



## Overcoming Challenges to Ensure the Truthworthiness and Impact of Evidence Synthesis

Isabelle Boutron

Cochrane France Centre for Research in Epidemiology and Statistics (CRESS) Université Paris Cité











# Medical doctors, health care providers, public health decision makers are overwhelmed with the amount of new information produced

- 37 millions articles in the field of life science and medicine
- The amount of information doubles every 3 years
- About 900 000 clinical trials indexed in PubMed
- 35 000 new clinical trials published every year





If you were to print out just the first page of every item indexed in Web of Science, **the stack of paper would reach almost to the top of Mt Kilimanjaro** 

There is a crucial need for a reliable, independent, accessible and up-to-date synthesis of the knowledge produced by research

An accurate, concise and unbiased synthesis of the available evidence is arguably one of the most valuable contributions a research community can offer decision-makers (Donnelly, Nature 2021)

# Rofecoxib example 2004 drug withdrawal





Jüni P, et al. Lancet. 2004 <u>Boutron I, Créquit P</u>, (...), <u>Ravaud P</u>. **J Clin Epidemiol**. 2020 <u>Créquit P, Boutron I</u>, (...), Ravaud P. **J Clin Epidemiol**. 2020 <u>Ravaud P, Créquit P</u>, (...) <u>Boutron I</u>. **J Clin Epidemiol**. 2020

## **Evidence** synthesis methods to increase trustworthiness

- A strong process is implemented to improve untrustworthiness
  - Protocol, registration
  - Identification of all information related to primary studies (registry data, protocol, SAP, publication, Clinical study reports, etc)
  - Standardized extraction of all information

### **Evidence** synthesis

Definition of the research question



- **Specific tools** have been developed to ensure trustworthiness
  - Risk of bias tool (RoB 2)
  - Tool for addressing reporting bias (RoB-ME)
  - Tool for addressing conflicts of interest in trials (TACIT)
  - Tool to identify problematic trials
    - INSPECT-SR to identify problematic trials
    - Research Integrity Assessment (RIA) tool
    - Trustworthiness in Randomized Controlled Trials (TRACT) screening tool
    - The IPD Integrity tool

## **Evidence synthesis methods to increase trustworthiness**

• Sources of information: clinical trial registries

**PLOS** MEDICINE Timing and Completeness of Trial Results Posted at ClinicalTrials.gov and Published in Journals

Completeness of reporting	ClinicalTrials.gov N=202	Published article N=202	P-value
Flow of participants	64%	48%	< 0.001
Efficacy results	79%	69%	0.02
Adverse events	73%	45%	< 0.001
Serious adverse events	99%	63%	<0.001

### 202 published RCTs with posted results

## **Evidence synthesis methods to increase trustworthiness**

## **Other sources of information: Clinical study reports (CSR)**



- 42 trials (2015-18)
- Harms reporting was more complete in CSRs than other sources.
- Marked discrepancies in harms data between sources



Paludan-Müller AS, Créquit P, Boutron I. BMC Med. 2021

## **Transparency of evidence synthesis**

## **Reporting guidelines**

1999: QUOROM (QUality Of Reporting Of Meta-analyses) Statement

### 2009: PRISMA Statement

7 extensions to the PRISMA

- 2012 Equity
- 2013 Abstracts
- 2015 Network Meta-analyses
- 2015 Individual Participant Data
- 2015 Protocols
- 2016 Harms
- 2017 Complex interventions



## **Completeness of reporting improved but remained insufficient**

### **Transparency of evidence synthesis**

detail to allow users to assess the

trustworthiness and applicability of the

matthew.page@monash.edu

(ORCID 0000-0002-4242-7526)

