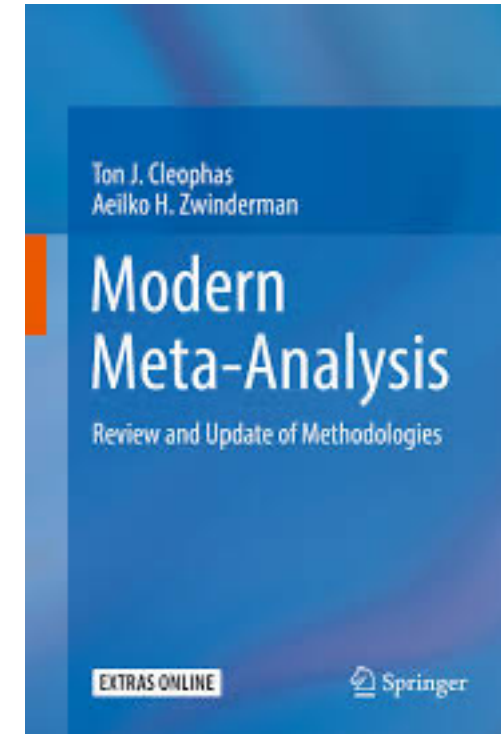
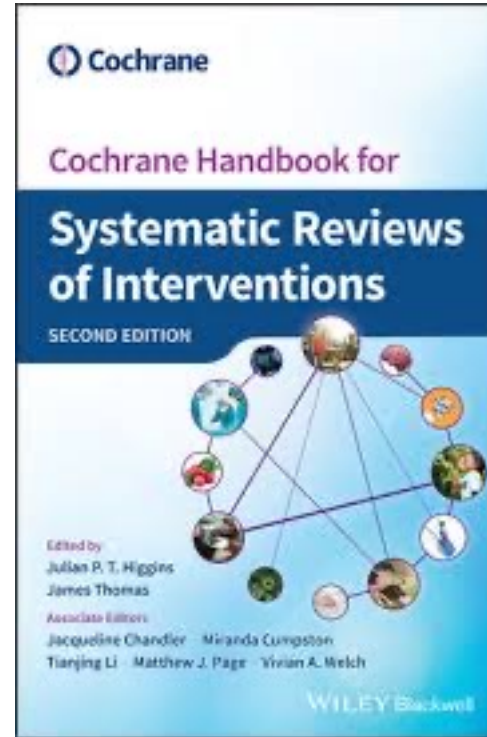
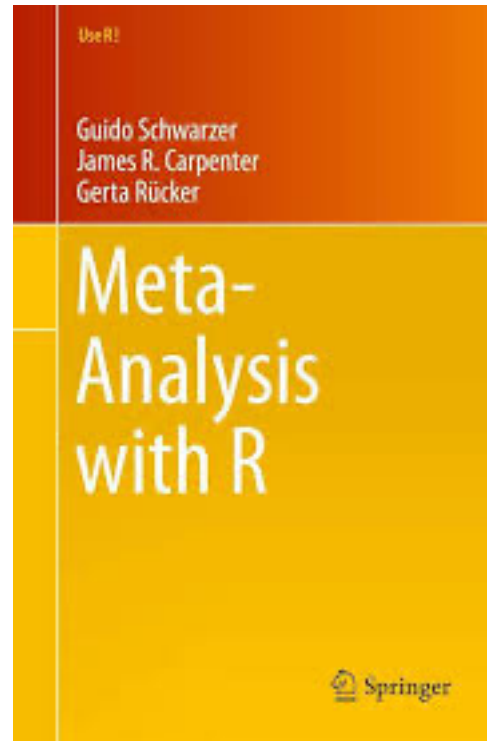
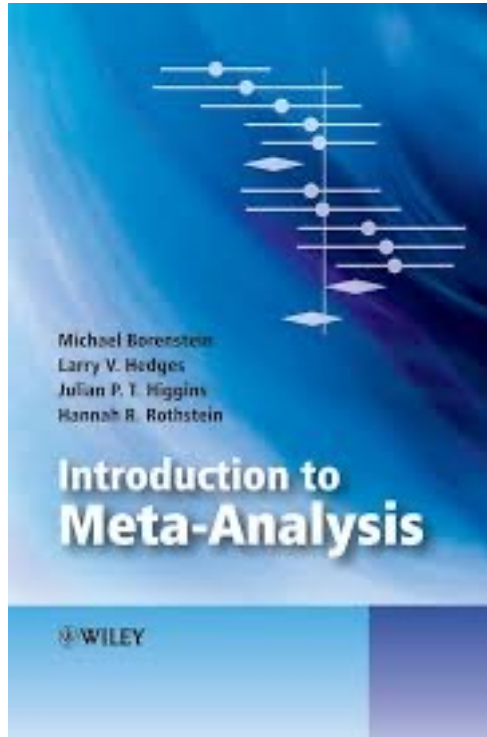




