Systematic reviews ‘Werken met referentiesoftware en tools’

14 Juni 2022

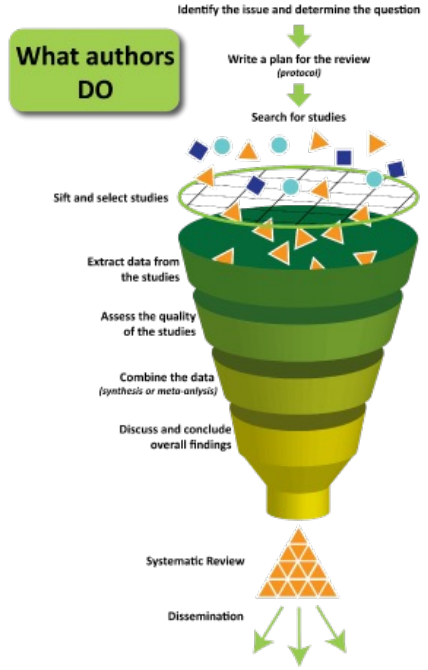
René Spijker

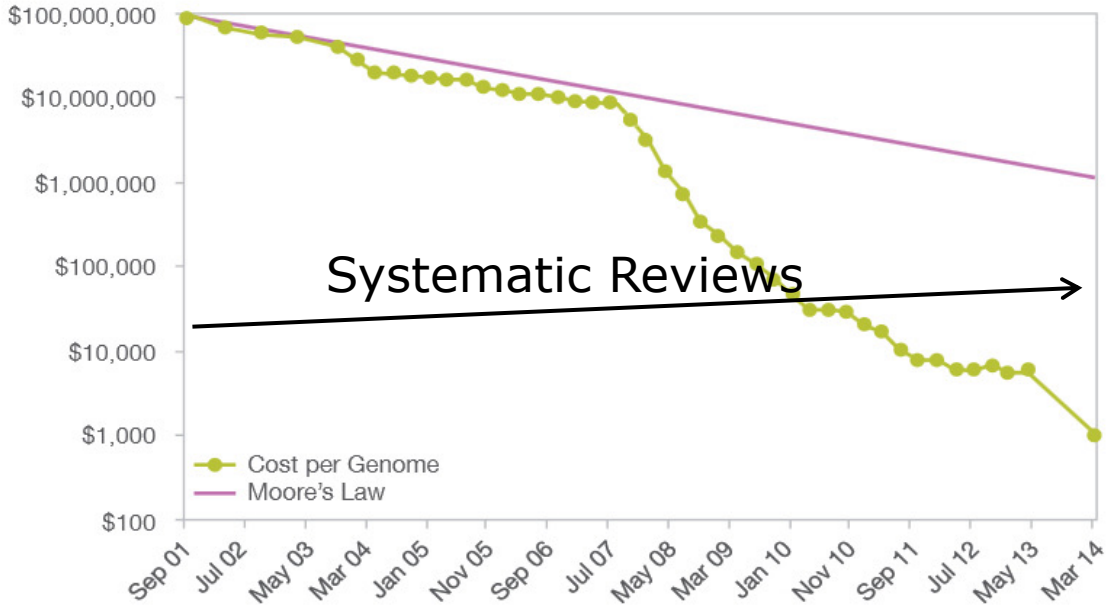
About me

- Informatiespecialist Amsterdam UMC, Universiteit van Amsterdam
- Informatiespecialist/senior onderzoeker Cochrane Netherlands
- Member Cochrane Council
- Member Cochrane Information specialist Executive
- Cochrane Methods groups
  - Screening and diagnostic test evaluation
  - Prognosis
  - Information Retrieval (IRMG)
  - Rapid Reviews
- EVBRES: co-lead work package on systematic review efficiency
- Organising committee ICASR (International collaboration for the automation of systematic reviews)

Systematic Reviews

- Een systematisch review neemt veel tijd in beslag en bestaat nog uit veel menselijk werk
- Vanuit morse law(computer chips) zien we dat automatiserig zowel kosten als tijd elke 2 jaar reduceerd
- Met het human genome project zagen we een zelfde ontwikkeling op het gebied van DNA sequencing.
- Is dit ook mogelijk voor het systematisch review process





Wat kunnen we definieren als technologische ondersteuning

- Alleen Machine learning en kunstmatige intelligentie?
- Elke technologische ondersteuning welke een menselijke handeling vervangt

- *Uiteindeijk gaat het om de kwaliteit en efficientie en niet om de techniek*

- Het genereren van een referentie lijst kan handmatig maar je kunt ook endnote gebruiken om je referenties in te voegen en met 1 klik aan te passen.
- Mbv simpele regex expressie dubbele referenties (bv in endnote) ontdekken
- High level deep learning applicaties om duplicaten te ontdekken

Hoe het begon

- Eerste serieuze introductie van technologie was het ontwikkelen van een specifieke database waarin alle studies relevant voor Cochrane reviews werden bijgehouden ➔ uiteindelijk zou dit CENTRAL worden

- De start van Revman om reviews te schrijven waarin ook de data in tabellen gezet kon worden en meta-analyses gedaan konden worden.

- Tegen woordig geeft Revman ook suggesties over items welke op specifieke plekken in het review benoemd dienen te worden (MERCIR standards)

Eerste kennismaking met geavanceerde technologie

- Cochrane Colloquium 2010 keystone
- Byron Wallace presenteerd abstrackr waarbij gebruik is gemaakt van een rudimentaire active learning component op basis van SVM (support vector machines) om het selecteren van abstracts te versnellen
- Op zelfde moment komt NIST (national institute of standards and technology) met een TREC in het juridische domein welke het ’total recall’ probleem wil aanpakken.
- Gordon Cormack van Waterloo university ontwikkeld daarvoor een geavanceerdere variant continuous active learning op basis van SVM classificatie.

- Vanaf dat moment worden op veel verschillende plekken in de wereld deze active learning methoden ontwikkeld (in R, python, java, etc) en toegepast op systematische reviews.