Additional material is published

### **RESEARCH METHODS AND REPORTING**

OPEN ACCESS	PRISMA 2020 explanation and elaboration: updated guidance					
Check for updates	Matthew J Page, <sup>1</sup> David Moher, <sup>2</sup> Patrick M Bossuyt, <sup>3</sup> Isabelle Boutron, <sup>4</sup> Tammy C Hoffmann, <sup>5</sup> Cynthia D Mulrow, <sup>6</sup> Larissa Shamseer, <sup>7</sup> Jennifer M Tetzlaff, <sup>8</sup> Elie A Akl, <sup>9</sup> Sue E Brennan, <sup>1</sup> Roger Chou, <sup>10</sup> Julie Glanville, <sup>11</sup> Jeremy M Grimshaw, <sup>12</sup> Asbjørn Hróbjartsson, <sup>13</sup> Manoj M Lalu, <sup>14</sup> Tianjing Li, <sup>15</sup> Elizabeth W Loder, <sup>16</sup> Evan Mayo-Wilson, <sup>17</sup> Steve McDonald, <sup>1</sup> Luke A McGuinness, <sup>18</sup> Lesley A Stewart, <sup>19</sup> James Thomas, <sup>20</sup> Andrea C Tricco, <sup>21</sup> Vivian A Welch, <sup>22</sup> Penny Whiting, <sup>18</sup> Joanne E McKenzie <sup>1</sup>					
For numbered affiliations see end of the article. Correspondence to: M Page	The methods and results of systematic reviews should be reported in sufficient decisions. To allow decision makers to assess the					

### **PRISMATIC** Project

trustworthiness and applicability of review findings,

reports of systematic reviews should be transparent



Fig. 1 Model of the PRISMA Web app. In the example depicted, a customised checklist and template for a systematic review with network meta-analysis of adverse events associated with COVID-19 vaccines are created. The items are drawn from three statements—the unified PRISMA statement; PRISMA Harms extension [15] and PRISMA NMA extension [7]. Adapted from Hopewell et al. [28]

### Accuracy and spin in evidence synthesis

Spin is defined as specific intentional or unintentional reporting that fails to faithfully reflect the findings and could affect the impression the results produce in readers

Distortion of results interpretation

25% SR reported with spin



```
- Januari of Funnois Symmetry and Pre20043 About
```

### A new classification of spin in systematic reviews and meta-analyses was developed and ranked according to the severity Amélie Yavchitz<sup>a,har,\*</sup>, Philippe Ravaud<sup>a,har,al</sup>, Douglas G. Altman<sup>\*</sup>, David Moher<sup>1,\*</sup>,

Journal of Clinical Epidemiology

Asbjørn Hrobjartsson<sup>16</sup>, Toby Lasserson<sup>1</sup>, Isabelle Boutron<sup>2,0,6</sup>

"Umme de Recherche Epidlimologie et Statistique Sorbonue Paris Chië (CRESS-UMR1155), Internallimorenté Paris Demartes, 1 place du Paris Forr Danne, Paris 73004, France

<sup>6</sup>Comme d'Épidémiologie Clinique, AP-HP (Assistance Publique des Hopitaux du Paris), Hopital Hirri Dina, I place du Parisi Notre Dume, Faris, France <sup>6</sup>French Cuchenne Conten Paris Desearres (lanvenity, Sarboune Paris Cité, Faculté du Mederine, I place du Parisi Notre Dume, Faris 75000, France <sup>6</sup>Department of Epidemiology, Columbia University Malimum Richard Helder Budie, New York, NY, USA <sup>7</sup>Centre for Nonview in Medicum, Nullevil Department of Orthoposodics, Resoundatory, and Marculask-etail Sciencer, University of Ostant, USport, Chi-<sup>7</sup>Centre for Nonview in Medicum, Nullevil Department of Orthoposodics, Resoundatory, and Marculask-etail Science, University of Ostant, USport, Chi-<sup>7</sup>Centre for Nonview, Ottawa, Canum.