Les différents types de méta-analyses

Charles Khouri
Centre régional de pharmacovigilance
Centre d'investigation clinique
CHU Grenoble Alpes

Qu'est-ce qu'un type de méta-analyse?



Qu'est-ce qu'un type de méta-analyse?



ChatGPT

Il existe plusieurs types de méta-analyses, chacune avec ses propres caractéristiques et applications. Voici les principaux types de méta-analyses

- 1. Méta-analyse classique (ou traditionnelle)**
- 2. Méta-analyse en réseau**
- 3. Méta-analyse de sous-groupes**
- 4. Méta-analyse cumulative**
- 5. Méta-analyse de la courbe de dose-réponse**
- 6. Méta-analyse bayésienne**
- 7. Méta-analyse de la prévalence**
- 8. Méta-analyse de la qualité de vie**
- 9. Méta-analyse qualitative**
- 10. Méta-analyse de la variance de traitement**

Chacune de ces méta-analyses a des applications spécifiques et peut être utilisée pour répondre à différentes questions de recherche en fonction du contexte et des données disponibles.

Qu'est-ce qu'un type de méta-analyse?

- **Finalité** synthèse effet intervention, analyse de l'hétérogénéité, quantification biais...
- **Type d'études:** RCT observationnelles, prévalence, test diagnostic, modèles de prédiction... **2592 types de méta-analyses....**
- **Type de données:** individuelles, agrégées, **analyses !!**
- **Modèle statistique:** MA classique fixe/aléatoire, multivariée, multiniveau, en réseau, en composante...
- **Inférence** bayésienne vs fréquentiste
- **Mais aussi** MA statique, prospective, vivante....

La méta-analyse en réseau bayésienne multivariée d'études observationnelles dose-réponse

Intérêt



THE LANCET

Limites



Plan

MA pairwise/classique

MA en réseau

MA sous groupe/metaregressions

MA sur données individuelles

Les méta-analyses inutiles

Perspectives (MA vivantes)

Finalités

Type de données et d'études

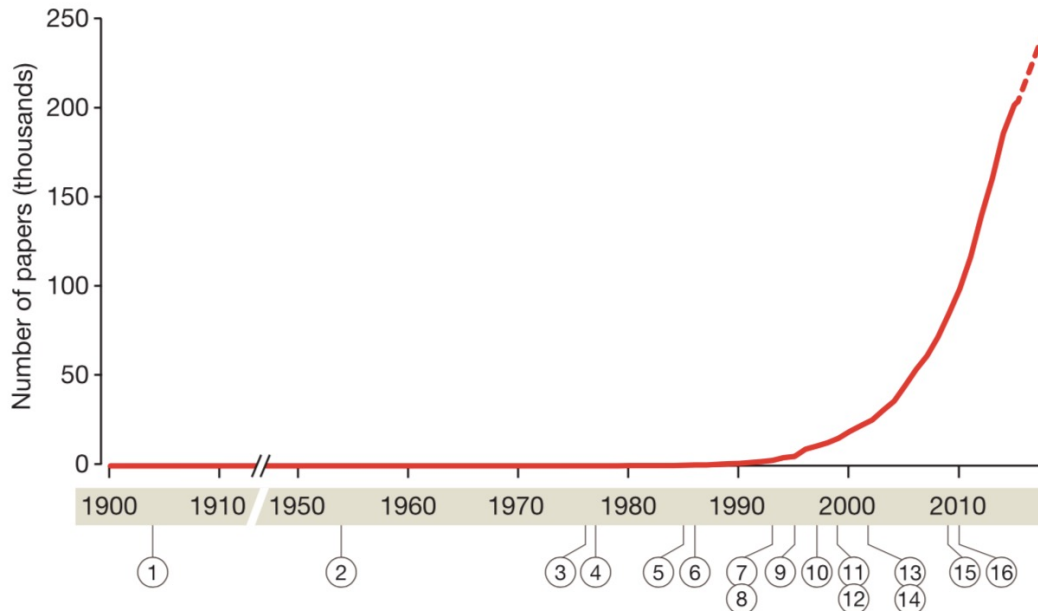
Méthodes

Limites

Une méta-analyse, pourquoi faire ?