- Meest actief hierin is het Eppi centre onder leiding van James Thomas. Al snel wordt een priority screening module ingebouwd in het reeds bestaande systematic review platform Eppi Reviewer. (10.1002/jrsm.1093) en bekeken hoe deze het best in te passen (10.1186/s13643-016-0315-4) als single screener, dubbel screener, machine alleen.

- Voor gebruik in scoping reviews waarbij je accepteerd een studie te missen werd dit dan ook al snel toegepast bv persoonlijk vpoorbeeld uit 2016: 10.1186/s12911-016-0281-8.

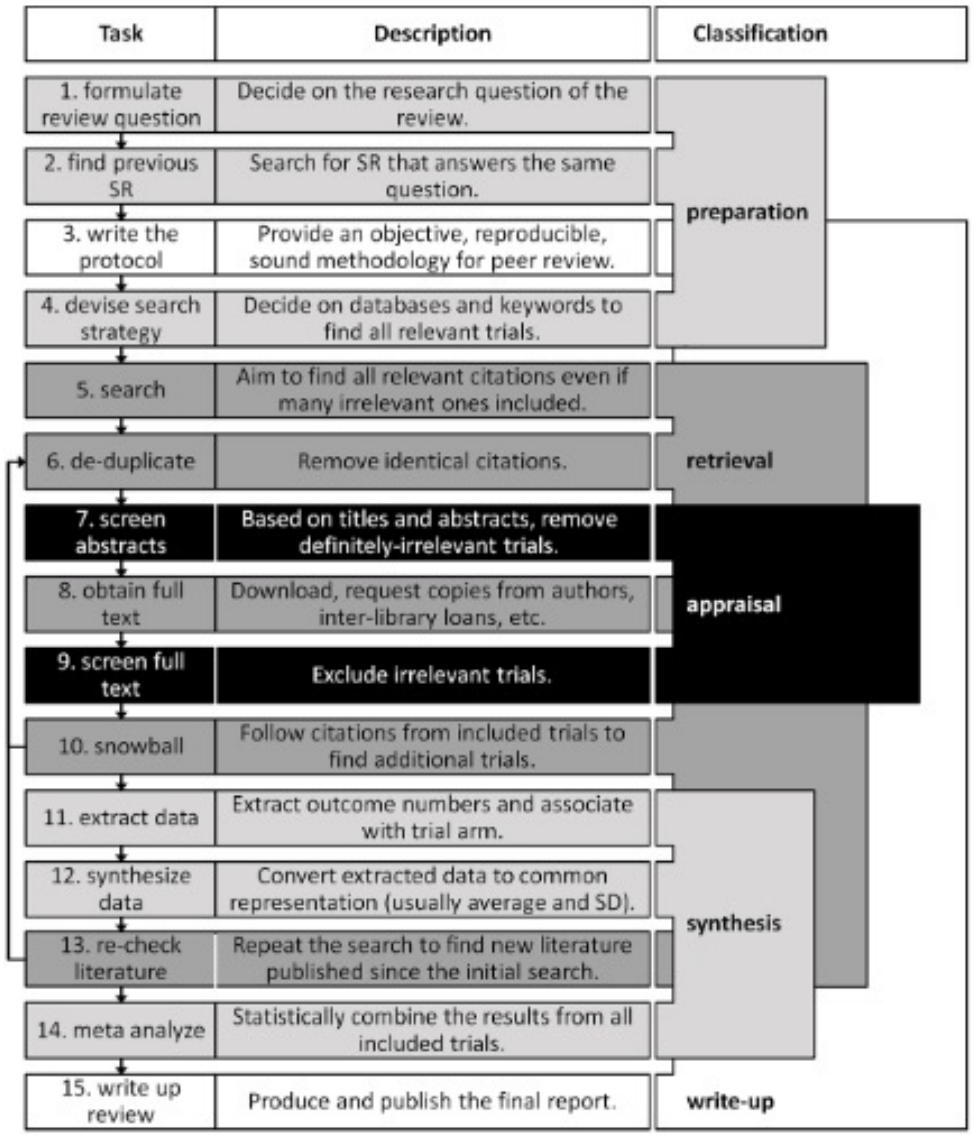
Steeds meer technologie

- 2015 Cochrane colloquium: Oprichting van ICASR door Paul Glasziou en Per Olav Vandvik
- ICASR International organisation for the automation of systematic reviews (verslagen van meerdere meetings: https://pubmed.ncbi.nlm.nih.gov/?term=icasr%5Bti%5D)
- Zeker met het oog op research waste graag eerst goed op de hoogte zijn middels systematic reviews voordat een nieuwe trial wordt uitgevoerd
- Daarvoor moet de tijd welke het kost om een review uit te voeren wel naar beneden

Vienna principles

- Systematic reviews involve multiple tasks, each with different issues, but all must be improved.
- Automation may assist with all tasks, from scoping reviews to identifying research gaps as well protocol development to writing and dissemination of the review.
- The processes for each task can and should be continuously improved, to be more efficient and more accurate.
- Automation can and should facilitate the production of systematic reviews that adhere to high standards for the reporting, conduct and updating of rigorous reviews.
- Developments should also provide for flexibility in combining and using, e.g. subdividing or merging steps and allowances for different users to use different interfaces.
- Different groups with different expertise are working on different parts of the problem; to improve reviews as a whole will require collaboration between these groups.
- Every automation technique should be shared, preferably by making code, evaluation data and corpora available for free.
- All automation techniques and tools should be evaluated using a recommended and replicable method with results and data reported.

