Dealing with a new kind of team: the crowd

EFSA Conference 2018, Parma, Italy
20 September 2018

Anna Noel-Storr
@annanoelstorr

Trusted evidence.
Informed decisions.
Better health.
Evidence synthesis

Question

Does X improve Y for condition Z?
Evidence synthesis

Does X improve Y for condition Z?

Question
Primary research
Evidence synthesis

Evidence synthesis comes along in the form of meta-analysis or systematic review and finds all the relevant studies to combine them in order to reach the best possible answer to the question.
The problem (well, one problem)

Turning information into knowledge and wisdom is challenging as the amount of information increases exponentially.
Another problem

Hello. Can I come in?

It can be challenging to enable people to contribute meaningfully.
We also struggle to retain contributors with meaningful ways to remain involved
So wait a minute…

We are struggling to keep up with the amount of information, yet we have people who want to help.
The solution

Cochrane Crowd: a citizen science platform that offers potential contributors ‘micro-tasks’
The ingredients

Three key elements: lots of **data that needs processing**, being able to create **doable tasks** to help process, and having a robust **agreement algorithm** to ensure collective accuracy.
Classifying or categorising

Vitamin D and the development and evolution of permanent black holes among patients with clinically isolated syndrome. [72058510]

Objective: To assess the relationship between vitamin D [25(OH) D] and irreversible brain tissue damage characterized by the occurrence of persistent T1-hypointensities (permanent black holes-PBHs) in patients with clinically isolated syndrome (CIS) who were followed for 5 years. Methods: BENEFIT was a randomized trial comparing early versus delayed interferon beta-1b (IFNB-1b) treatment in patients with a first event suggestive of MS (CIS). Serum 25(OH)D concentrations were measured at baseline, 6, 12, and 24 months. 433 of the 468 patients had at least one 25(OH)D measurement and had lesion follow-up for at least 1 year. We calculated a seasonadjusted 25(OH)D and estimated the association between the time-dependent cumulative average of 25(OH)D and the number of new PBHs after 6 months. We modeled lesion counts using negative binomial models and logistic regression models to assess the proportion of lesions evolving into PBHs accounting for intrapatient correlation using generalized estimating equations. We also assessed the association
You are not alone!

No record is just looked at once. Most records need 4 agreements for it to either be deemed an RCT or not.
Types of task

*Identify* health research

*Describe* health research

In Cochrane Crowd the micro-tasks are about *identifying* and *describing* health research.
Has it worked?

Cochrane Crowd (crowd.cochrane.org) was launched in May 2016
High level metrics

Cochrane Crowd (crowd.cochrane.org) was launched in May 2016
High level metrics

Cochrane Crowd (crowd.cochrane.org) was launched in May 2016
High level metrics

- 700,000+ records screened
- 5 mainstream tasks
- 2.1m classifications

Data as of 09 Sept 2018
High level metrics

- 10,000+ signed up
- 700,000+ records screened
- 2.1m classifications
- 5 mainstream tasks

Data as of 09 Sept 2018
The contributors

Contributors come from 180 countries (56% from lower and middle income countries)
The contributors

“I like that I can make a small contribution to Cochrane from my own home, on my own schedule”

“I hope to be able to contribute as much as I can in college and in my life after graduation, but I’d like to do something now. When I saw that I can work directly with Cochrane I was really excited”

“Feels like something useful you can do if you have just a few minutes and can’t get into something bigger workwise. It’s great being part of the team”

“It’s fun, and is actually useful.”
## Crowd accuracy

### RCT Identification

<table>
<thead>
<tr>
<th></th>
<th>Info specialist and methodologist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cochrane Crowd</td>
<td>TP 457, FP 58, FN 4, TN 5522</td>
</tr>
</tbody>
</table>

Sensitivity: **99.1%** Specificity: **99%**

### DTA Identification

<table>
<thead>
<tr>
<th></th>
<th>Info specialist and methodologist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cochrane Crowd</td>
<td>TP 100, FP 22, FN 20, TN 978</td>
</tr>
</tbody>
</table>

Sensitivity: **83%** Specificity: **97%**

### CTgov Identification

<table>
<thead>
<tr>
<th></th>
<th>Info specialist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cochrane Crowd</td>
<td>TP 8191, FP 77</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>FP 77</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cochrane Crowd</td>
<td>FN 17</td>
</tr>
<tr>
<td></td>
<td>TN 5823</td>
</tr>
</tbody>
</table>

Sensitivity: **99.7%** Specificity: **98.6%**

With the right agreement algorithm in place very high collective accuracy is possible.
The Crowd has identified 42,500 reports of randomised trials that had not been indexed as randomised trials. This has significantly enriched our central repository of trials.
The Crowd has also labelled 465,000 records that might have been RCTs as ‘not RCTs’
Machine learning

This dataset, generated by the Crowd has enabled us to train a machine to do the task (well, a significant proportion of the task)
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.
Increasing capacity

Crowd and machine working together enables scale-up
As the Crowd generate more data, it is fed to the machine who continues to learn.
In summary

Cochrane Crowd is helping us meet the challenge of information overload
In summary

Cochrane Crowd is helping us meet the challenge of information overload.

It provides potential contributors with meaningful ways to get involved.
Thank you
Anna Noel-Storr
anna.noel-storr@rdm.ox.ac.uk

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.