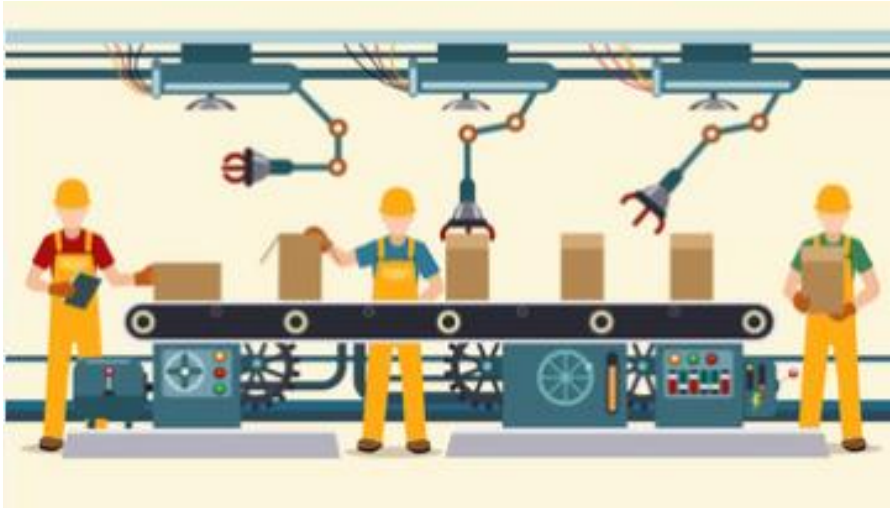
Semi-automation in evidence production

Introducing Screen4Me

David Moher Cochrane Methods Symposium:
Bias and Beyond
Saturday 15 September 2018

Trusted evidence.
Informed decisions.
Better health.

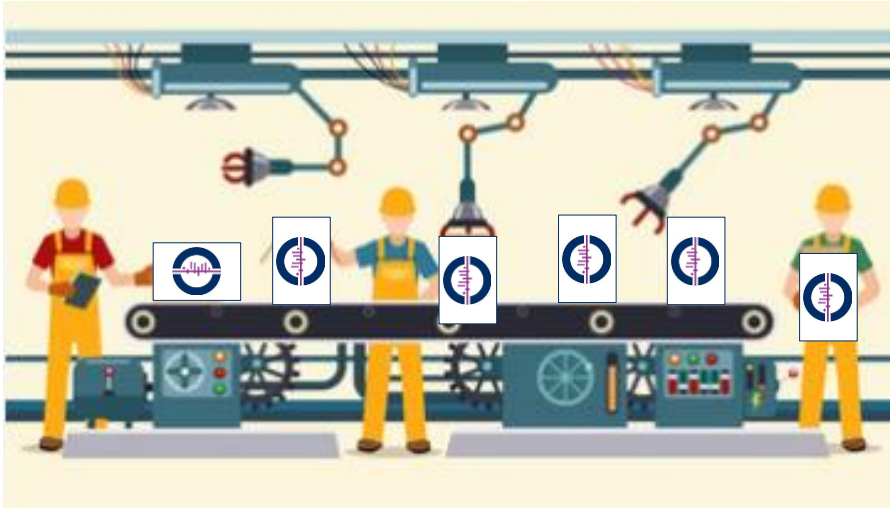
Semi-automation



First coined in manufacturing with the development of robotics in assembly lines.

Cochrane

Semi-automation



We are also using semi-automation in our review production line

Challenges

In Cochrane we are under pressure to produce **high quality** evidence **quickly**

Challenges

Information overload

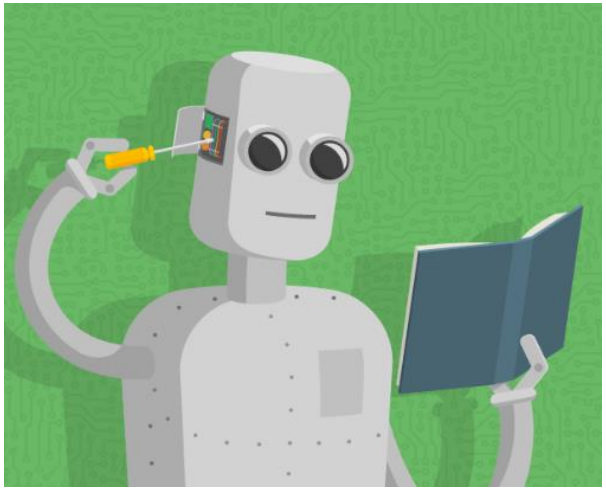
Global scientific output doubling every 9 years. 4000+ articles published every day

Challenges

Processes not working: too siloed leading to inefficiency and duplication of effort

New approaches

Crowdsourcing



Machine learning

New approaches like crowdsourcing and machine learning could help us make better use of data (metadata) and better use of people

Crowdsourcing

"Crowdsourcing is the process of obtaining needed services, ideas, or content by soliciting contributions from a large group of people, especially an online community, rather than from employees or suppliers."

