

Machine learning tools are now available
for use in Cochrane reviews!
Try them out and discuss how they should
— and shouldn't — be used

James Thomas, Claire Stansfield, Alison O'Mara-Eves, Ian Shemilt
Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre)
Social Science Research Unit
UCL Institute of Education
University College London

Declaration of interests and funding

- James Thomas is co-lead of the Cochrane ‘Transform’ project, which is implementing some of the technologies discussed here. He also directs development & management of EPPI-Reviewer, the EPPI-Centre’s software for systematic reviews.
- Parts of this work funded by: Cochrane, JISC, Medical Research Council, National Health & Medical Research Council (Australia), Wellcome Trust. All views expressed are my own, and not necessarily those of these funders.

Objectives

- Demonstrate the range of machine learning tools which Cochrane authors can use in their reviews
- Try out machine learning technologies
- Discuss their use in Cochrane reviews
- Links to tools: <http://eppi.ioe.ac.uk/> (under 'resources' tab)

Automation in systematic reviews – what can be done?

- Study identification:
 - Assisting search development
 - Citation screening
 - Updating reviews
 - RCT classifier
- Mapping research activity
- Data extraction
 - Risk of Bias assessment
 - Other study characteristics
 - Extraction of statistical data
- Synthesis and conclusions



**What is a
classifier?**



What does a classifier do?

- It takes as its input the title and abstract describing a publication
- It outputs a ‘probability’ score – between 0 and 1 which indicates how likely the publication is to being the ‘positive class’ (e.g. is an RCT)
- Classification is an integral part of the ‘evidence pipeline’

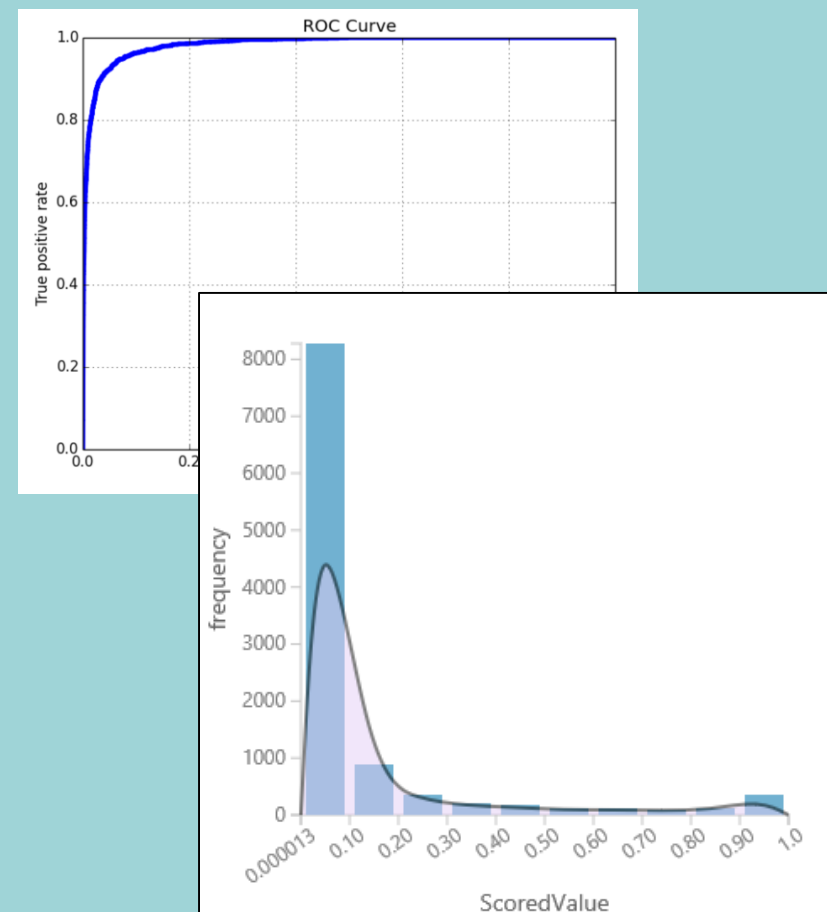
Pre-built or build your own

- Pre-built
 - Developed from established datasets
 - RCT model
 - Systematic review model
 - Economic evaluation
- Build your own



Pre-built classifier

- An RCT classifier was built using more than 280,000 records from Cochrane Crowd
- 60% of the studies have scores < 0.1
- If we trust the machine, and automatically exclude these citations, we're left with 99.897% of the RCTs (i.e. we lose 0.1%)
- Is that good enough?
- Systematic review community needs to discuss appropriate uses of automation



**Demo - RCT
classifier
*EPPI-Reviewer 4***



<http://eppi.ioe.ac.uk/eppireviewer4/>

Testing three models for TROPHI register of controlled trials

N=9,431 records

	Pre-built RCT classifier		Build your own classifier			
			Best		Second best	
	RCTs	NonRCTs	RCTs	NonRCTs	RCTs	NonRCTs
Precision = relevant items scored 11-99/total number of items scored 11-99	12%	3%	17%	5%	12%	4%
Recall = relevant items scored 11-99/all relevant items	99%	86%	99%	99%	99%	100%
Screening reduction	43%		58%		41%	

Build your own classifier



**Demo - DIY
classifier
*EPPI-Reviewer 4***



<http://eppi.ioe.ac.uk/eppireviewer4/>

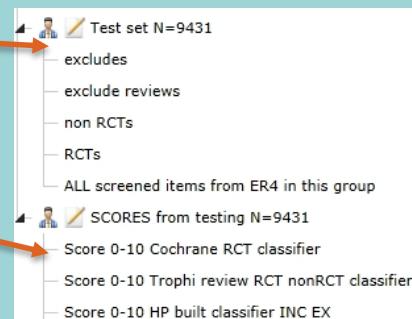
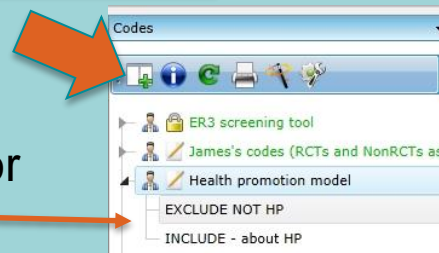
How to build your own

To build a classifier you need a **development set** of known includes and excludes

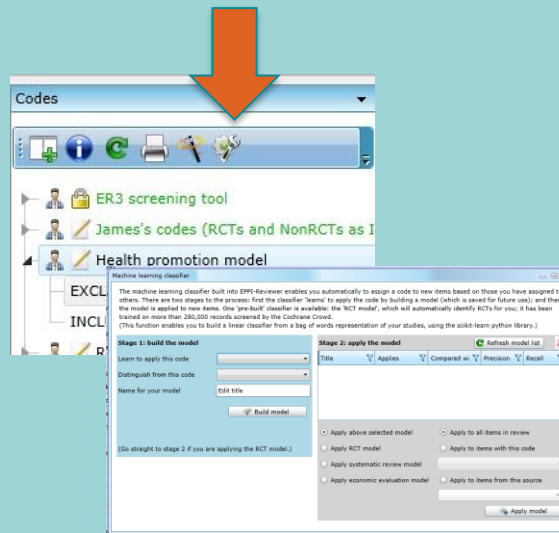
To test the classifier you need a **test set** of includes and excludes

1. Create codesets

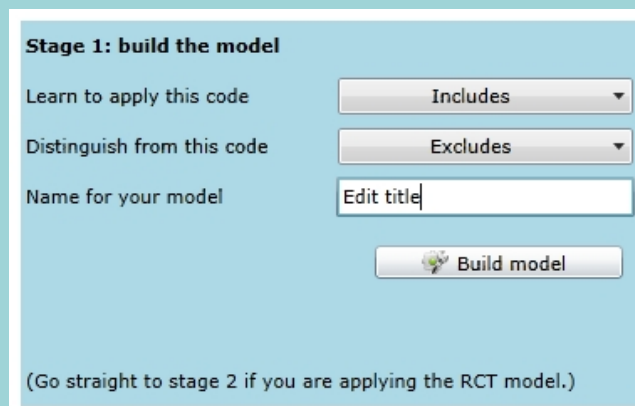
- i) include and exclude codes for the development set
- ii) a test codeset
- iii) a score codeset



2. **Click** on the spanner 'classifier' icon to get the Machine building classifier menu



3. Build the model.
Apply the include code from exclude code.
Name the model.



Go to stage 2

4. Select a model

5. Select the items to apply to the model

6. Choose the Search tab for the results.

7. Click 'Select'

Stage 2: apply the model Refresh model list

Title	Applies	Compared wi	Precision	Recall
ER3 TROPHI m	EXCLUDE	INCLUDE	0.99	0.95
RCT and nonRC	INCLUDE	EXCLUDE	0.46	0.79
HP model	INCLUDE - about	EXCLUDE NOT H	0.59	0.78
Reviews only	Includes	Excludes	0.35	0.86

☒ Apply above selected model
 ☐ Apply to all items in review
☐ Apply RCT model
 ☐ Apply to items with this code
☐ Apply systematic review model
 ☐ Apply economic evaluation model
 ☐ Apply to items from this source