<sup>3</sup>Schnel of Epidemiology, Public lookli and Preventive Medicine, Faculty of Medicine, University of Ottava, Canada <sup>8</sup>Ninetic, Cachrane Contre, Right-optimient Department 7811, Copyentagen, Denmath Research Unit for Evidence-Based Medicine, Odence University Haspital/Durversity of Southers Dominac, Odence, Denmatk (Cochrane Editorial Univ, Lodence, UK)

Accepted 19 January 2016: Published online 2 February 2016

#### Abstract

Objectives: We aimed to (1) identify and classify spin (i.e., a description that overstates efficacy and/or understates harm) in systematic reviews and (2) rank spin in abstracts of systematic reviews according to their severity (i.e., the likelihood of distorting readers) interpre-

META-RESEARCH ARTICLE

# 'Spin' in published biomedical literature: A methodological systematic review

### Kellia Chiu, Quinn Grundy, Lisa Bero\*

Charles Perkins Centre, Faculty of Pharmacy, The University of Sydney, Sydney, New South Wales, Australia

\* lisa.bero@sydney.edu.au

# Spin disseminates from the publication to the press release and the news



Yavchitz A, Boutron I, Bafeta A, Marroun I, Charles P, Mantz J, Ravaud P. PLoS Med. 2012

10

### **Spin impacts readers' interpretation**



**Based on this abstract, do you think treatment A would be beneficial to patients?** Scale, 0 [very unlikely] to 10 [very likely)

Mean difference = 0.71 (95% CI, 0.07-1.35); P .03



"What do you think is the probability that 'treatment X' would be beneficial to patients?" (scale, 0 [very unlikely] to 10 [very likely])

Boutron I, Altman DG, Hopewell S, Vera-Badillo F, Tannock I, Ravaud P. J Clin Oncol. 2014 Boutron I, Haneef R, Yavchitz A, Baron G, Novack J, Oransky I, Schwitzer G, Ravaud P. BMC Med. 2019

Rapid growth in the publication of systematic reviews



THE MILBANKQUARTERLY A MULTIDISCIPLINARY JOURNAL OF POPULATION HEALTH AND HEALTH POLICY

**Original Investigation** 

### The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses

JOHN P.A. IOANNIDIS 🔀



The problems with systematic reviews: a living systematic review Lesley Uttley<sup>a,\*</sup>, Daniel S. Quintana<sup>b,e,d</sup>, Paul Montgomery<sup>e</sup>, Christopher Carroll<sup>a</sup>, Matthew J. Page<sup>f</sup>, Louise Falzon<sup>a</sup>, Anthea Sutton<sup>a</sup>, David Moher<sup>g</sup>



2023

Journal of Clinical

Epidemiology



### Systematic reviews are rarely updated



#### Distribution of Duration between 2 versions

Each dot represents one duration between 2 versions; boxplot shows the distribution.



## **Evidence syntheses are incomplete and fragmented**

• What proportion of randomized evidence is included in systematic reviews. The example of lung cancer?





- From 2009 to 2015 the evidence covered by existing systematic reviews was consistently incomplete and did not consider
  - 45 % to 70 % of trials;
  - 30 % to 58 % of patients;
  - 40 % to 66 % of treatments;
  - 38 % to 71 % of comparisons

<u>Créquit,..,Ravaud</u> **BMC Med**. 2017 <u>Créquit,..,Ravaud</u> **BMC Med**. 2016 <u>Crequit..,Ravaud</u> **BMJ Open** 2016

## From meta-analysis to living meta-analysis to living network meta-analysis



### **Decision makers need** 'living' evidence synthesis

Julian H. Elliott, Rebecca Lawrence, Jan C. Minx, Olufemi T. Oladapo, Philippe Ravaud, Britta Tendal Jeppesen, James Thomas, Tari Turner, Per Olav Vandvik & Jeremy M. Grimshaw

Fund and use dynamic evidence summaries of the latest data to steer research, practice and policy.