Cochrane

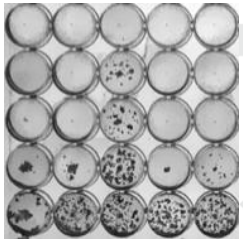
Crowdsourcing

There are different types of crowdsourcing and various names have been given to the different types.

Crowd Creation

Crowd Wisdom

Crowd Funding



Crowd Voting

Cochrane

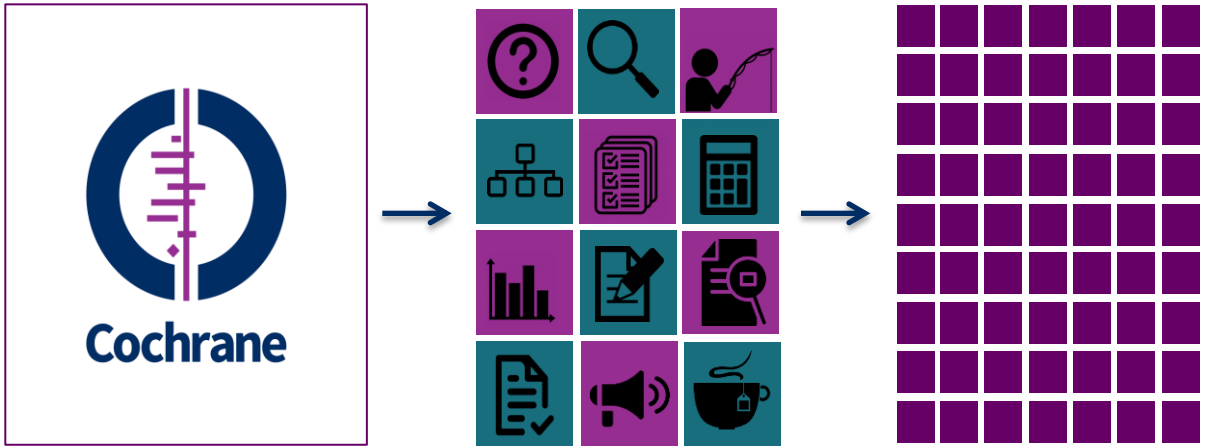
Cochrane Crowd

Cochrane Crowd: crowd,cochrane.org

Classification tasks about identifying and describing health research

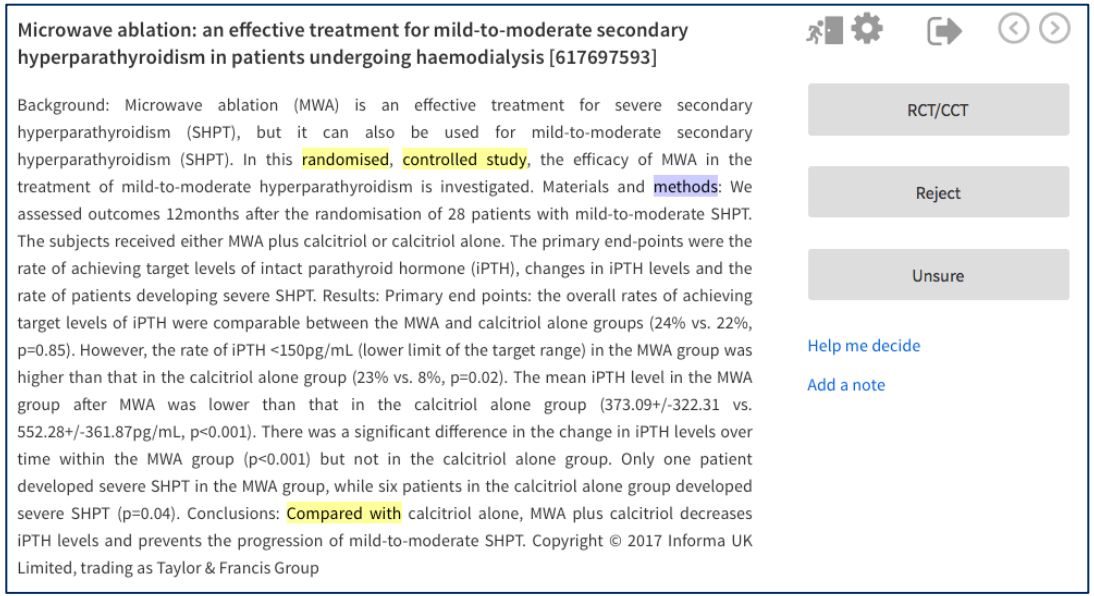
Cochrane

Cochrane Crowd



It breaks tasks down to a micro form

Cochrane Crowd



Breaking a task down to one single question: is it an RCT?