Apply model

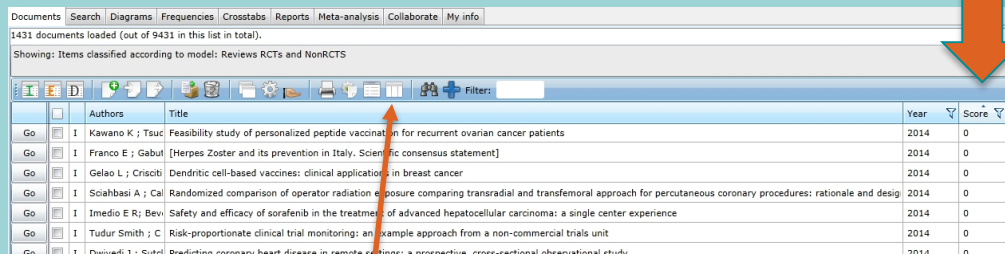
Search Diagrams Frequencies Crosstabs Reports Meta-analysis Collaborate My info

New search Refresh search list Delete selected Combine AND OR NOT (included)

Title	Created by	Date	Hits	List
62 Items classified according to model: RCTs only	Claire Stansfield	30/11/2016	9431	Select
61 Items classified according to model: NonRCTs only	Claire Stansfield	30/11/2016	9431	Select
60 Items classified according to model: Reviews RCTs and NonRCTs	Claire Stansfield	30/11/2016	9431	Select
59 Items classified according to model: Reviews only	Claire Stansfield	30/11/2016	9431	Select
58 Items classified according to model: Reviews only	Claire Stansfield	30/11/2016	9431	Select
57 Items classified according to model: RCT and nonRCTs model TROPHI	Claire Stansfield	30/11/2016	9431	Select
56 Items classified according to model: RCT and nonRCTs model TROPHI	Claire Stansfield	30/11/2016	9431	Select
55 Items classified according to model: HP model	Claire Stansfield	30/11/2016	9431	Select
54 Items classified according to model: HP model	Claire Stansfield	30/11/2016	9431	Select
53 Items classified according to model: RCT	Claire Stansfield	29/11/2016	9431	Select
52 Items classified according to model: RCT	Claire Stansfield	29/11/2016	9431	Select

ER3 screening tool
 James's codes (RCT)
 Health promotion m
 Review model
 Review RCT and no
 nonRCTmodel
 RCTmodel
 Screened items from
 SCORES from testin

The results are displayed.
A Score tab has appeared. The
items are ranked from 0 to 99



Documents | Search | Diagrams | Frequencies | Crosstabs | Reports | Meta-analysis | Collaborate | My info

1431 documents loaded (out of 9431 in this list in total).

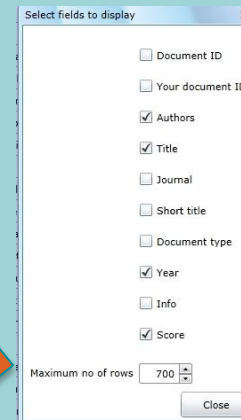
Showing: Items classified according to model: Reviews RCTs and NonRCTs

Filter:

	Authors	Title	Year	Score
Go	1 Kawano K ; Tsuc	Feasibility study of personalized peptide vaccination for recurrent ovarian cancer patients	2014	0
Go	1 Franco E ; Gabul	[Herpes Zoster and its prevention in Italy. Scientific consensus statement]	2014	0
Go	1 Gelao L ; Criscio	Dendritic cell-based vaccines: clinical applications in breast cancer	2014	0
Go	1 Sciahbasi A ; Cal	Randomized comparison of operator radiation exposure comparing transradial and transfemoral approach for percutaneous coronary procedures: rationale and design	2014	0
Go	1 Imedio E R ; Bevi	Safety and efficacy of sorafenib in the treatment of advanced hepatocellular carcinoma: a single center experience	2014	0
Go	1 Tudur Smith ; C	Risk-proportionate clinical trial monitoring: an example approach from a non-commercial trials unit	2014	0
Go	1 Delucchi T ; Sured	Prediction coronary heart disease in remote regions: a retrospective, cross-sectional observational study	2014	0

8. **Click** on the Column icon.

9. **Change** the maximum no. of
rows to 4,000.



Select fields to display

☐ Document ID

☐ Your document ID

☒ Authors

☒ Title

☐ Journal

☐ Short title

☐ Document type

☒ Year

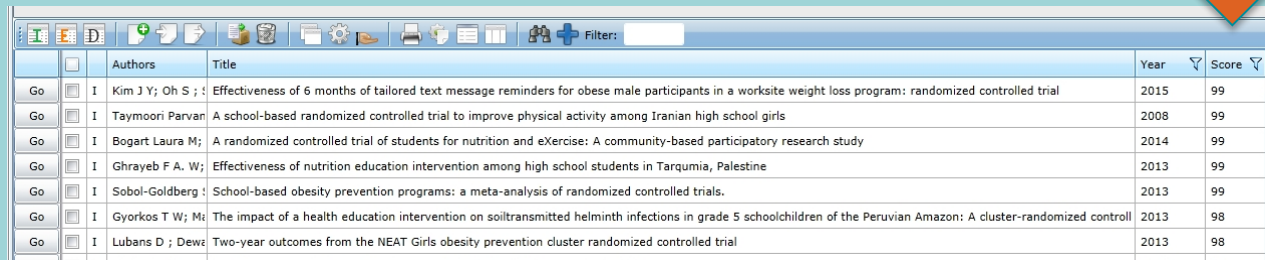
☐ Info

☒ Score

Maximum no of rows 700

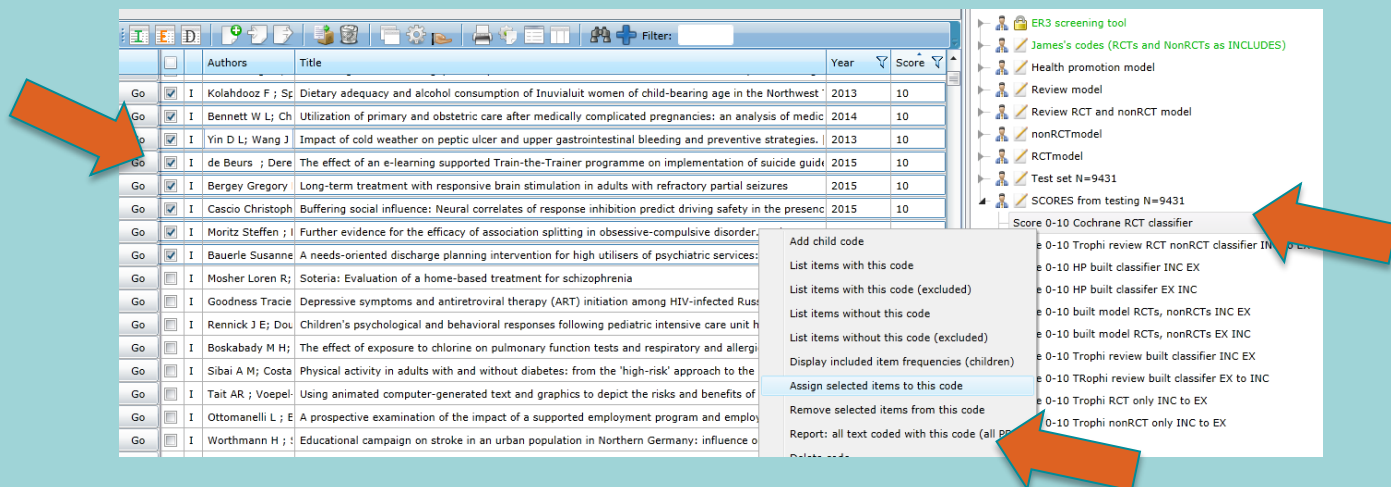
Close

10. Click on **score**. This orders items by score



	Authors	Title	Year	Score
Go	Kim J Y; Oh S ;	Effectiveness of 6 months of tailored text message reminders for obese male participants in a worksite weight loss program: randomized controlled trial	2015	99
Go	Taymoori Parvan	A school-based randomized controlled trial to improve physical activity among Iranian high school girls	2008	99
Go	Bogart Laura M;	A randomized controlled trial of students for nutrition and eExercise: A community-based participatory research study	2014	99
Go	Ghrayeb F A. W;	Effectiveness of nutrition education intervention among high school students in Tarqumia, Palestine	2013	99
Go	Sobol-Goldberg	School-based obesity prevention programs: a meta-analysis of randomized controlled trials.	2013	99
Go	Gyorkos T W; M;	The impact of a health education intervention on soiltransmitted helminth infections in grade 5 schoolchildren of the Peruvian Amazon: A cluster-randomized control	2013	98
Go	Lubans D ; Dewe	Two-year outcomes from the NEAT Girls obesity prevention cluster randomized controlled trial	2013	98

11. for **each page** of citations,
highlight the items coded 0-10 (Ctrl and drag with mouse)
assign to the score code
 (left click on code and click 'Assign selected items to this code')