Council, worried that the cacophony would create confusion and anxiety among alreadystressed clinicians. We argued for key bodies

the evidence pipeline'<sup>1</sup>. Take the example of remdesivir, an intravenous treatment originally developed for Ebola virus. In May 2020, weak to come together quickly and use robust, but promising data suggested it could be used evidence-based processes to find signals in the to treat COVID-19. Over the next 18 months,

### EVIDENCE ACCELERATED

Using a living-evidence approach, researchers find, appraise and incorporate research in frequent cycles, rather than always starting from scratch.

#### Stroke

The Australian Stroke Foundation reduced the time between guideline updates from 7 years to under 3 months.





Chaimani A, Porcher R, Sbidian E, Mavridis D. **Stat in Med** 2021 Boutron,..,Ravaud **J Clin Epidemiol**. 2020 Ravaud..,Boutron **J Clin Epidemiol**. 2020 **Boutron**, Chaimani..., Ravaud. **Ann Internal Med**, 2020 Oikonomidi, Boutron, ..., Ravaud, **BMC Med**, 2020

## The ecosystem is evolving rapidly



# Can we trust results published as preprint?

Comparison	Outcome	Preprint n(N)	Journal n(N)		Weights	ROR [95% CI]
Sofosbuvir-ledipasvir vs. control [Hosp]†24	ACM	1 (250)	1 (82)	A	0,41%	0.08 (0.00, 2.16
Intravenous immunoglobulin vs. control [Hosp]2#	ACM	1 (33)	1 (84)	<	0.70%	0.28 [0.02, 3.57
Hydraxychioroquine vs. control (OP)+	AE	1 (105)	3 (1203)		0.35%	0.31 [0.01, 11.70
ivermectin vs. control (Hosp)	ACM	3 (365)	2 (233)	• • • • • • • • • • • • • • • • • • •	0.95%	0.33 (0.04. 2.84
REGN-CoV2 vs. control [OP]*12	SAE	1 (311)	1 (209)	< · · · · · · · · · · · · · · · · · · ·	0.32%	0.42 [0.01, 18.67
Colchicine vs. control [OP]12#	ACM	1 (276)	1 (4488)		0.40%	0.46 (0.02, 13.52
Hydroxychloroquine vs. control [Hosp]	ACM	1 (247)	7 (7871)	to a second seco	3.57%	0.46 (0.16. 1.31
Anakinra vs. control [Hosp]†2*	WHO?	1 (594)	1(114)	i and	3.95%	0.47 (0.18. 1.23
Ciclesonide vs. control [OP]*12+	AE	1 (400)	1 (203)	the second second	4.23%	0.49 (0.19, 1.23
Convalescent plasma vs. control (OP)11	ACM	1 (782)	2 (1679)	4	0.13%	0.51 [0.00, 202.02
Baricitinib vs. control (Hosp]*114	Cim	1 (101)	1 (1512)	h	4.83%	0.51 (0.22, 1.21
Favipiravir vs. control (OP)#1	AE	1 (119)	2 (380)		3.78%	0.56 [0.21, 1.52