Cochrane Crowd

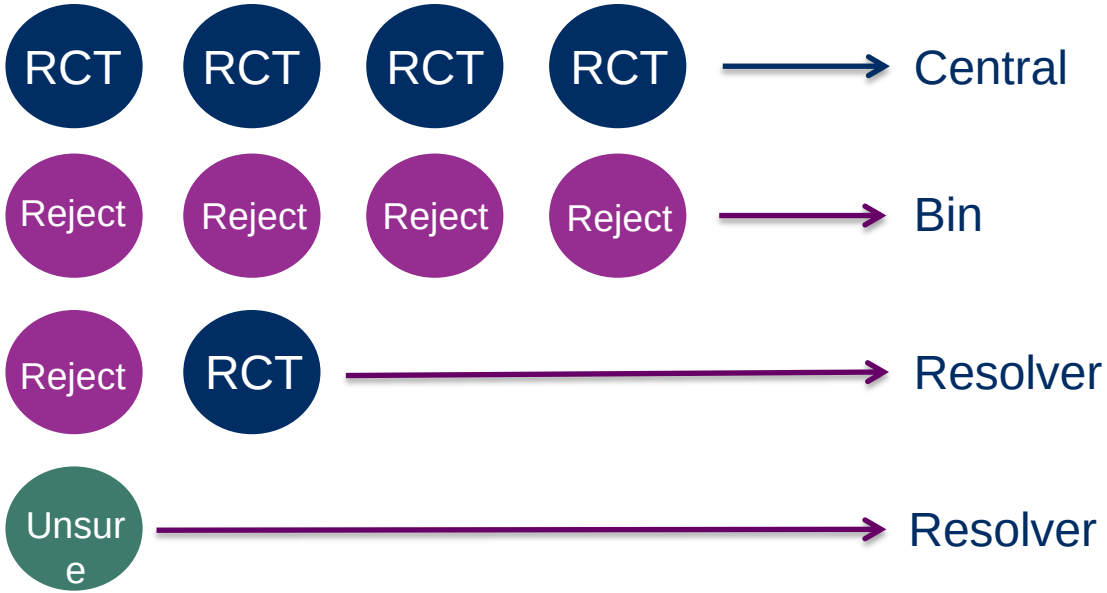
Brief, interactive training

Agreement algorithm

Individual accuracy is achieved through training

Collective accuracy is achieved through the agreement algorithm

Agreement algorithm



The agreement algorithm for RCT ID

Cochrane

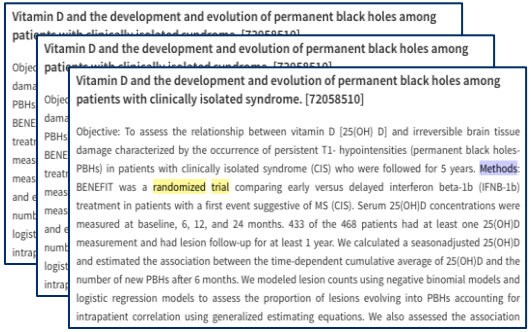
Cochrane Crowd

6 mainstream micro-tasks

10,000+ crowd

150+ countries

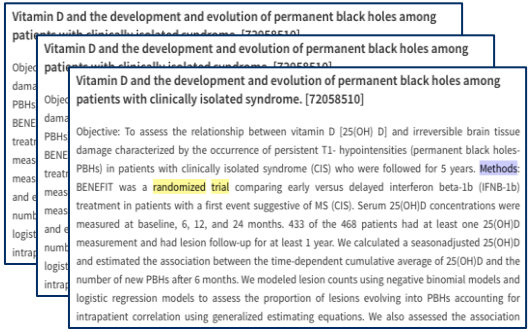
Cochrane Crowd: RCTs



42,706 records have been identified as RCTs by the crowd

The Crowd have identified 42,706 RCTs for CENTRAL

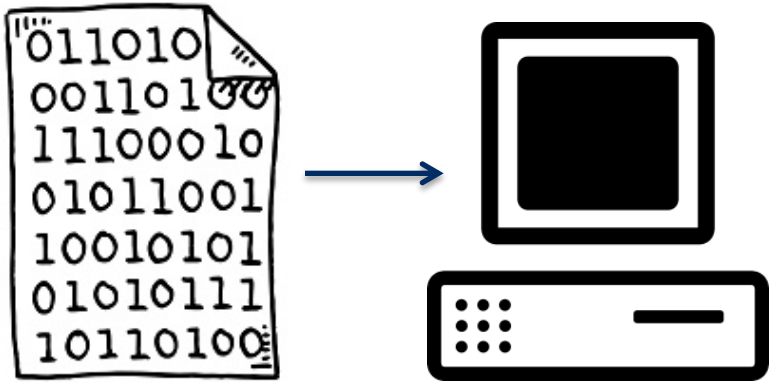
Cochrane Crowd: Rejects