	Authors	Title	Year	Score
Go	Kolahdoz F ; Sp	Dietary adequacy and alcohol consumption of Inuvialuit women of child-bearing age in the Northwest	2013	10
Go	Bennett W L; Ch	Utilization of primary and obstetric care after medically complicated pregnancies: an analysis of medic	2014	10
Go	Yin D L; Wang J	Impact of cold weather on peptic ulcer and upper gastrointestinal bleeding and preventive strategies.	2013	10
Go	de Beurs ; Dere	The effect of an e-learning supported Train-the-Trainer programme on implementation of suicide guide	2015	10
Go	Bergey Gregory	Long-term treatment with responsive brain stimulation in adults with refractory partial seizures	2015	10
Go	Cascio Christoph	Buffering social influence: Neural correlates of response inhibition predict driving safety in the presenc	2015	10
Go	Moritz Steffen ;	Further evidence for the efficacy of association splitting in obsessive-compulsive disorder,		
Go	Bauerle Susanne	A needs-oriented discharge planning intervention for high utilisers of psychiatric services:		
Go	Mosher Loren R;	Soteria: Evaluation of a home-based treatment for schizophrenia		
Go	Goodness Tracie	Depressive symptoms and antiretroviral therapy (ART) initiation among HIV-infected Rus		
Go	Rennick J E; Dou	Children's psychological and behavioral responses following pediatric intensive care unit h		
Go	Boskabady M H;	The effect of exposure to chlorine on pulmonary function tests and respiratory and allergi		
Go	Sibai A M; Costa	Physical activity in adults with and without diabetes: from the 'high-risk' approach to the		
Go	Tait AR ; Voepel	Using animated computer-generated text and graphics to depict the risks and benefits of		
Go	Ottomanelli L ; E	A prospective examination of the impact of a supported employment program and employ		
Go	Worthmann H ;	Educational campaign on stroke in an urban population in Northern Germany: influence o		

ER3 screening tool

James's codes (RCTs and NonRCTs as INCLUDES)

Health promotion model

Review model

Review RCT and nonRCT model

nonRCTmodel

RCTmodel

Test set N=9431

SCORES from testing N=9431

Score 0-10 Cochrane RCT classifier

0-10 Trophi review RCT nonRCT classifier INC EX

0-10 HP built classifier INC EX

0-10 HP built classifier EX INC

0-10 built model RCTs, nonRCTs INC EX

0-10 built model RCTs, nonRCTs EX INC

0-10 Trophi review built classifier INC EX

0-10 Trophi review built classifier EX to INC

0-10 Trophi RCT only INC to EX

0-10 Trophi nonRCT only INC to EX

Add child code

List items with this code

List items with this code (excluded)

List items without this code

List items without this code (excluded)

Display included item frequencies (children)

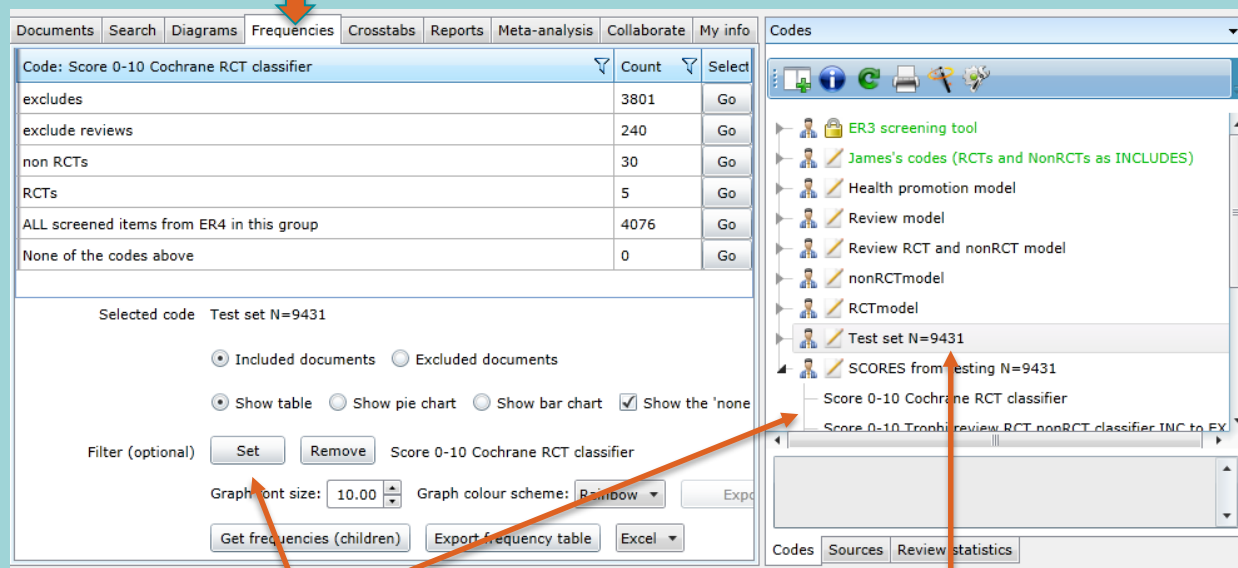
Assign selected items to this code

Remove selected items from this code

Report: all text coded with this code (all PR

Delete code

12. Use the **frequency tab** to compare results for the code (these are excluded items with a score of 0-10)



Code: Score 0-10 Cochrane RCT classifier	Count	Select
excludes	3801	Go
exclude reviews	240	Go
non RCTs	30	Go
RCTs	5	Go
ALL screened items from ER4 in this group	4076	Go
None of the codes above	0	Go

Selected code Test set N=9431

☒ Included documents ☐ Excluded documents

☒ Show table ☐ Show pie chart ☐ Show bar chart ☒ Show the 'none'

Filter (optional) Score 0-10 Cochrane RCT classifier

Graph font size: 10.00 Graph colour scheme: Rainbow

Codes Sources Review statistics

Click on Score code, and on 'Set'

Click on test set codeset

Click on

Study identification



Using text mining for study identification in systematic reviews: a systematic review of current approaches

Alison O'Mara-Eves¹, James Thomas^{1*}, John McNaught², Makoto Miwa³ and Sophia Ananiadou²

Abstract

Background: The large and growing number of published studies, and their increasing rate of publication, makes the task of identifying relevant studies in an unbiased way for inclusion in systematic reviews both complex and time consuming. Text mining has been offered as a potential solution: through automating some of the screening process, reviewer time can be saved. The evidence base around the use of text mining for screening has not yet been pulled together systematically; this systematic review fills that research gap. Focusing mainly on non-technical issues, the review aims to increase awareness of the potential of these technologies and promote further collaborative research between the computer science and systematic review communities.

Methods: Five research questions led our review: what is the state of the evidence base; how has workload reduction been evaluated; what are the purposes of semi-automation and how effective are they; how have key contextual problems of applying text mining to the systematic review field been addressed; and what challenges to

Citation screening

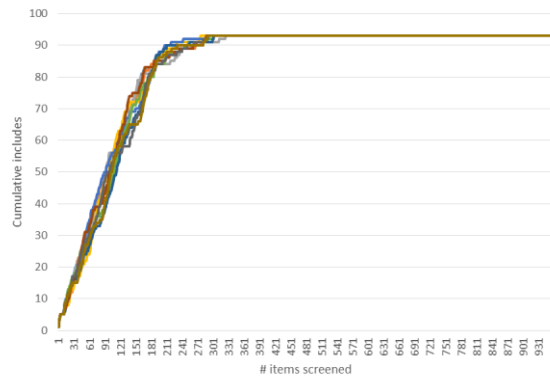
- Has received most R&D attention
- Diverse evidence base; difficult to compare evaluations
- 'semi-automated' approaches are the most common
- Possible reductions in workload in excess of 30% (and up to 97%)

Summary of conclusions

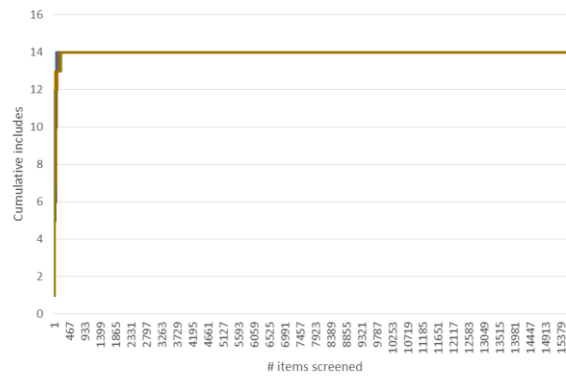
- Screening prioritisation
 - 'safe to use'
- Machine as a 'second screener'
 - Use with care
- Automatic study exclusion
 - Highly promising in many areas, but performance varies significantly depending on the domain of literature being screened