Welke taken zijn er in een systematisch review



Er is al veel software voor de verschillende stappen I

Search	2Dsearch	visuele search builder
	Inciteful.xyz	tool om gerelateerd artikelen te vinden
	MeSH on Demand	Deze tool geeft je aan welke potentiele MeSH termen bij een artikel zouden passen
	NCI metathesaurus	openbare toegang tot de synoniemen lijst uit de UMLS
	polyglot search	tool om automatische syntax vertaling te doen tussen database searches
	PubReMiner	tool om frequentie analyse uit te oeren op een pubmed search
	Vosviewer	Tool om artikelen te clusteren naar onderwerpen
	SearchRefinery	tool om tov een referentie set je search te onderzoeken naar overbodige links om de precissie te verhogen
	litsearchr	R-package for text analysis and search construction
	Swift Review	text analyses waaronder topic modelling
	multitagger	identificatie van verschillende study designs in pubmed artikelen
	Robotsearch	RCT classifier om uit een search alle potentiele RCT studies te filteren
deduplicate	Deduplicator(arrowsmith)	
	deduplication function	Eppi, rayyan, covidence, distillerSR
	ASySD	shiny app deduplicator CAMARADES
	Deduplicator(SRA)	

Er is al veel software voor de verschillende stappen II

Screening	Rayyan	screenen en conflict oplossen
	Covidence	screenen en conflict oplossen
	screenatron	screenen en conflict oplossen
	Swift active screener	Re-ranking of screening results om de meest relevante eerst te krijgen zodat rest van team alvast door kan
	DistillerSR	classifier builder and re-ranking
	Colandr	re-ranking
	EppiReviewer	classifier builder and re-ranking
	ASReview	re-ranking
obtain full text	endnote	kan full text finden op basis van doi, pmid, openurl
	zotero	kan full text finden op basis van doi, pmid, openurl
	quosa	full text vinden op basis van doi, crossref
	PMC	open access database
citation tracking	citationchaser	aan de hand van lens.org forward/backward citations
	spidercite	aan de hand van lens.org forward/backward citations
	scopus	zowel forward als backward staan in de database
	web of science	zowel forward als backward staan in de database

Er is al veel software voor de verschillende stappen III

extract data	Dexter	semi-automatische data extractie
	excel	
Risk Of Bias	Robotreviewer	
meta-analyse	metafor package	
	review manager	
write-up	methods wizard	
	revman replicant	

Eerste evaluaties in het veld

Rathbone *et al. Systematic Reviews* (2015) 4:80
DOI 10.1186/s13643-015-0067-6

SYSTEMATIC REVIEWS

RESEARCH Open Access

Faster title and abstract screening? Evaluating Abstrackr, a semi-automated online screening program for systematic reviewers

John Rathbone*, Tammy Hoffmann and Paul Glasziou

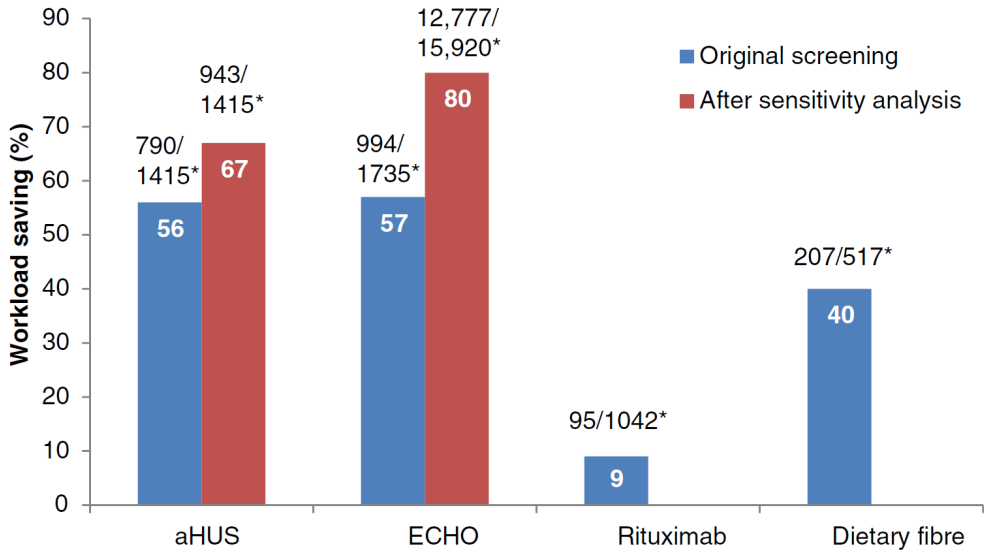


Fig. 4 Workload saving (%) when using Abstrackr in each of the four datasets. *Raw numbers of the proportion of citations predicted not relevant from the total

The CLEF Initiative

Conference and Labs of the Evaluation Forum

ehealth

- **Task 1**. Multilingual Information Extraction
- **Task 2.** Technologically Assisted Reviews in Empirical Medicine
- **Task 3.** Patient-centred information retrieval

- The lab will focus on Diagnostic Test Accuracy (DTA) reviews. Search in this area is generally considered the hardest, and a breakthrough in this field would likely be applicable to other areas as well.

- The task will have a focus on the second stage of the process, i.e. given the results of a Boolean Search how to make Abstract and Title Screening more effective.

Methoden gebruikt

- Keyword extraction
- Learning to rank
- Logic regression
- **Continuous active learning SVM**
- Neural network
- Deep-learning

CLEF 2018 Technologically Assisted Reviews in Empirical Medicine Overview

Evangelos Kanoulas[1], Dan Li[1], Leif Azzopardi[2], and Rene Spijker[3]

[1] Informatics Institute, University of Amsterdam, Netherlands,
E.Kanoulas@uva.nl, D.Li@uva.nl
[2] Computer and Information Sciences, University of Strathclyde, Glasgow, UK,
leif.azzopardi@strath.ac.uk
[3] Cochrane Netherlands and UMC Utrecht, Julius Center for Health Sciences and Primary Care, Netherlands, R.Spijker-2@umcutrecht.nl

Figure 6 shows the recall-effort curves for the participats' runs, that is the recall value at different percentage of abstracts shown to the user.