464,294 records have been identified as NOT describing RCTs by the crowd

The Crowd have identified 464,294 REJECTs

Cochrane Crowd



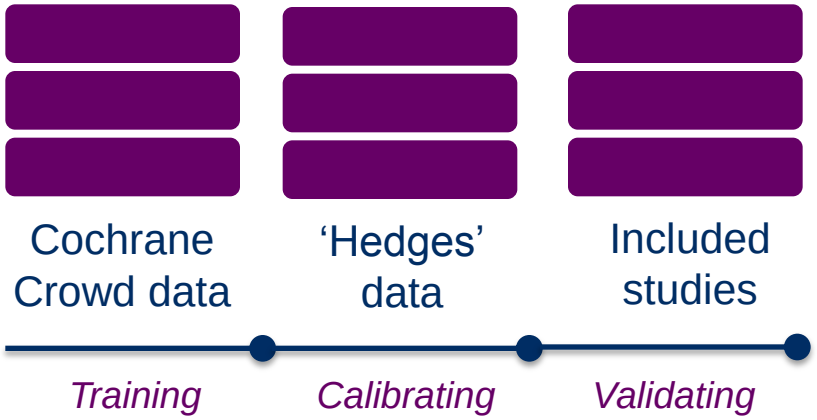
With the Crowd dataset we were able to train the machine

Cochrane

Machine learning

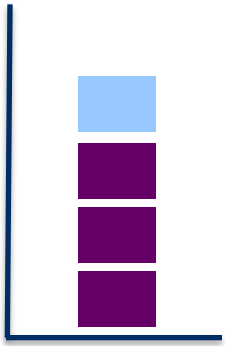
Machine learning gives “computers the ability to learn without being explicitly programmed”. In the context of Cochrane, this is about building classifiers that provide likelihood scores

Machine learning



The RCT classifier built, calibrated and validated

Perfect partnership

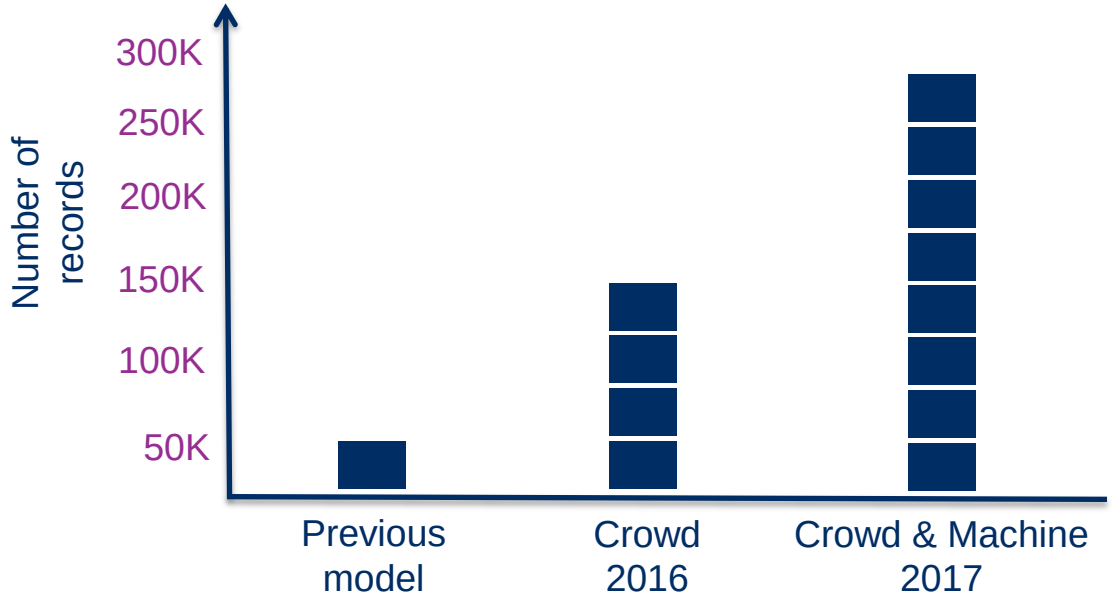


The classifier is been used in CSS processes.