Does it work? e.g. reviews from Cochrane Heart Group

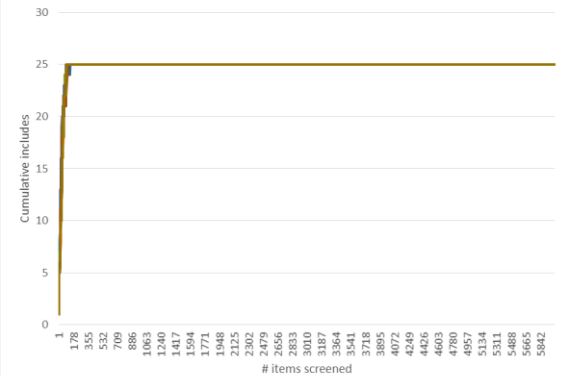
Review 0902



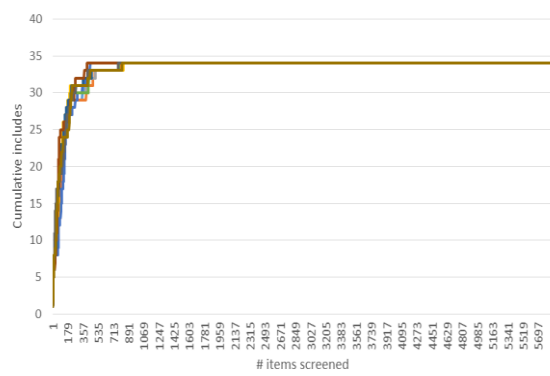
Review 1006



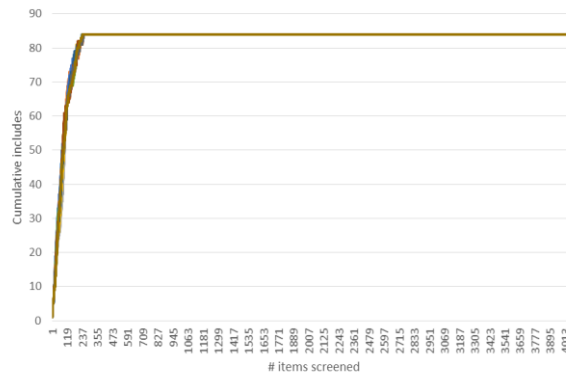
Review 1007



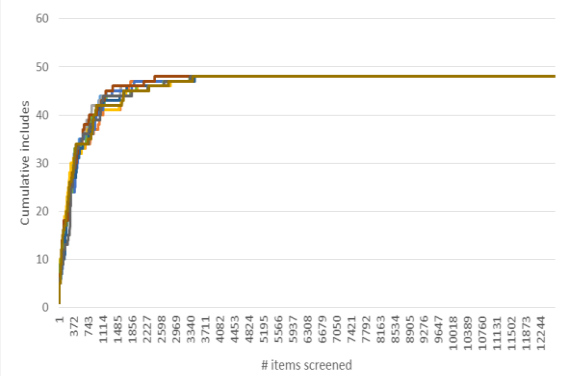
Review 1004



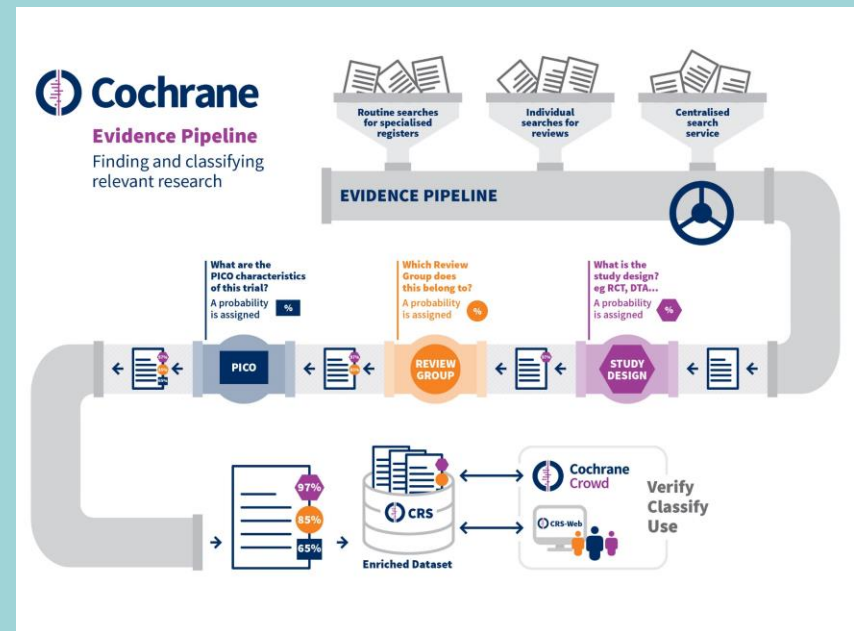
Review 1125

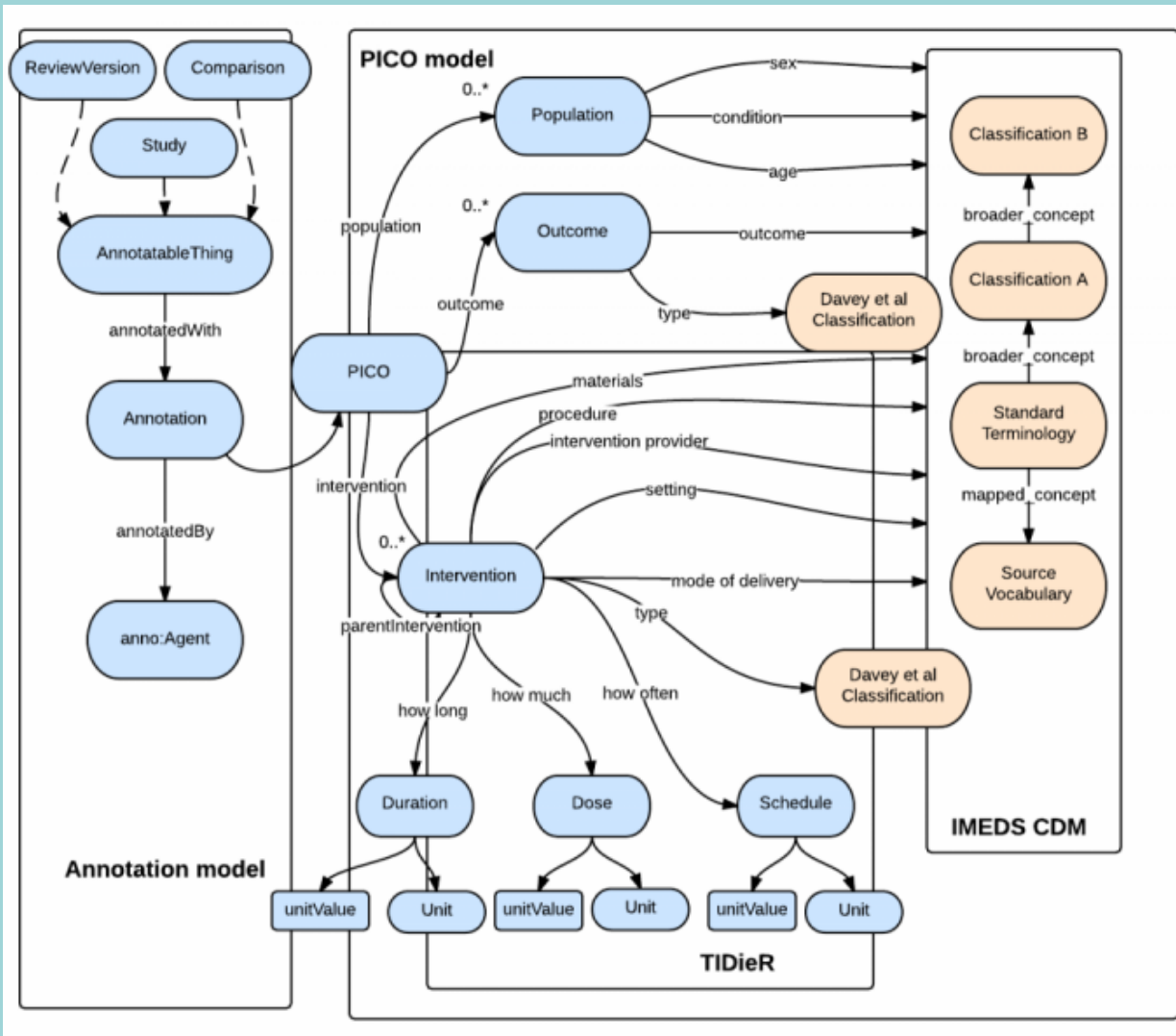


Review 1309



Cochrane Evidence Pipeline

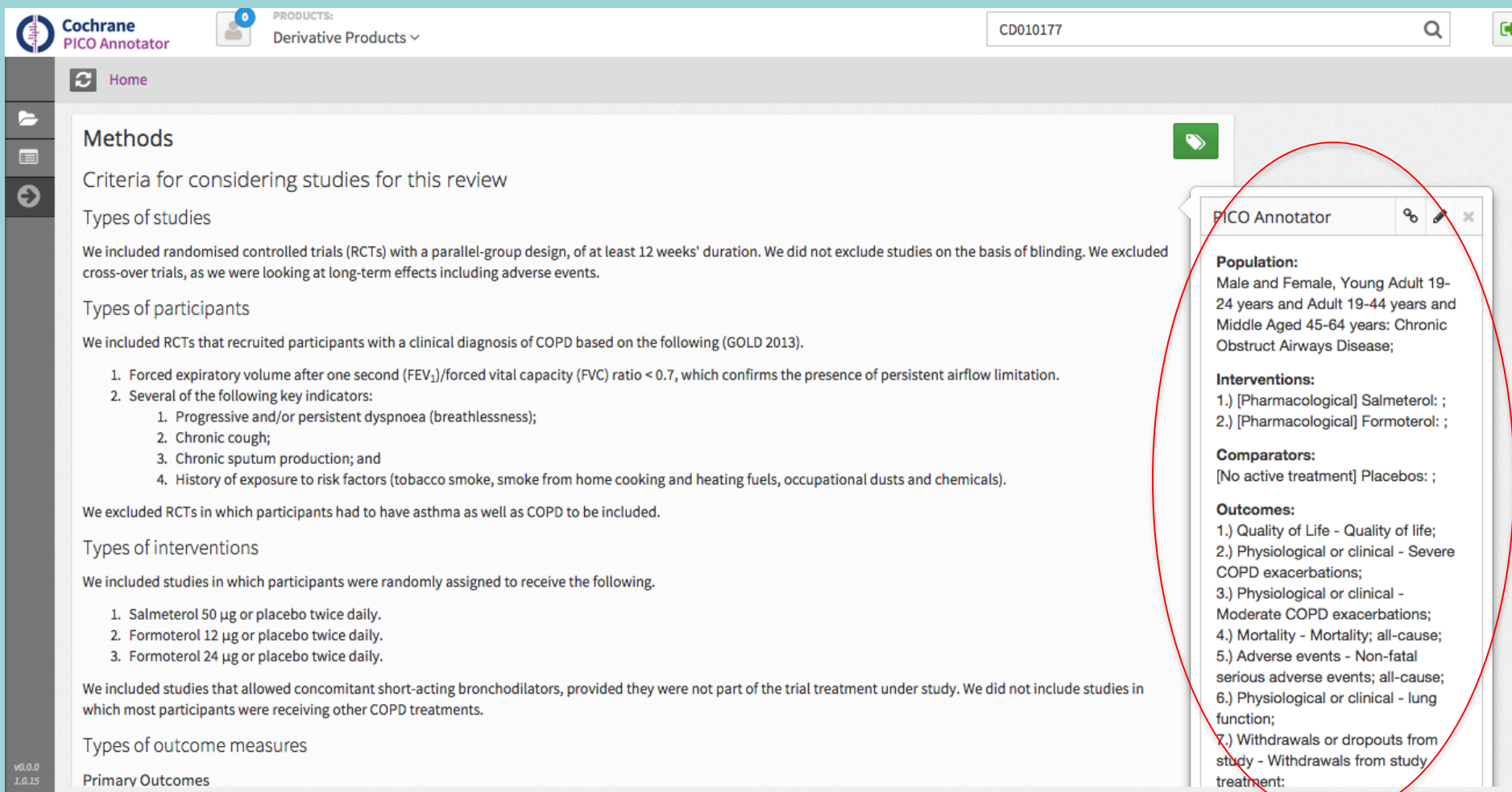




A PICO 'ontology' is being developed in Cochrane

... and is being applied to...

... all Cochrane reviews and all the trials they contain



Cochrane PICO Annotator PRODUCTS: Derivative Products ▾ CD010177

Home

Methods

Criteria for considering studies for this review

Types of studies

We included randomised controlled trials (RCTs) with a parallel-group design, of at least 12 weeks' duration. We did not exclude studies on the basis of blinding. We excluded cross-over trials, as we were looking at long-term effects including adverse events.

Types of participants

We included RCTs that recruited participants with a clinical diagnosis of COPD based on the following (GOLD 2013).

1. Forced expiratory volume after one second (FEV₁)/forced vital capacity (FVC) ratio < 0.7, which confirms the presence of persistent airflow limitation.
2. Several of the following key indicators:
 1. Progressive and/or persistent dyspnoea (breathlessness);
 2. Chronic cough;
 3. Chronic sputum production; and
 4. History of exposure to risk factors (tobacco smoke, smoke from home cooking and heating fuels, occupational dusts and chemicals).

We excluded RCTs in which participants had to have asthma as well as COPD to be included.

Types of interventions

We included studies in which participants were randomly assigned to receive the following.

1. Salmeterol 50 µg or placebo twice daily.
2. Formoterol 12 µg or placebo twice daily.
3. Formoterol 24 µg or placebo twice daily.