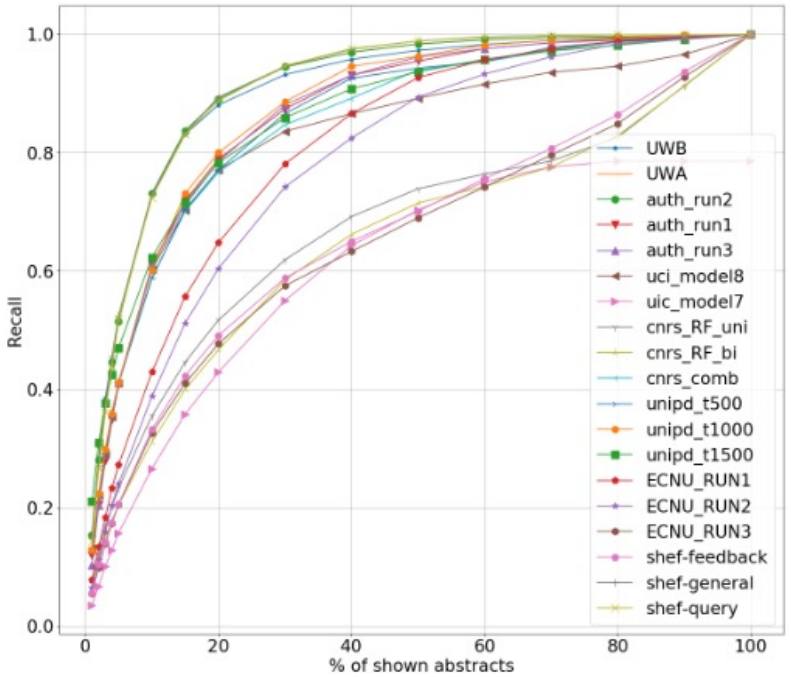


Fig. 6. Recall at different ranks.

