNA expression levels in prostate cancer patients	Exclude By Title
1129	PMID: 21348634	2021-09-01	Polygenic modeling of genome-wide association studies: an application to prostate and breast cancer	Exclude By Title
1128	PMID: 21160075	2021-09-01	Personalized prostate cancer screening: improving PSA tests with genomic information	Exclude By Title
1127	N/DOI: 10.1042/C...	2021-09-01	Prostate cancer, PI3K, PTEN and prognosis	Exclude By Title
1125	PMID: 2575485	2021-09-01	Chromosomal localization to 3q21----qter and two TaqI RFLPs of the human prostate-specific acid phosphatase gene (ACPP)	Exclude By Title
1124	N/DOI: 10.4267/2...	2021-09-01	CDK2 (cyclin dependent kinase 2)	Exclude By Title
1123	PMID: 9852240	2021-09-01	A transgenic strategy for analyzing the regulatory regions of the human prostate-specific antigen gene: potential applications for the treatment of prostate cancer (Review)	Exclude By Title

Showing 1 to 10 of 1,073 entries

Previous 1 2 3 4 5 ... 108 Next

Included References: 46 10 4

- Metastatic Re... 35 11
- First-line Tre... 31 15
- Adjuvant Tyr... 9 37

Excluded References: By Title 12 10 1, By Abstract 4 3 0, By Full Text 4 0 0

Screeener Tools: Update Original/Followup, Export Reference List, Show PRISMA

Summary

To improve the efficiency of screening in a Living SRMA

- Incremental search + conventional start for living update
- Automatically generated labels based on NLP and ML models for decision making
- Interactive UI based on user centered design and HCI principles

Thank you for your attention :)