We included studies that allowed concomitant short-acting bronchodilators, provided they were not part of the trial treatment under study. We did not include studies in which most participants were receiving other COPD treatments.

Types of outcome measures

Primary Outcomes

PICO Annotator

Population:
Male and Female, Young Adult 19-24 years and Adult 19-44 years and Middle Aged 45-64 years: Chronic Obstruct Airways Disease;


Interventions:
1.) [Pharmacological] Salmeterol; ;
2.) [Pharmacological] Formoterol; ;

Comparators:
[No active treatment] Placebos; ;

Outcomes:
1.) Quality of Life - Quality of life;
2.) Physiological or clinical - Severe COPD exacerbations;
3.) Physiological or clinical - Moderate COPD exacerbations;
4.) Mortality - Mortality; all-cause;
5.) Adverse events - Non-fatal serious adverse events; all-cause;
6.) Physiological or clinical - lung function;
7.) Withdrawals or dropouts from study - Withdrawals from study treatment;

vd.0.0
1.0.15

... Boolean searches are replaced by the specification of the 'PICO' of interest


Cochrane PICOfinder
 Powered by Cochrane linked data

Population

⚡ condition

🕒 age

👤 sex


Intervention / Comparator

⚙️ classification

💧 materials / procedures

Outcome

♥️ classification


Cochrane PICOfinder
 Powered by Cochrane linked data

Population

⚡ condition

Asthma
SNOMED 195967001 112

Dementia
SNOMED 52448006 26

Dementia Due To Alzheimer's Disease
SNOMED 142811000119104 28

Acute Asthma
SNOMED 304527002 25

Mild Cognitive Impairment
SNOMED 888271000000101 17

Asthma
MedDRA 10003553 12

Vascular Dementia
SNOMED 429998004 11

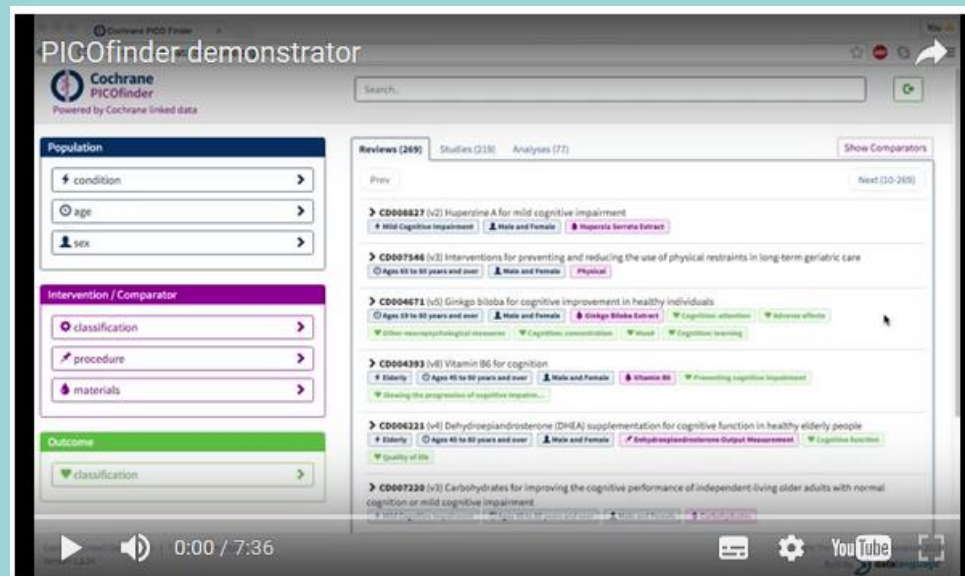
Chronic Obstructive Pulmonary Disease
MedDRA 10009033 7

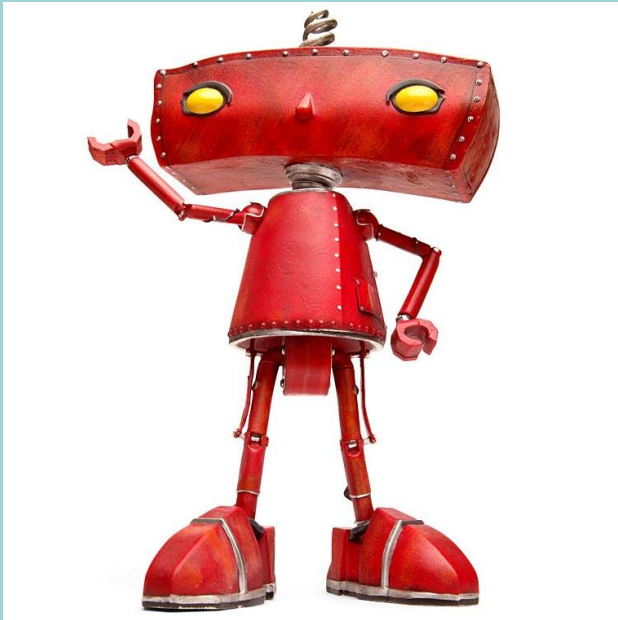
Elderly
MedDRA 10014348 6

Mixed Dementia
SNOMED 79341000119107 3

PICOfinder

<https://youtu.be/WtqAnL6QPt4>





Through a combination of human and machine effort the aim is to identify and classify ALL trials using this system.