Input: zoekresultaat

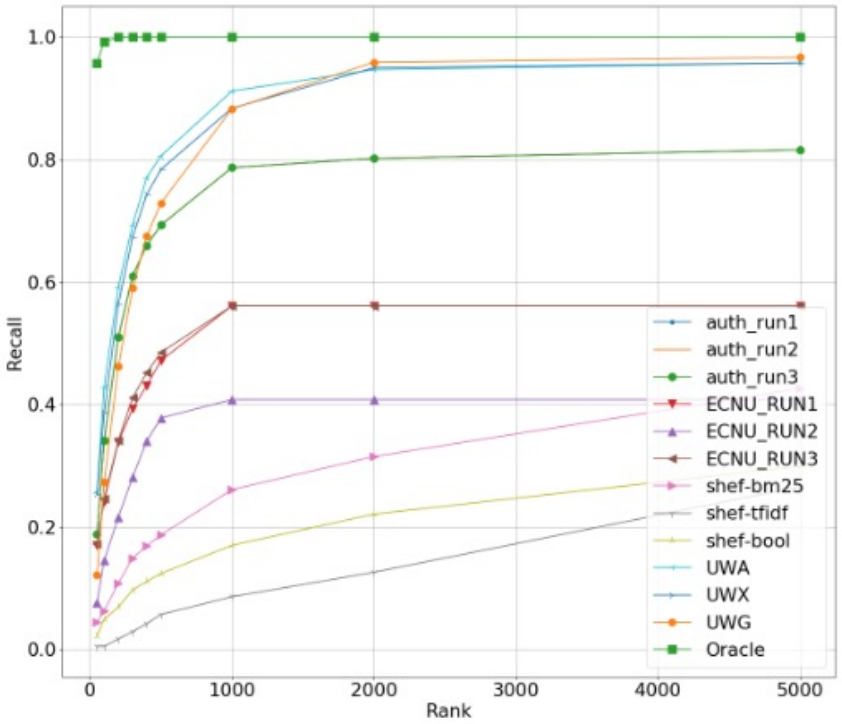
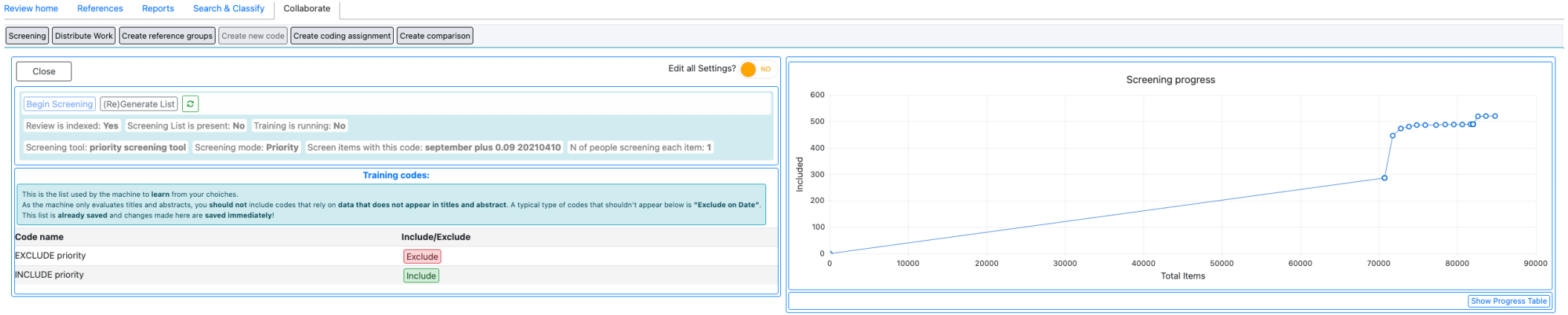
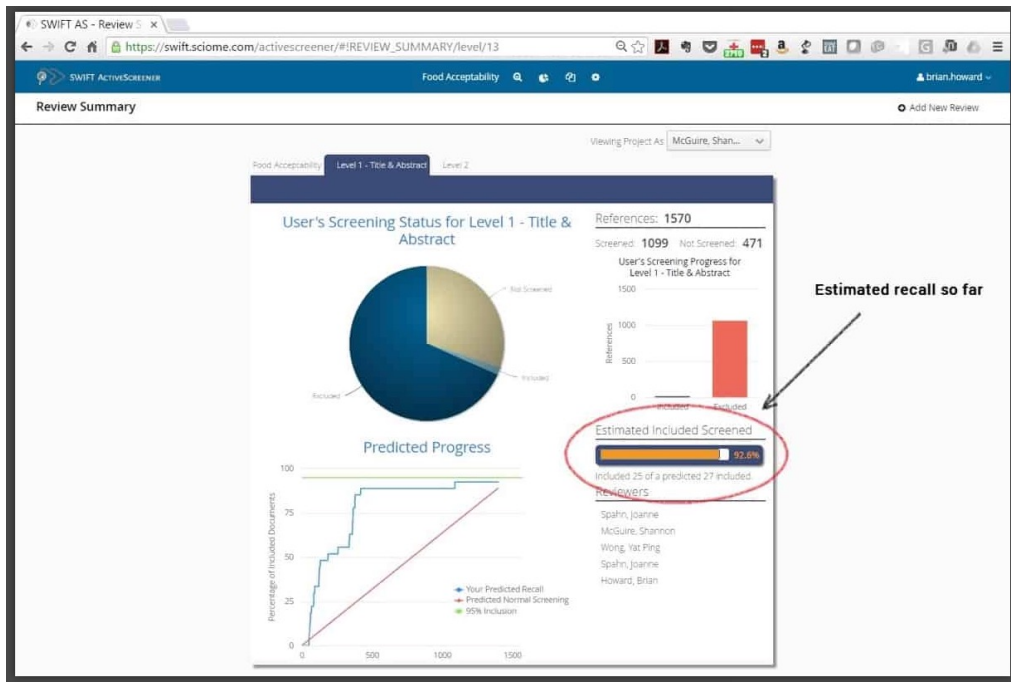
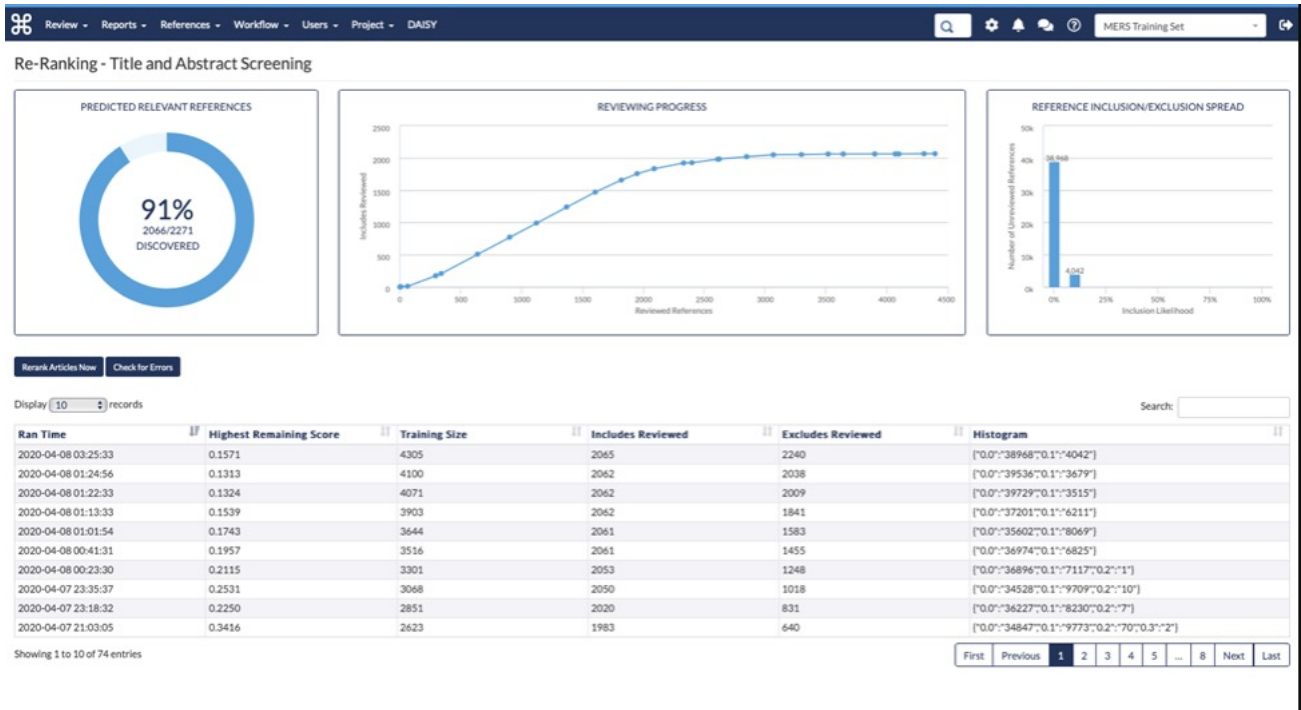


Fig. 2. Recall at different top-k percentages of shown abstracts. Recall is computed using the abstract level relevance labels.

Input: heel pubmed


Review home
References
Reports
Search & Classify
Collaborate
Screening
Distribute Work
Create reference groups
Create new code
Create coding assignment
Create comparison
Close
Edit all Settings?
NO
Begin Screening
(Re)Generate List
Review is indexed: Yes
Screening List is present: No
Training is running: No
Screening tool: priority screening tool
Screening mode: Priority
Screen items with this code: september plus 0.09 20210410
N of people screening each item: 1
Training codes:
This is the list used by the machine to learn from your choiches.
As the machine only evaluates titles and abstracts, you should not include codes that rely on data that does not appear in titles and abstract. A typical type of codes that shouldn't appear below is "Exclude on Date".
This list is already saved and changes made here are saved immediately!
Code name
Include/Exclude
EXCLUDE priority
Exclude
INCLUDE priority
Include
Screening progress
Included
Total Items
Show Progress Table