REGN-CoV2 vs. control [Hoso]*11*	ACM	1 (1233)	1 (9875)	H	11.24%	0.63 (0.42, 0.96
Molnupiravit vs. control (OP)***	AE	1 (117)	2 (1559)	the second se	4 58%	0.67 10.28. 1.63
Interferon beta vs. control [Hosp]	ACM	1 (60)	3 (4247)		0.99%	0.73 (0.09. 6.03
Ivermectin vs. control (OP)	AE	1 (116)	4 (973)	Inclusion in the second second	1.27%	0.77 10.12. 4.91
Nitazoxanide vs. control IOPI*14	AE	1 (835)	1 (392)	the second se	8.42%	0771044 138
Topligumablys, control [Hoso]	ACM	2 (390)	9 (6482)		8.44%	07910.45 1.39
Auxora vs. control Prospi*te	ACM	1 (201)	1 (26)		0.85%	0.64 (0.08. 8.32
Favoravir vs. control Diosol***	AF	1 (163)	2 (254)		0.39%	0.85 10.03, 28.04
Saniumab vs. control (Hosp)*11*	BAE	2 (1787)	2 (866)		7.92%	0.92 10.51 1.66
Pasisterteros inmistis un control (CO)***	SAF	1 (120)	1.(60)		0.30%	1.00/0.03 31.50
Convolenced plasma us control biosolt	ACM	5 (631)	5 (101)		1 5656	1 14 10 22 6 14
Loninguisethoragie un control Bidenti*	4/34	1 (293)	3 (8010)		5.16%	1 19 10 53 2 69
Remaining on control Datasetter	ACM	1/54511	3 (1882)		12.50%	130/030 187
Campatat mediate vs. control (Hospi'tts	SAF	1 (153)	1 (205)		0.42%	170 00.05 47 03
Locinetic densities an enter (ODI+t)	RAE	1 (120)	1 (462)	Solution and solutions of the	0.42%	1.87 0 07 50.81
Nitesperante us control lideal**	ACM	1 (54)	1 (50)	Contraction of the second second	0.41%	1 66 10 07 55 50
Anthromotic us, control (CDI++	140407	1 (292)	1 (5121)	and the second	0.695	2.02 0 15 26 36
Colositivos dariatesuirus postrol Islassila	ACM	2 (1172)	2 (220)		5.85%	2.05.00.07 4.33
Asthomatic us costrol liberaliti	ACM	1 (2264)	5 (515)		0.44%	3 10 10 12 78 03
Minacimenta alabela in control Decada	101	1 (1104)	1 (50)		0.45%	4 34 10 15 103 75
highmentation general fidewal*th	4014	1 (42)	1 (107)		0.45%	4.34 (0.10, 103.71
LIC-LICC us control (Lines)+	AE	1 (100)	1 (94)	and the second s	0.07%	4 85 10 30 60 97
Control of a control (house)	4044	1 (100)	1 (24)		4.700	4,00 [0,00, 00,07
Decisione vs. comos (Hosp)	PALE	7(11340)	3 (200)	and the second se	1.70%	5.04 [1.09, 23.00
Vitamin D vs. control (Hosp)**	ACM	1 (240)	1 (76)	the second se	0.43%	13.99 [0.54, 365.09
Overall (Q = 38.49, df = 36, p = 0.38; $r^2 = 17.8\%, r^2 = 0.06$ )				•	100%	0.85 [0.71, 1.09
				· · · · · · · · · · · · · · · · · · ·		