1/ Estimer l'effet central (moyen) d'une intervention et sa variabilité

- Interprétation quantitative globale des informations disponibles

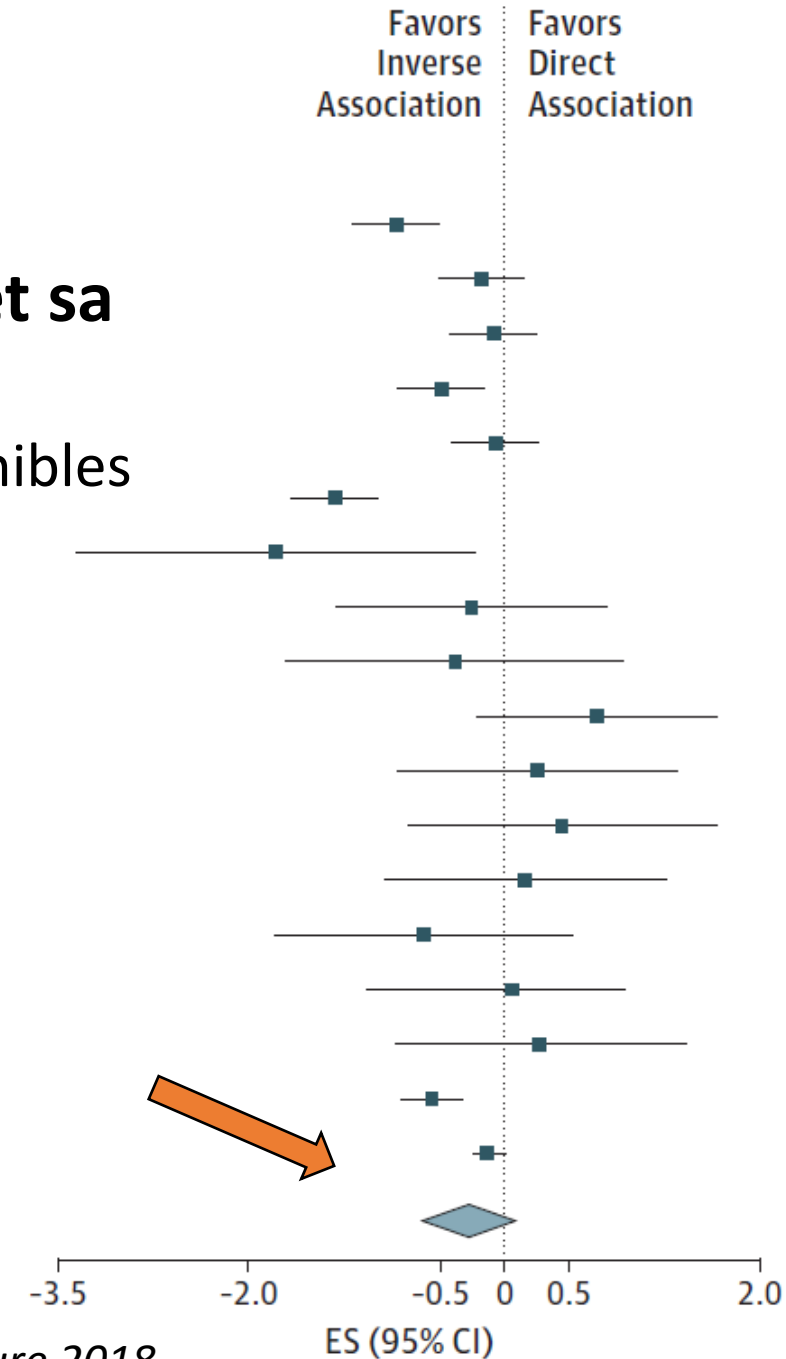


① **1904** First (medical) meta-analysis published (effect of inoculation against typhoid)

② **1954** First meta-analytic methods formalized (fixed- and random-effects models)