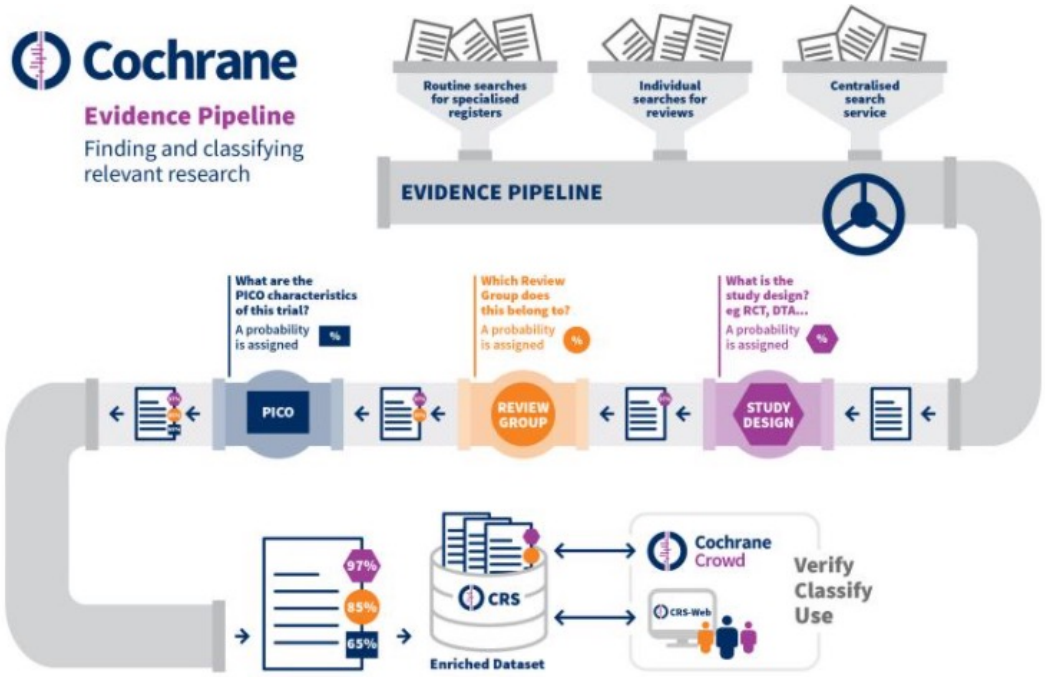

SWIFT AS - Review
https://swift.sciome.com/activescreener/#!REVIEW_SUMMARY/level/13
SWIFT ACTIVESCREENER
Food Acceptability
brian.howard
Review Summary
Add New Review
Viewing Project As
McGuire, Shan...
Food Acceptability
Level 1 - Title & Abstract
Level 2
User's Screening Status for Level 1 - Title & Abstract
References: 1570
Screened: 1099
Not Screened: 471
User's Screening Progress for Level 1 - Title & Abstract
Estimated recall so far
Predicted Progress
Estimated Included Screened
92.6%
Included 25 of a predicted 27 included
Reviewers
Spahn, Joanne
McGuire, Shannon
Wong, Yat Ping
Spahn, Joanne
Howard, Brian
Your Predicted Recall
Predicted Normal Screening
95% Inclusion


Review
Reports
References
Workflow
Users
Project
DAISY
MERS Training Set
Re-Ranking - Title and Abstract Screening
PREDICTED RELEVANT REFERENCES
91%
2066/2271
DISCOVERED
REVIEWING PROGRESS
REFERENCE INCLUSION/EXCLUSION SPREAD
Rerank Articles Now
Check for Errors
Display 10 records
Search:
Ran Time
Highest Remaining Score
Training Size
Includes Reviewed
Excludes Reviewed
Histogram
2020-04-08 03:25:33 0.1571 4305 2065 2240 {"0.0":"38968","0.1":"4042"}
2020-04-08 01:24:56 0.1313 4100 2062 2038 {"0.0":"39536","0.1":"3679"}
2020-04-08 01:22:33 0.1324 4071 2062 2009 {"0.0":"39729","0.1":"3515"}
2020-04-08 01:13:33 0.1539 3903 2062 1841 {"0.0":"37201","0.1":"6211"}
2020-04-08 01:01:54 0.1743 3644 2061 1583 {"0.0":"35602","0.1":"8069"}
2020-04-08 00:41:31 0.1957 3516 2061 1455 {"0.0":"36974","0.1":"6825"}
2020-04-08 00:23:30 0.2115 3301 2053 1248 {"0.0":"36896","0.1":"7117","0.2":"1"}
2020-04-07 23:35:37 0.2531 3068 2050 1018 {"0.0":"34528","0.1":"9709","0.2":"10"}
2020-04-07 23:18:32 0.2250 2851 2020 831 {"0.0":"36227","0.1":"8230","0.2":"7"}
2020-04-07 21:03:05 0.3416 2623 1983 640 {"0.0":"34847","0.1":"9773","0.2":"70","0.3":"2"}
Showing 1 to 10 of 74 entries
First
Previous
1
2
3
4
5
...
8
Next
Last

Implementatie?