### Preprint are living documents -> Iterative assessment

Holly Else, Nature 2020 Davidson et al JCE 2023 Davidson et al BMC Med Res Methodol. 2024

# **Paper mills and predatory journals**

### 53,000 papers analysed (6 publishers)



## Avoiding fake journals and judging the work in real ones



NEWS | 11 March 2020

### Hundreds of scientists have peerreviewed for predatory journals

Many of these titles have some editorial oversight – but the quality of reviews is in question.

By Richard Van Noorden

nature

Paper Mills – Research report from COPE & STM 2022

## The role of the peer-review process is questioned

- Costly and not sustainable
  - Estimation that 63,4 million hours devoted to peer-review in 2015<sup>1</sup>
  - US\$ 6 billion in  $2020^3$
- Delay access to results
  - COVID-19 pandemic: a 1-month delay in the results on the effect of steroids (RECOVERY trial) could cost about 200 lives
- Subject to biases
  - Peer-reviewers randomized to assess a study report differing only in the direction of the findings for the primary outcome
    - were **more likely to recommend the positive** version of the manuscript for publication (97% vs 80%, P<0.001)
    - detected **more errors** and **lower methods scores** in the no-difference version than in the positive version

<sup>1</sup> Kovanis et al. Plos One 2016
<sup>2</sup> Knowlson, Torgerson F1000 2020
<sup>3</sup> LeBlanc et al. Res Integr Peer Rev 2023

# The role of the peer-review process is questioned

### **Impact on completeness of reporting (adherence to CONSORT)**



### **Impact on spin (distorted reporting and interpretation of results)**



## **Generative AI is transforming research practices**



A world of AI-assisted writing and reviewing might transform the nature of the scientific paper. **By Gemma Conroy** 

234 | Nature | Vol 622 | 12 October 2023

### Is ChatGPT making scientists hyperproductive? The highs and lows of using AI

Large language models are transforming scientific writing and publishing. But the productivity boost that these tools bring could have a downside.

ChatGPT continues to steal the spotlight, more than a year after its public debut.

The artificial intelligence (AI) chatbot was released as a free-to-use tool in November 2022 by technology company OpenAI in San Francisco, California. Two months later, ChatGPT had already been listed as an author on a handful of research papers. use of LLMs in research writing. In a 2023 Nature survey of more than 1,600 scientists, almost 30% said that they had used generative AI tools to help write manuscripts, and about 15% said they had used them to help write grant applications.

And LLMs have many other uses. They can help scientists to write code, brainstorm research ideas and conduct literature reviews. LLMs from other developers are improving, as well – Google's Gemini, for example, and Claude 2 by Anthropic, an AI company in San Francisco. Researchers with the right skills



Artificial intelligence software is increasingly involved in reviewing papers – provoking interest and unease. By Miryam Naddaf

852 | Nature | Vol 639 | 27 March 2025

### Nature's survey of 1,600 respondents on artificial intelligence (AI)

### **HOW RESEARCHERS USE** LARGE LANGUAGE MODELS

Q: What do you use generative AI tools (such as ChatGPT and other large language models) for? (Choose all that apply.)

For creative fun not related to my research

To help write code

To brainstorm research ideas

To help write research manuscripts

To help do research

To conduct literature reviews

Within scientific search engines

To help fill out work-related administrative e-mails

To help write presentations

To help write grant applications

To help review research manuscripts

To help create graphics or pictures

To help write coursework or exam questions

Other

0





**AI ANTICIPATIONS** 

Q: How useful do you think AI tools

are for researchers in your field?

Respondents who use AI in research

Respondents who don't use AI in research

Ó

## INTERNATIONAL COMMITTEE of MEDICAL JOURNAL EDITORS

### 4. Artificial Intelligence (AI)-Assisted Technology

At submission, the journal should require authors to disclose whether they used Artificial Intelligence (AI)assisted technologies (such as Large Language Models [LLMs], chatbots, or image creators) in the production of submitted work. Authors who use such technology nitshe  $\frac{1}{1}$  (...) the journal should require ow the sisauthors to **disclose whether they** tar see Se used Artificial Intelligence ((...) sis, e in  $th_{th_{\epsilon}}^{or}$  in the production of submitted as <sup>Ch</sup> work. ney nd car

originality of the work, and these responsibilities are required for authorship (see Section II.A.1). Therefore, humans are responsible for any submitted material that included the use of AI-assisted technologies. Authors should carefully review and edit the result because AI can generate authoritative-sounding output that can be incorrect, incomplete, or biased. Authors should not list AI and AI-assisted technologies as an author or co-author, nor cite AI as an author. Authors should be able to assert that there is no plagiarism in their paper, including in text and images produced by the AI. Humans must ensure there is appropriate attribution of all quoted material, including full citations.

### 3. Peer Reviewers

(...) Reviewers must maintain the confidentiality of the manuscript as outlined above, which may prohibit the uploading of the manuscript to software or other AI technologies (...)

### Reviewers must **request permission from the journal prior to using AI technology to facilitate their review.**

Reviewers should be **aware** that AI can generate authoritative-sounding output that **can be incorrect, incomplete, or biased**.