Every month, the classifier reduces the amount of records that need to go to Cochrane Crowd by **around 75%**

Crowd and machine working together

Machine-Crowd in centralised searches



Crowd-machine system as part of the centralised search service has enabled significant scale up

Cochrane

Introducing Screen4Me

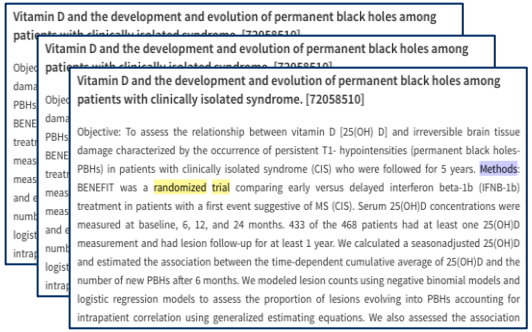
Known assessments

RCT classifier

Cochrane Crowd

Brining three components together and enabling review teams to access them for their specific reviews

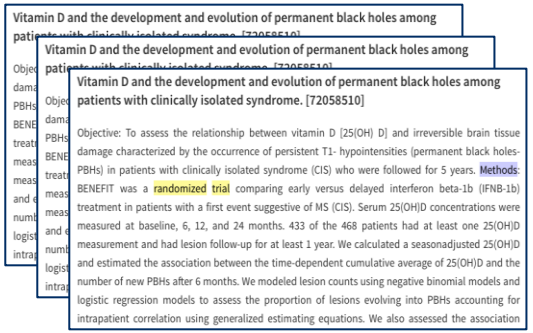
Known Assessments



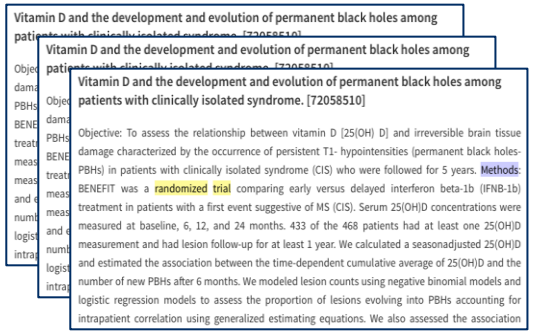
To date, over 500,000 bibliographic records have been through Cochrane Crowd