④ **1977** First social science meta-analysis published (efficacy of psychotherapy) (ref. 87)

⑤ **1985** Statistics textbook dedicated to meta-analytic methods released (ref. 16) *Gurevitch, Nature 2018*





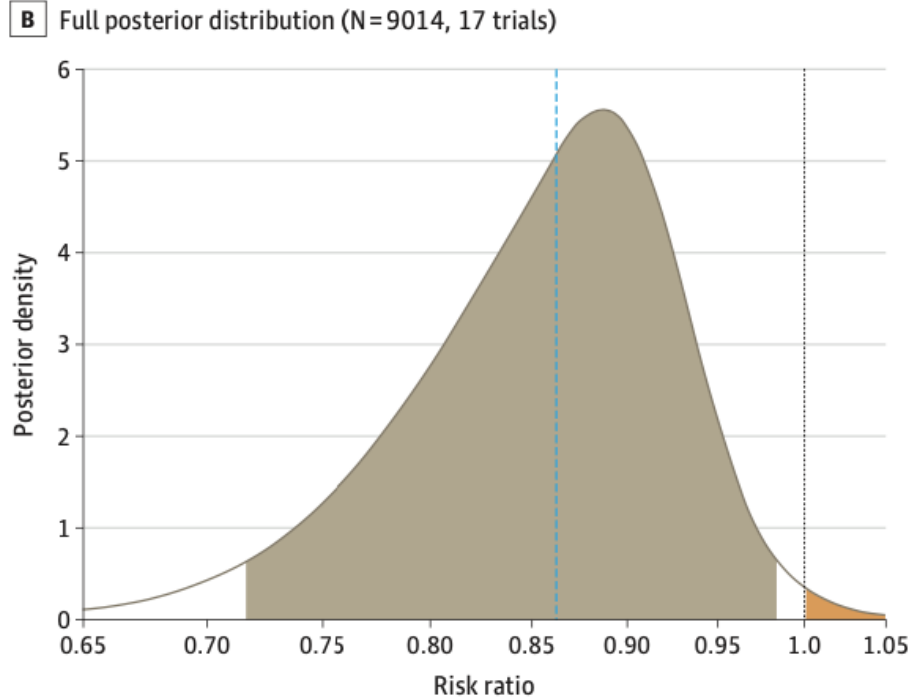
Que peut-on méta-analyser ?

Prolonged vs Intermittent Infusions of β -Lactam Antibiotics in Adults With Sepsis or Septic Shock