Identifying studies for systematic reviews* will then be a simple process of specifying the relevant PICO

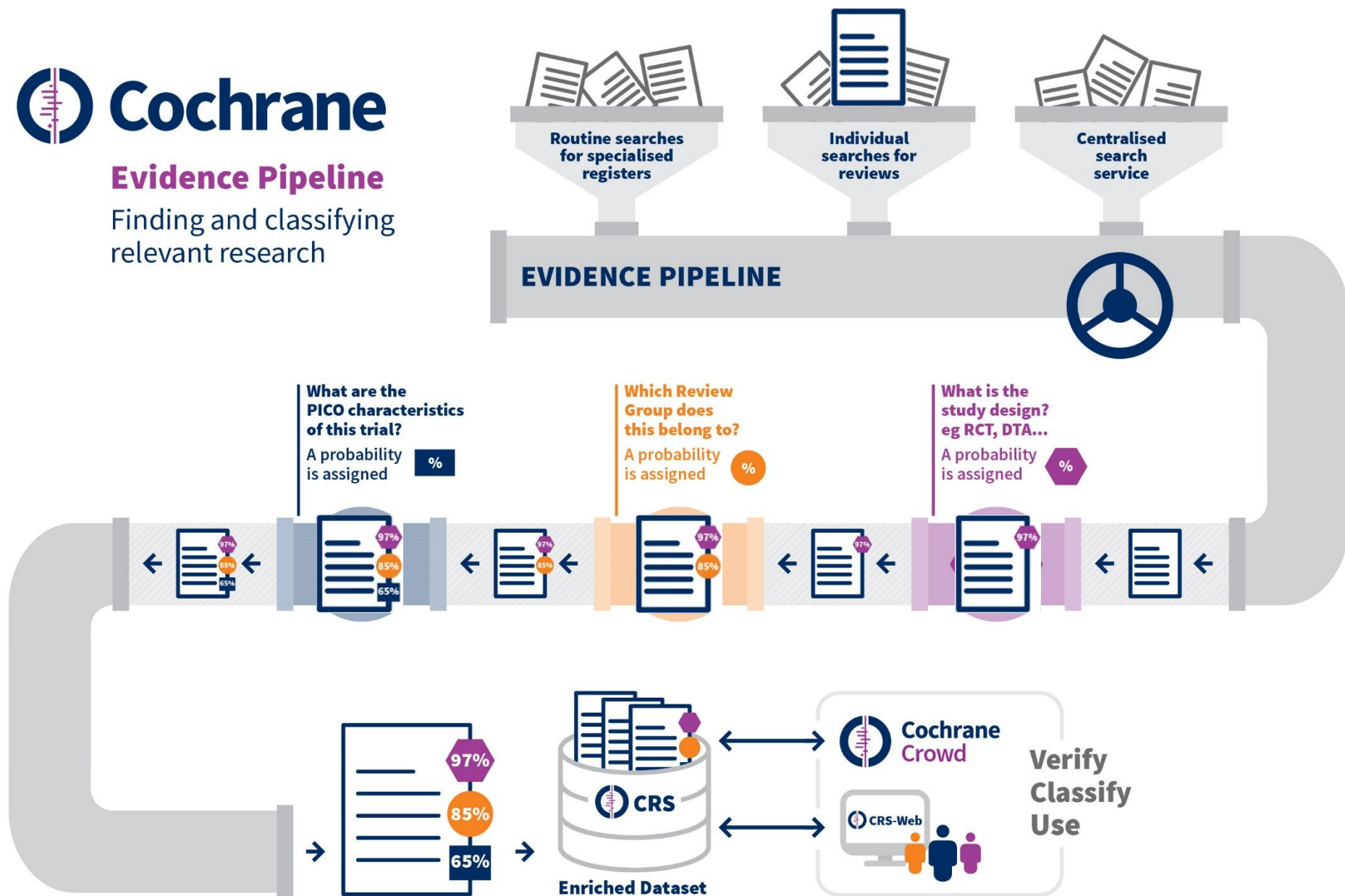
* Of RCTs

A screenshot of the Cochrane Crowd website. The browser address bar shows 'crowd.cochrane.org/index.html'. The website has a white header with the Cochrane Crowd logo on the left, the tagline 'Trusted evidence. Informed decisions. Better health.' in the center, and 'Login' and 'Sign up' buttons on the right. The main content area has a purple overlay with a background image of a woman and a child. The text 'You can make a difference' is prominently displayed. Below it, a paragraph invites users to become Cochrane citizen scientists. A 'Give it a try' button is centered. At the bottom, three statistics are shown: '2480 Contributors', '31 Countries', and '796040 Classifications'.



Evidence Pipeline

Finding and classifying relevant research



CRS-Web

CRS Web (online)

crsdemo.metaxis.com/index.php#Search

Search Simple MeSH Classifier Saved Tracking

Classifier search

Records that have been through the classifier have probabilities assigned to them to indicate how likely they are to have certain properties, like whether they are of interest to a review group, or whether they are likely to be an RCT. Choose the classifier model you are interested in, set the model parameters and click Search to find the records

RCT

Approximately 32129 records that are between 99 and 100 percent likely to be of interest

Search

You can find your records that are currently being processed by the classifier by searching for INPROCESS:CLASSIFIER

Find those records

Cochrane Register of Studies

Anna Noel-Storr [DEMENTIA] | logout

Dashboard Records Import Journals CT.GOV Reports To do

Search Layout1 Layout2 Layout3 Layout4 Deduplication New reference New study Users Help

Search results

(399 records) Page 1 of 8

#	Title	Author
1	Cognitive effects of treating obstructive sleep apnea in Alzheimer's disease: a randomized controlled study	Ancoli-Israel S // Palmer BW // Cooke
2	Efficacy of galantamine in probable vascular dementia and Alzheimer's disease combined with cerebrovascular disease: a rando...	Erkinjuntti T // Kurz A // Gauthier S //
3	Donepezil improved memory in multiple sclerosis in a randomized clinical trial	Krupp LB // Christodoulou C // Melvil
4	A randomized, 26-week, double-blind, placebo-controlled trial to evaluate the safety and efficacy of galantamine in the treatme...	Auchus A
5	A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease. Donepezil Study Group	Rogers SL // Farlow MR // Doody RS /
6	A Controlled, Double-Blind, Randomized Pilot Clinical Trial of Hydroxysafflor Yellow a on Cognitive Function in Patients With Vas...	Tian J

Record

Fields Duplicates Links Reviews Classifier Files Audit CENTRAL REGISTER

The bar chart below shows the classifier scores for this record. Scores are presented in the range 0 -1 00 where higher scores mean a higher likelihood that the record is of interest to the group. You can tell a group about this record if it doesn't already have it in its segment by clicking the bar for that group.

In register In segment Not in segment Not relevant to my group

There is a 99% likelihood that this record is an RCT [Confirm this is **not** an RCT] [Confirm this is an RCT]

Mapping research activity



Mapping research activity

- It is possible to apply ‘keywords’ to text automatically, without needing to ‘teach’ the machine beforehand
- This relies on ‘clustering’ technology – which groups studies which use similar combinations of words
- Very few evaluations
 - Can be promising, especially when time is short
 - But users have no control on the terms actually used

Original Article

Research Synthesis Methods

Received 23 November 2012,

Revised 21 March 2013,

Accepted 21 April 2013

Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/jrsm.1082

‘Clustering’ documents automatically to support scoping reviews of research: a case study

Claire Stansfield,^{*,†} James Thomas[†] and Josephine Kavanagh[†]