Cochrane

Known Assessments



42,000 RCTs

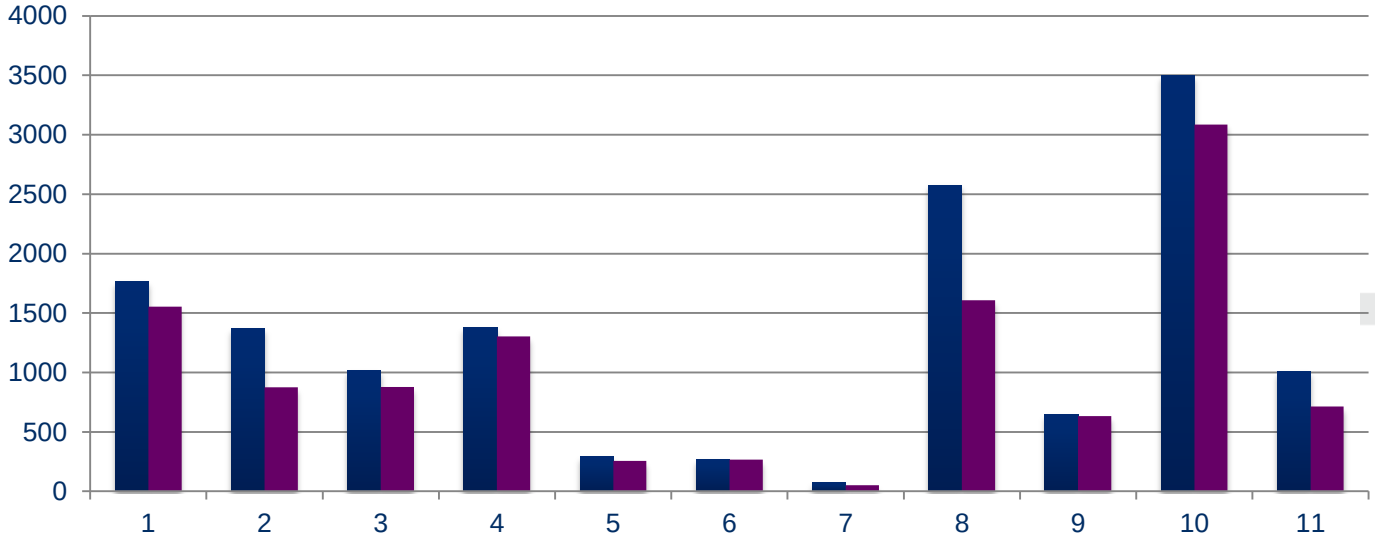


465,000 Rejects

We reject the same records time and again

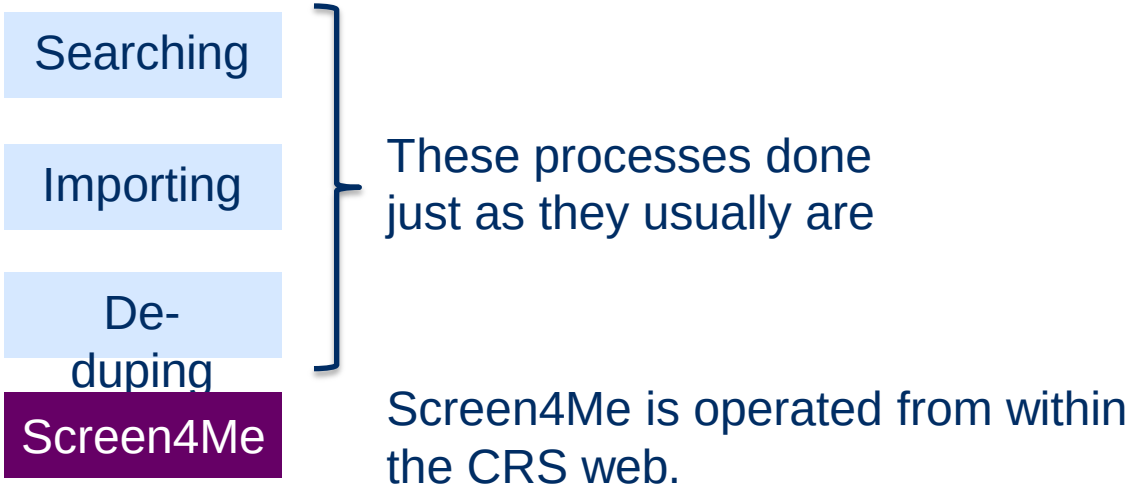
In S4M we will make use of the metadata we already have