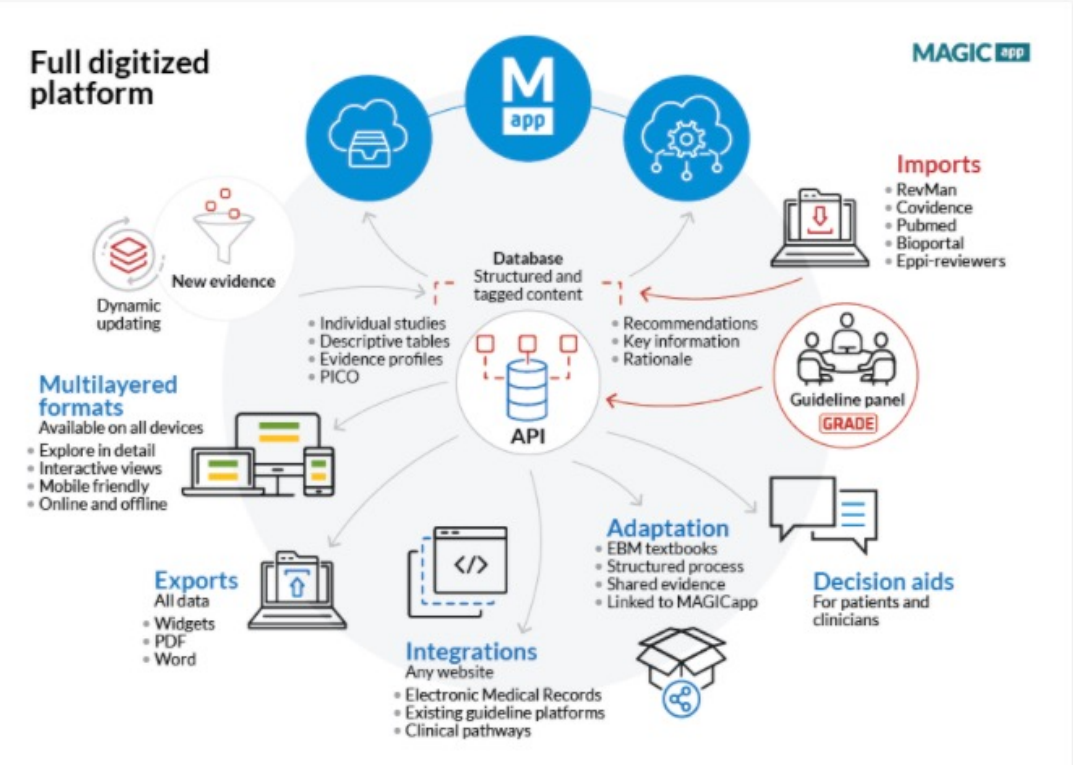
- Nog een weg te gaan
- Cochrane maakt op dit moment nog geen gebruik van priority screenen in systematische reviews (omdat er te weining data is over hoe of wanneer te stoppen en wat de gevolgen zijn)
- Wat dan wel

Evidence pipeline



- Classifiers getraind op de grote hoeveelheid data reeds beschikbaar binnen Cochrane
- Voorbeeld RCT classifier gebasseerd op 300000 beslissingen binnen Cochrane crowd
- Onderwerps specifieke classifiers op basis van Jaren lange opgebouwde registers

Alleen Cochrane?



- Magic app Noorwegen(NIPH) om informatie mbt evidence tussen richtlijn groepen beter uit te kunnen wisselen

Problemen bij implementatie

Received: 24 April 2018 | Revised: 28 September 2018 | Accepted: 11 December 2018
DOI: 10.1002/jrsm.1335

RESEARCH ARTICLE

WILEY Research Synthesis Methods

Usage of automation tools in systematic reviews

A.J. van Altena[1] | R. Spijker[2,3] | S.D. Olabarriaga[1]

Title

Systematic review automation tool use by systematic reviewers, health technology assessors and clinical guideline developers: tools used, abandoned, and desired.

Authors:

Anna Mae Scott[1], Connor Forbes[1], Justin Clark[1], Matt Carter[1], Paul Glasziou[1], Zachary Munn[2]

Affiliations:

[1] Institute for Evidence-Based Healthcare, Bond University
[2] JBI, Faculty of Health and Medical Sciences, The University of Adelaide

- Organisatorisch en niet een evidence problem
- Beter matchen met huidige workflow, liefst zoveel mogelijk binnen 1 platform
- Meer praktijk evaluaties
- Er bestaat inmiddels ook al een graveyard of tools (ooit ontwikkeld maar al lang niet meer up to date gehouden)

Vedere taken in het review proces

Ontdubbelen van referenties

- Gebruik van reference management software (eg endnote) voor ontdubbelen (bv Bramer et al 10.3163/1536-5050.104.3.014)

- Maar kan ook automatisch via deduplicator van Systematic review accelerator(https://sr-accelerator.com/#/), rayyan.ai, eppi reviewer, distillerSR, zelf ontwikkeld (bv AMC)

Hulp bij ontwikkelen van zoekacties

- SearchRefinder, MeSH analyzer, pubmed pubreminer, antconc, voyant, vos viewer etc.

- Vertalen van zoekacties
  - Macros' in word (bv Wichor's methode)
  - Polyglot search: https://sr-accelerator.com/#/polyglot

Full text download

- Reference management software (endnote, Zotero, jabref, etc)
- Scopus document download manager plugin (https://chrome.google.com/webstore/detail/scopus-document-download/ojplelelocihfchkdaebocpankipadmp)
- unpaywall

Critical appraisal

- Robot reviewer: automatisch genereren RoB van RCT's en highlight PDF waar informatie staat

physical activity and their confidence (i.e. self-efficacy), commitment, and intention to carry out the new goal. A complete description of the goal-setting procedures and the overall intervention is reported elsewhere [12].

Measures of goal-related physical activity

Goal-related physical activity was corroborated in the intervention group using concurrent records of daily pedometer steps and 10-min blocks of MVPA. Steps were assessed using the Yamax SW-200 pedometer (Yamax Corp, Tokyo, Japan), which has high accuracy for assessing level walking (±3% of actual steps) [27, 28], a target of the intervention.

Correlates of goal setting

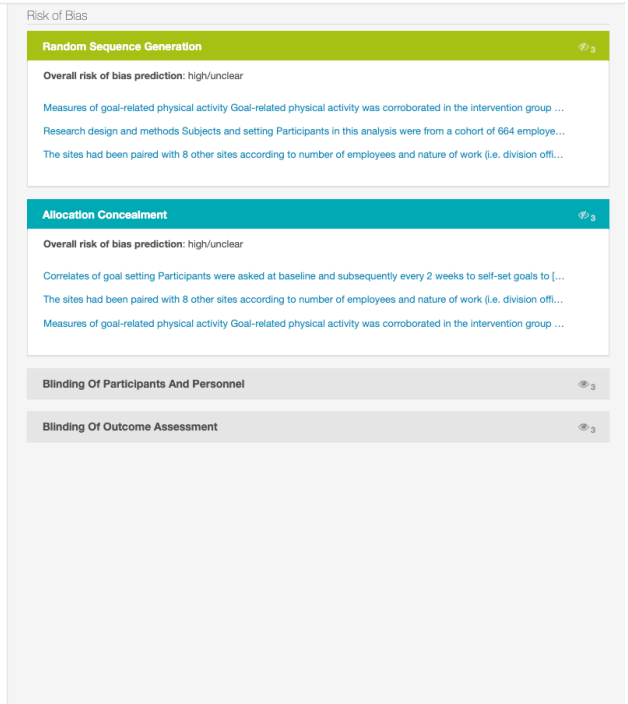
Participants were asked at baseline and subsequently every 2 weeks to self-set goals to [1] increase daily pedometer steps and the weekly number of 10-min blocks of MVPA accumulated during the subsequent 2-week period. They also provided biweekly self-ratings (1 = not at all, 2 = somewhat, 3 = moderately, 4 = very) of single-

0.514 across all paired weeks.