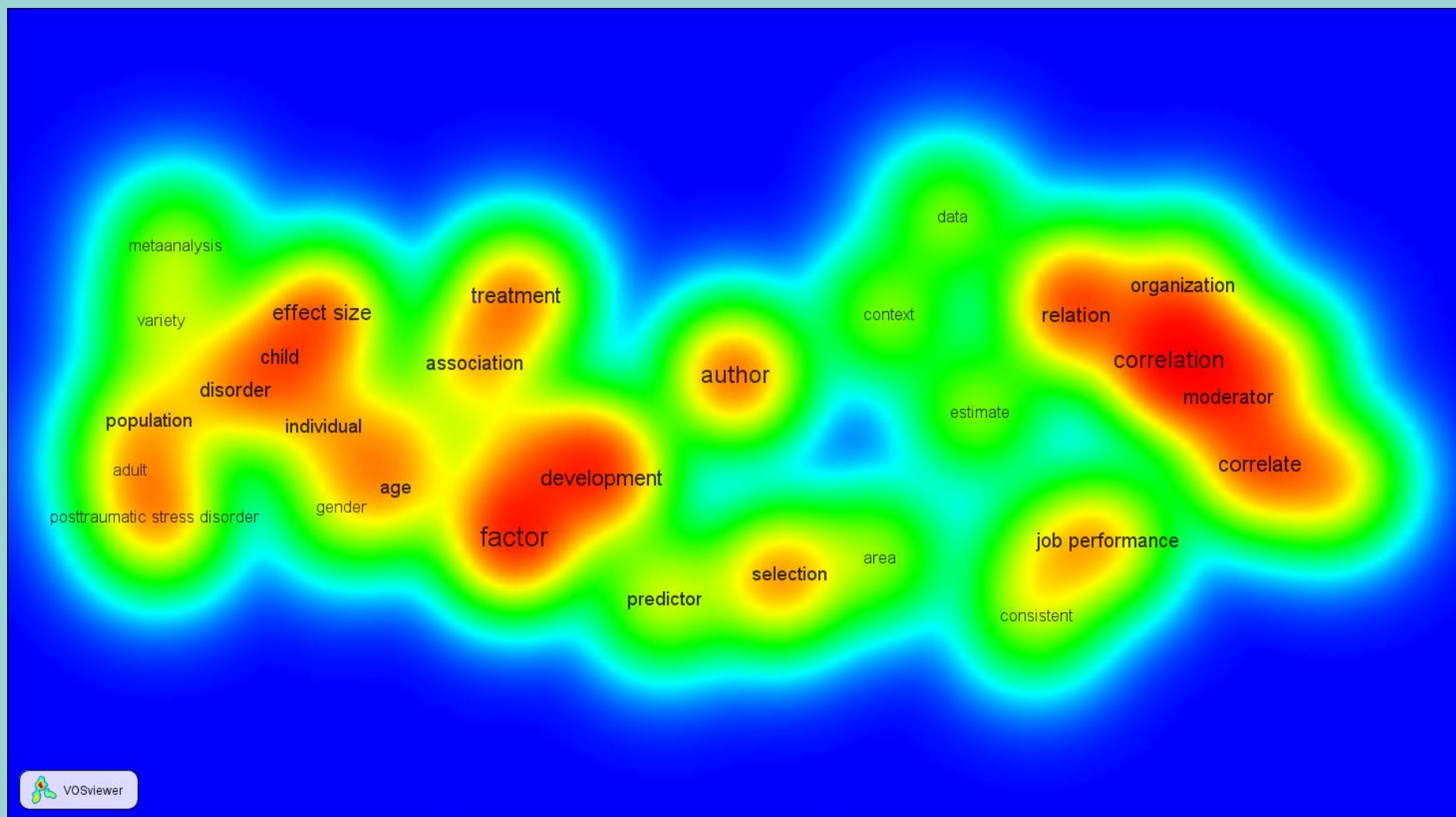
Background: Scoping reviews of research help determine the feasibility and the resource requirements of conducting a systematic review, and the potential to generate a description of the literature quickly is attractive.

Aims: To test the utility and applicability of an automated clustering tool to describe and group research studies to improve the efficiency of scoping reviews.

Methods: A retrospective study of two completed scoping reviews was conducted. This compared the

Technologies for identifying subsets of citations

- Different families of techniques
 - Fairly simple approaches which examine term frequencies to group similar citations
 - More complex approaches, such as Latent Dirichlet Allocation (LDA)
- The difficult part is finding good labels to describe the clusters
 - But are labels always needed?
- Visualisations are often incorporated into tools



Demo – Topic modelling *pyLDAvis*



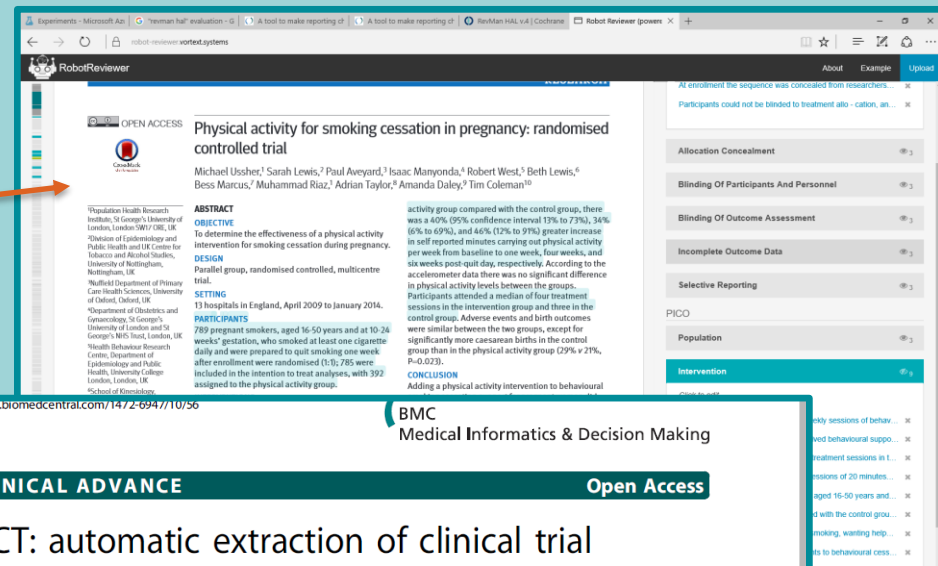
<http://eppi.ioe.ac.uk/ldavis/index.html#topic=6&lambda=0.63&term=>

Data extraction; synthesis and conclusions



Data extraction

- RobotReviewer can identify phrases relating to study PICO characteristics
- ExaCT extracts trial characteristics (e.g. eligibility criteria)
- Systematic review found that no unified framework yet exists
- More evaluative work is needed on larger datasets
- Further challenges include extraction of data from tables and graphs



TECHNICAL ADVANCE

Open Access

ExaCT: automatic extraction of clinical trial characteristics from journal publications

Svetlana Kiritchenko^{1*}, Berry de Bruijn¹, Simona Carini², Joel Martin¹, Ida Sim²

Abstract

Background: Clinical trial publications are a key source of information for systematic reviews. However, the process of extracting data from these publications is often laborious and error-prone.

Methods: We developed ExaCT, a system for automatically extracting clinical trial characteristics from journal publications.

Results: We evaluated ExaCT on a dataset of 1,000 clinical trial publications. The system achieved a precision of 91% and a recall of 91%.

Conclusions: ExaCT is a promising tool for automatically extracting clinical trial characteristics from journal publications.

Jonnalagadda et al. *Systematic Reviews* (2015) 4:78
DOI 10.1186/s13643-015-0066-7

RESEARCH



Open Access



Automating data extraction in systematic reviews: a systematic review

Siddhartha R. Jonnalagadda^{1*}, Pawan Goyal² and Mark D. Huffman³

Abstract

Background: Automation of the parts of systematic review process, specifically the data extraction step, may be an important strategy to reduce the time necessary to complete a systematic review. However, the state of the science of automatically extracting data elements from full texts has not been well described. This paper performs a systematic review of published and unpublished methods to automate data extraction for systematic reviews.

Methods: We systematically searched PubMed, IEEEExplore, and ACM Digital Library to identify potentially relevant

Risk of Bias assessment

- Emerging area; e.g.
 - RobotReviewer
 - Millard, Flach and Higgins
- Tools can accomplish two purposes:
 - 1. identify relevant text in the document
 - 2. automatically assess risk of bias
- Can perform very well though authors do not yet suggest well enough to replace humans

Int. J. Epidemiol. Advance Access published December 8, 2015

IEA International Journal of Epidemiology, 2015, 1-12
doi: 10.1093/ije/dyv306
Original article

Original article

Machine learning to assist risk-of-bias assessments in systematic reviews

Louise A.C. Millard,^{1,2,3*} Peter A. Flach^{1,3} and Julian P.T. Higgins^{1,2}

¹MRC Integrative Epidemiology Unit, ²School of Social and Community Medicine and ³Intelligent Systems Laboratory, University of Bristol, Bristol, UK

RobotReviewer

pregnancy.

Methods
Study population
Between April 2009 and November 2012, we recruited pregnant women by telephone after their first antenatal booking visit (conducted at either hospital or a community clinic), from 13 hospital antenatal clinics in London, Surrey, Kent, and Cheshire. In the United Kingdom all pregnant women are booked for delivery in the secondary care setting, although a small fraction of women deliver at home or in primary care (midwife-led) units. Pregnant smokers are, by definition, high risk and would be expected to be cared for in hospital. Inclusion criteria were wanting to stop smoking, wanting help with stopping smoking, agreeing to set a date for quitting smoking within one week of the baseline visit, age 16-50 years, being at 10-24 weeks of gestation, cigarette consumption of five or more daily before pregnancy, currently smoking one or more cigarettes daily, and being able to walk continuously for 15 minutes. Exclusion criteria were medical conditions potentially exacerbated by exercise or advised against exercise by a doctor, inability to provide informed consent or complete questionnaires in English, drug or alcohol dependence, and currently using or wanting to use nicotine replacement therapy. We recruited women irrespective of their current level of physical activity or motivation towards increasing their activity.