Known Assessments

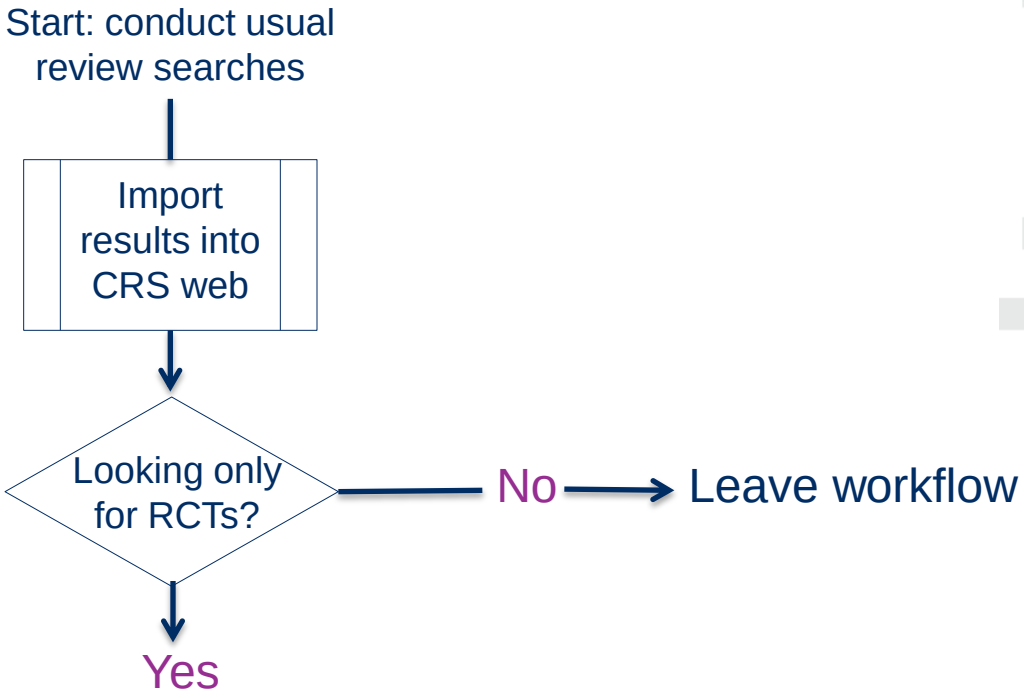


In a month's worth of updated reviews, the results identified in Embase searches for those updates had already been screened by the Crowd: 62%-98%

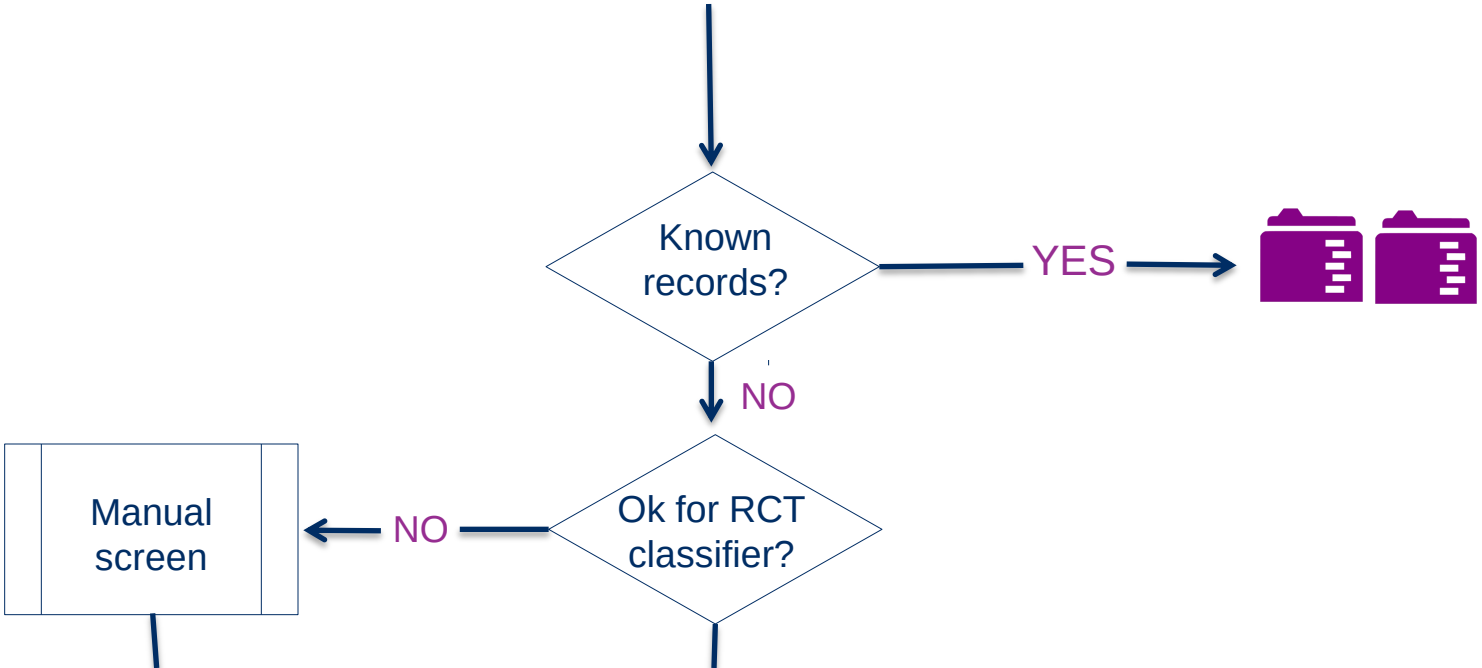
Screen4Me: how will it work?