Latent growth modeling

Changes in the study variables across the 12-week intervention and the relations between those changes were analyzed using latent growth modeling (LGM) performed using M*plus* 5.2 [30]. Missing data were 1% overall (range from 0% to 3%) and were estimated using full-information maximum likelihood estimation, which yields accurate fit indices and parameter estimates with up to 25% simulated missing data [31]. Covariances could be computed for >75% of the variables. LGM applies confirmatory factor analysis to variables measured longitudinally [32, 33]. It estimates parameters and their standard errors (SEs) for latent factors of initial status (i.e. the latent mean at the baseline in the study), change (i.e. the slope or trajectory of change from the baseline across the 12 weeks of the study) and the variances (i.e. interindividual differences) of initial status and change. Critical *z* scores (parameter estimate/standard error) are used to test significance. Two-tailed probabilities are reported.

LGM was used as a three-stage process to esti-



European Journal of Epidemiology (2021) 36:165–178
https://doi.org/10.1007/s10654-020-00677-6

METHODS

Data extraction for epidemiological research (DExtER): a novel tool for automated clinical epidemiology studies

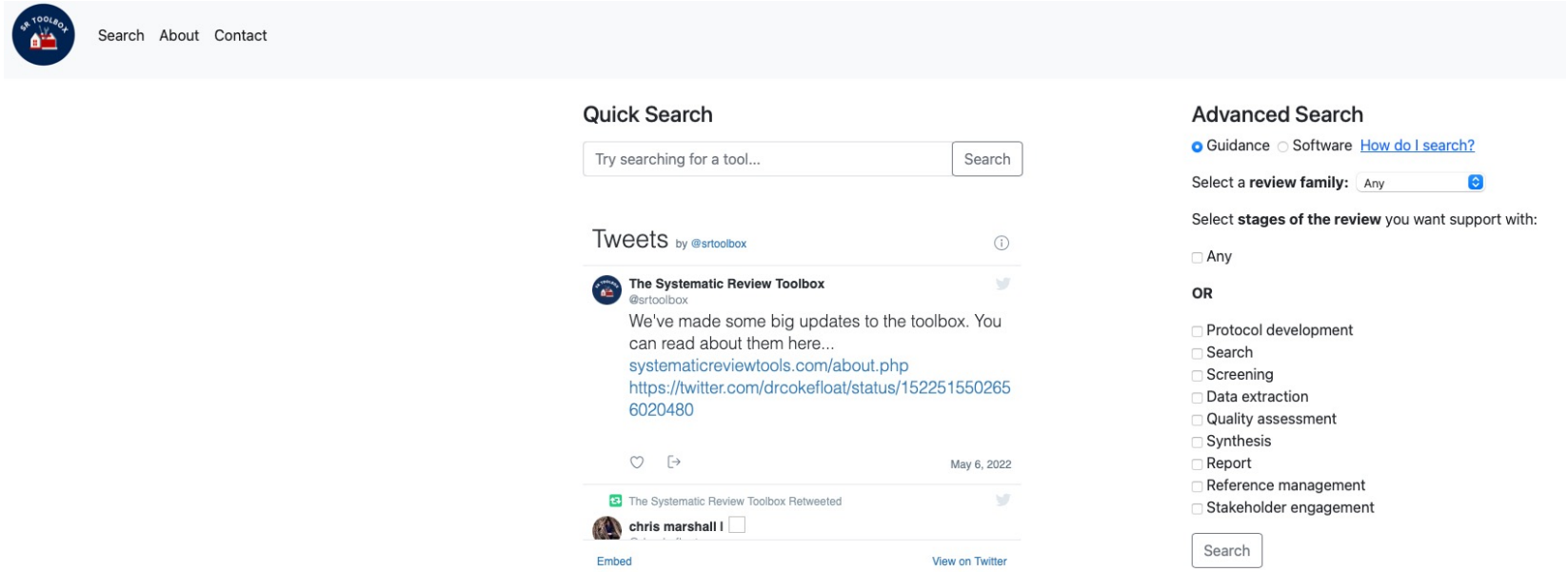
Krishna Margadhamane Gokhale[1,2,4] · Joht Singh Chandan[2] · Konstantinos Toulis[2] · Georgios Gkoutos[3,4] · Peter Tino[1] · Krishnarajah Nirantharakumar[2,4]

Received: 12 June 2019 / Accepted: 12 August 2020 / Published online: 27 August 2020

A distantly supervised dataset for automated data extraction from diagnostic studies

Christopher Norman,[1,2] Mariska Leeflang,[2] René Spijker,[3,4] Evangelos Kanoulas,[5] and Aurélie Névéol[1]
[1] LIMSI, CNRS, Université Paris-Saclay
[2] KEBB, Amsterdam Public Health, Amsterdam UMC, University of Amsterdam
[3] Medical Library, Amsterdam Public Health, Amsterdam UMC, University of Amsterdam
[4] Cochrane Netherlands, Julius Center for Health Sciences and Primary Care, UMCU, Utrecht University
[5] Informatics Institute and Amsterdam Business School, University of Amsterdam
norman@limsi.fr, m.m.leeflang@amc.uva.nl,
R.Spijker-2@umcutrecht.nl, E.Kanoulas@uva.nl,
neveol@limsi.fr

En omdat we nog niet voldoende tools hebben



- http://systematicreviewtools.com/