Study protocol and interventions
Wandsworth research ethics committee approved the published protocol¹⁸ (available at www.trialsjournal.com/content/13/1/186). All participants received written

physical activity into women's lives, to motivate them to use physical activity to reduce the urge to smoke, and to help them use behavioural strategies to improve adherence to these plans. These 20 minute consultations incorporated 19 behaviour change techniques, as described in the study protocol.¹⁵ The women were advised to be active for at least 10 minutes at a time, progressing towards 30 minutes of activity on at least five days a week. The emphasis was on brisk walking, and pedometers (Digi-Walker SW200; Yamax, Nottingham, UK) were supplied, with researchers setting individualised step count targets. We also provided a DVD on antenatal exercise. On the other occasion the women received behavioural support for smoking sessions (up to six sessions) as for the control group. For each session attended, participants were paid £7 towards travel.

Randomisation
An independent statistician generated a randomisation list using Stata, with random permuted blocks of random size stratified by recruitment centre, in a 1:1 ratio. At enrollment the sequence was concealed from researchers who confirmed consent and eligibility on an online database before allocation was revealed. It was not feasible to mask participants or researchers to group allocation.

Data collection
We collected personal and smoking characteristics of the women at baseline, including score on the Fagerstrom test for cigarette dependence,¹⁹ self reports of moderate-vigorous intensity physical activity in the previous week (bouts of ≥10 minutes) using the seven day physical activity recall interview,¹⁸ Edinburgh postnatal depression scale score,²⁰ confidence about taking up physical activity²¹ and stopping smoking,²² alcohol consumption,²³ weekly cigarette withdrawal symptoms,²⁴ and weekly smoking urges (combining ratings strength and frequency).²⁵ At the first antenatal booking visit the midwife measured the women's clothed weight (without shoes) on a digital scale. During all visits, the women were asked about adverse events. Research midwives examined the women's medical cards monthly for adverse events. The records were

Risk of Bias

Random Sequence Generation

Allocation Concealment

Overall risk of bias prediction: low

At enrollment the sequence was concealed from researchers...

Randomisation An independent statistician generated a random...

Participants could not be blinded to treatment allocation, an...

Blinding Of Participants And Personnel

Blinding Of Outcome Assessment

Incomplete Outcome Data

Selective Reporting

PICO

Population

Intervention

Click to edit

Participants were randomised to six weekly sessions of behav...

On the other occasion the women received behavioural suppo...

RobotReviewer

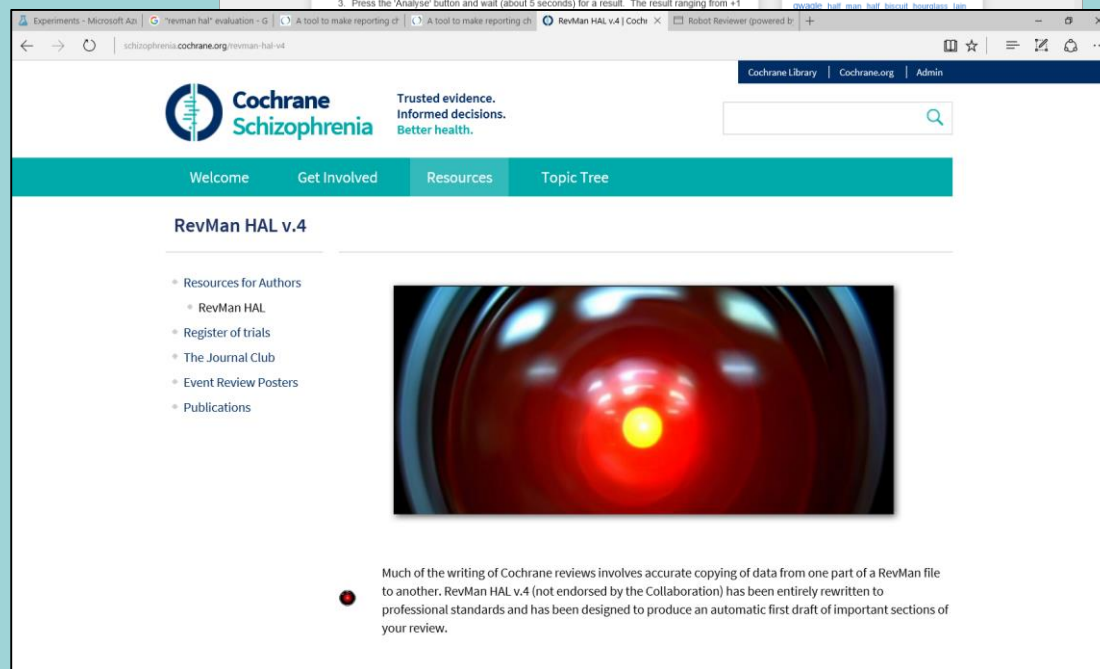
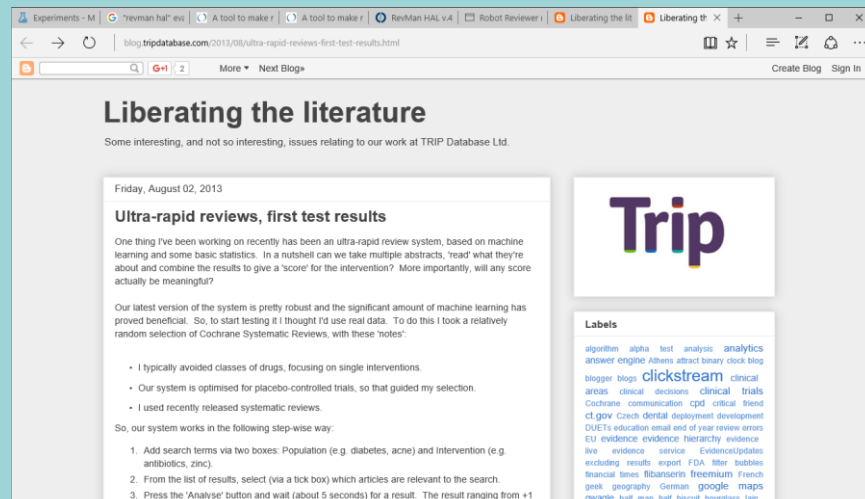
Demo - Data extraction *RobotReviewer*



<https://robot-reviewer.vortext.systems/>

Synthesis and conclusions

- Summarisation and synthesis of text is an active area for development in computer science
- Many hurdles to overcome before this technology can be used routinely
- Some systems automate parts of the process



Much of the writing of Cochrane reviews involves accurate copying of data from one part of a RevMan file to another. RevMan HAL v.4 (not endorsed by the Collaboration) has been entirely rewritten to professional standards and has been designed to produce an automatic first draft of important sections of your review.

Discussion



The wider picture: part of a wider evolution of systematic review methods

- Systematic reviews (as currently known) might change quite substantially
- From 'search strategy' to PICO definition
- From 'data extraction' to structured data (and IPD)
- We may choose to link trial data in new ways (e.g. via IPD to patient medical records)
- The 'systematic review' will become a matter of ascertaining the validity and utility of combining particular sets of studies at particular points in time, rather than the tedious trawling for, and extraction of, data – that they currently entail

Discussion and experimentation: in small groups:

How can Cochrane reviewers take advantage of the efficiencies these tools offer?

What methods and processes will need to be developed? How can we build an evidence base around them?

What are your concerns?

Are there other limitations?

Links to tools: <http://eppi.ioe.ac.uk/> (under 'resources' tab)

Thank you

SSRU website: <http://www.ioe.ac.uk/ssru/>

SSRU's EPPI website: <http://eppi.ioe.ac.uk>

Email

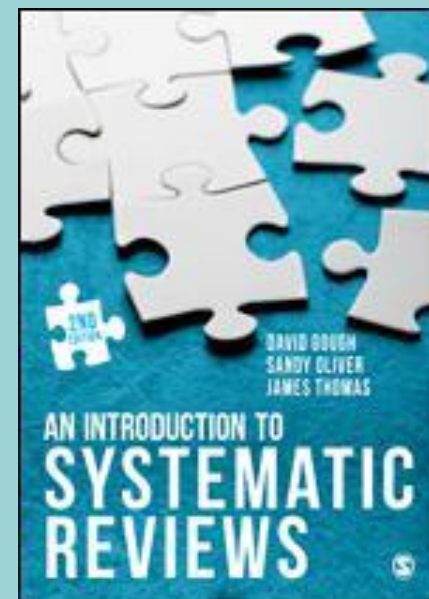
j.thomas@ucl.ac.uk

c.stansfield@ucl.ac.uk

a.omara-eves@ucl.ac.uk

i.Shemilt@ucl.ac.uk

The EPPI-Centre is part of the Social Science Research Unit at the UCL Institute of Education, University College London



EPPI-Centre

Social Science Research Unit
Institute of Education
University of London
18 Woburn Square
London WC1H 0NR

Tel +44 (0)20 7612 6397
Fax +44 (0)20 7612 6400
Email eppi@ioe.ac.uk
Web eppi.ioe.ac.uk/