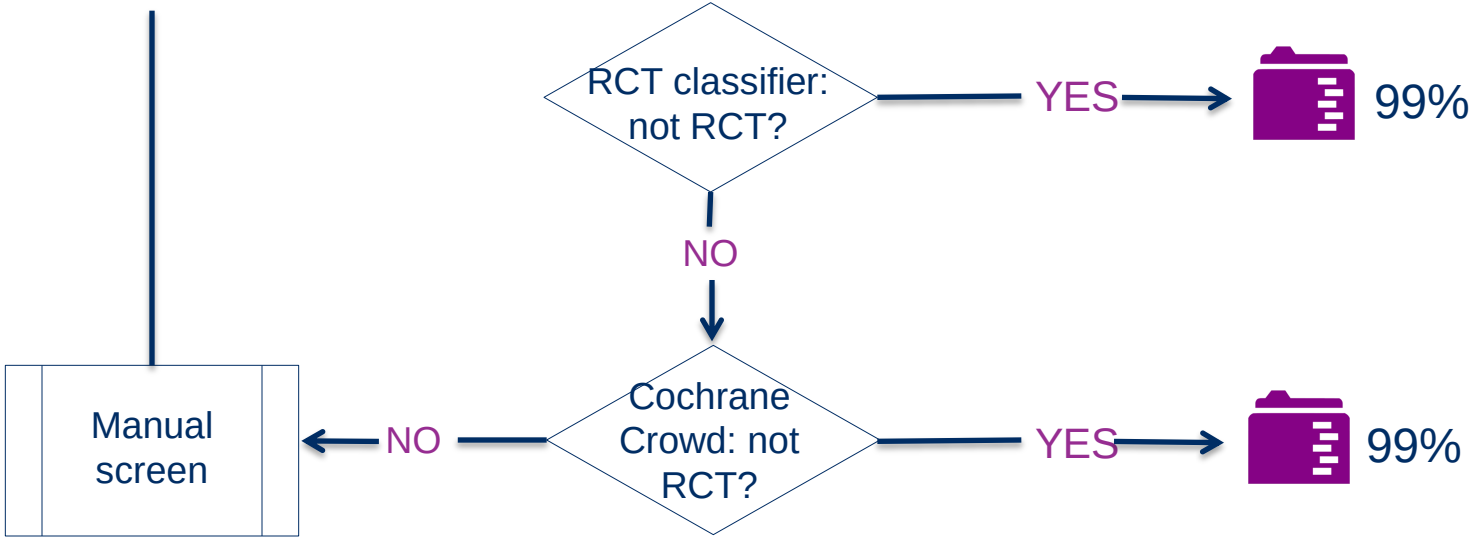
Screen4Me: workflow



Screen4Me: workflow *cont.*



Screen4Me: workflow *cont.*



Cochrane

Screen4Me

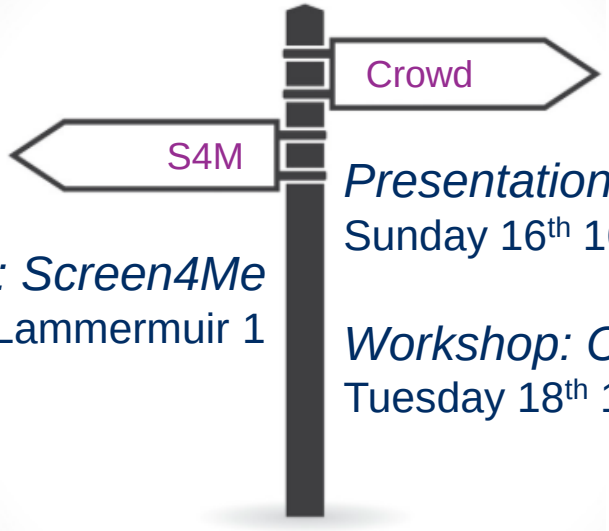
Estimated reduction in records for author teams to have to screen: 50-85%

Cochrane

In summary

- We have successfully incorporated semi-automation in the form of Crowd and machine into CSS processes
- This has enriched our central repositories
- Screen4Me is a new workflow available to review teams via CRS web

Want more detail?



Workshop: Screen4Me
Monday 17th 11.00, Lammermuir 1

Presentation: Cochrane Crowd
Sunday 16th 16.20, Carrick 1

Workshop: Cochrane Crowd
Tuesday 18th 11.00, Harris 2

Cochrane

Log-in to Cochrane Crowd crowd.cochrane.org anytime during the Colloquium to join the challenge!

Thank you

Anna Noel-Storr
anna.noel-storr@rdm.ox.ac.uk

With huge acknowledgements to the Project Transform team and to the Crowd

Support for Project Transform was provided by Cochrane and the National Health and Medical Research Council of Australia (APP1114605). The contents of the published material are solely the responsibility of the Administering Institution, a Participating Institution or individual authors and do not reflect the views of the NHMRC.