### MORE THAN 10,000 RESEARCH PAPERS WERE RETRACTED IN 2023 — A NEW RECORD

The number of articles being retracted rose sharply this year. Integrity experts say that this is only the tip of the iceberg.

Richard Van Noorden Nature 2023

## **BIOMED RETRACTIONS HAVE QUADRUPLED IN 20 YEARS — WHY?**

Data falsification and other forms of misconduct are driving a growing proportion of retractions.

### Holly Else, Nature 2024



# **Development of new tools to detect problematic articles**



## Problematic Paper Screener Est. February 27<sup>th</sup>, 2021

- Automated tools are being developed to detect AI generated articles (Guillaume Cabanac, Cyril Labbé)
- Relies on a network of researchers ("research integrity sleuth") checking the papers and commenting on PubPeer and on social media

## ➢ Hundreds of retraction

Springer Nature retracted >300 articles

# **Development of new tools to detect studies citating retracted studies**

**Goal**: Stop bad research propagating through the scientific literature

### Feet of Clay Detector

(more on the origin of the metaphor) G Cabanac, University of Toulouse

Search by journal allows identification of articles citing retracted studies

=> Detection of retracted studies included in meta-analysis

## RetractoBot

Because researchers need to know about retracted papers N DeVito, B Goldacre University of Oxford

More than 100,000 researchers infomed by email they cite retracted studies A randomized trial is ongoing

Grana, ... Boutron Jama Intern Med March 31 2025

# AI is also an opportunity for systematic reviews

## **Annals of Internal Medicine**<sup>®</sup>

Research and Reporting Methods | 21 May 2024

Sensitivity and Specificity of Using GPT-3.5 Turbo Models for Title and Abstract Screening in Systematic Reviews and Metaanalyses

Authors: Viet-Thi Tran, MD, PhD S, Gerald Gartlehner, MD, MPH S, Sally Yaacoub, PhD S, Isabelle Boutron, MD, PhD S, Lukas Schwingshackl, PhD, MSc, Julia Stadelmaier, MSc S, Isolde Sommer, PhD S, Farzaneh Alebouyeh, MSc S, Sivem Afach, PhD S Joerg Meerpohl, MD, PhD S, and Philippe Ravaud, MD, PhD R AUTHOR, ARTICLE, & DISCLOSURE INFORMATION



# **Evidence synthesis**

« The ultimate goal of systematic reviews and meta-analysis is to create an effective marketplace for synthesis in which policymakers (..)always seek the best evidence because they know it will be available, and researchers synthesize evidence because they know it will make a difference. » (C A. Donnelly, Nature 2018)

The current evidence synthesis ecosystem does not fulfil this goal.

Mass production of systematic reviews and meta-analyses

- Low quality
- Redundant
- Not covering all evidence
- Delay in producing the review (2-3 years)
- Rarely updated
- Rely on primary evidence of low quality

# **Opportunities and challenges**

- Access to new source of data
  - Preprint, clinical trial registries, protocols, and clinical study reports from regulatory agencies or pharmaceutical companies
- Access to new types of data
  - IPD
  - Non randomized studies of routinely collected data (rich data, new design such as emulated trials, new statistical methods)
- New technology
  - AI tools, Large language models

## **Rethinking the role of evidence synthesis**

### Toward a new research ecosystem relying on a culture of continuous improvement

A living community responsible for one condition Systematic reviewers, Researchers, Clinicians, Patients, Trialists, Methodologists, Statisticians, Guidelines experts Health policy-makers Health systems managers



Collaboration with IT experts to develop the relevant tools to accelerate the process

Ravaud et al. JCE 2020

- Trustworthy and up-to-date evidence synthesis is essential to support decision making
- The landscape is changing rapidly with new challenges and opportunities arising
- Novel approaches are needed to fulfil stakeholders' needs