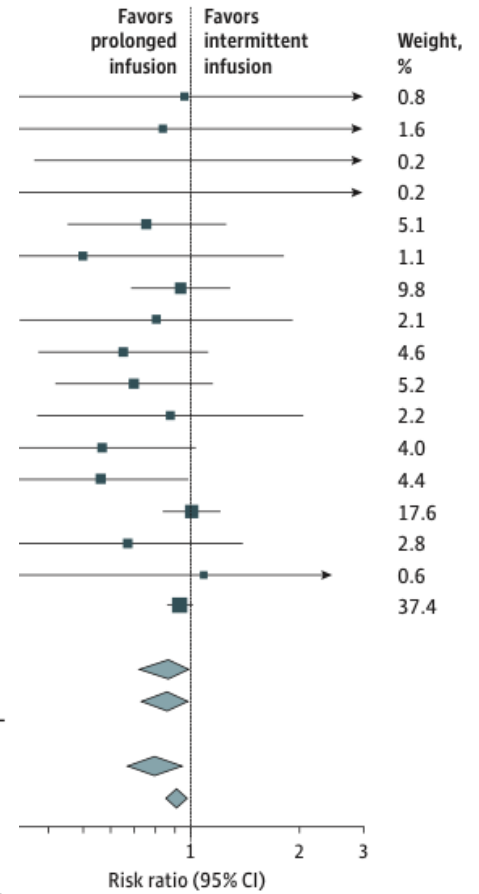
A Systematic Review and Meta-Analysis

Figure 1. All-Cause 90-Day Mortality for the Comparison B

Study	Dead (prolonged)
Georges et al, ³³ 2005	3
Rafati et al, ³⁴ 2006	5
Roberts et al, ³⁵ 2007	3
Roberts et al, ³⁶ 2009	2
Chytra et al, ³⁸ 2012	21
Dulhunty et al, ³⁹ 2013	3
Dulhunty et al, ⁴⁰ 2015	54
Jamal et al, ⁴¹ 2015	4
Jamal et al, ⁴² 2015	5
Abdul-Aziz et al, ⁴³ 2016	18
Zhao et al, ⁴⁴ 2017	7
Khan and Omar, ²² 2023	12
Mirjalili et al, ⁴⁵ 2023	14
Monti et al, ¹⁴ 2023	127
Saad et al, ⁴⁶ 2024	8
Álvarez-Moreno et al, ⁴⁷ 2024	2
Dulhunty et al, ¹⁵ 2024	864
Bayesian	
Vague priors ^a	
Semi-informative priors ^a	
Frequentist	
Hartung-Knapp-Sidik-Jonkman	
DerSimonian-Laird	



The orange area is related to an RR greater than 1 (ie, the intervention is associated with higher mortality vs standard care). The dotted line at an RR of 1 indicates no treatment effect. The figure demonstrates that the probability that prolonged infusions of β -lactam antibiotics is associated with a reduced risk of all-cause 90-day mortality (to any extent) compared with intermittent infusions is more than 99%.



s with vague priors. The whiskers represent CIs. Width of

The black boxes represent point estimates, and the areas of the diamonds represents the trials' pooled estimate CI, and the mid

^aCredible intervals are presented for bayesian analysis.

Efficacy and Tolerability of Gefapixant for Treatment of Refractory or Unexplained Chronic Cough

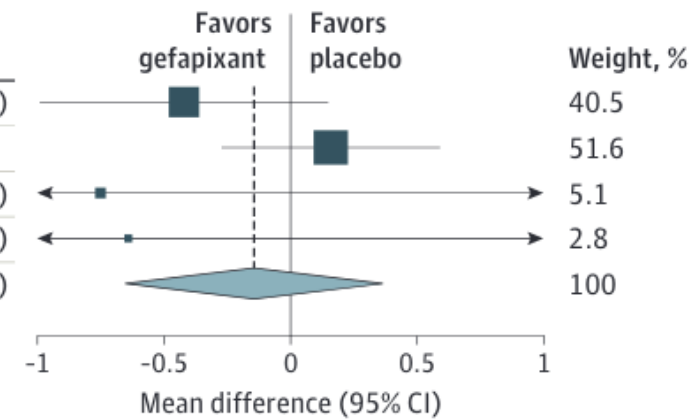
A Systematic Review and Dose-Response Meta-Analysis

Elena Kum, BSc; Matthew Patel, MD; Nermin Diab, MD, MPH; Mustafaa Wahab, BSc; Dena Zeraatkar, PhD; Derek K. Chu, MD, PhD; Paul M. O’Byrne, MD; Gordon H. Guyatt, MD, MSc; Imran Satia, MD, PhD

C Pooled mean difference in log-transformed sleep cough frequency among patients with refractory or unexplained chronic cough who received gefapixant vs placebo in 4 RCTs

Source	Gefapixant		Placebo		Mean difference (95% CI)
	Total	Mean (SD)	Total	Mean (SD)	
Abdulqawi et al, ¹² 2015	18	-0.35 (0.98)	20	0.07 (0.79)	-0.42 (-0.99 to 0.15)
Smith et al, ⁴² 2020	175	-0.56 (1.49)	61	-0.72 (1.48)	0.16 (-0.27 to 0.59)
Smith et al, ⁴⁰ 2020	25	-0.83 (4.09)	25	-0.08 (3.80)	-0.75 (-2.94 to 1.44)
Smith et al, ⁴⁰ 2020	29	-0.26 (6.52)	27	0.38 (4.87)	-0.64 (-3.64 to 2.36)
Random-effects model	247		133		-0.14 (-0.65 to 0.37)

Heterogeneity: $I^2 = 1\%$, $\tau^2 = 0.0822$, $P = .38$



A and B, Percent reductions are in comparison with placebo (dose = 0). Circles represent percent reductions observed in individual trials, and the diameters are proportional to the inverse variance of the study. Randomized clinical trials (RCTs) evaluating different doses may have contributed more than 1 circle in the graph. Doses are truncated at the largest dose for use in clinical practice

(60 mg twice daily). C, Effects for sleep cough frequency are represented as mean differences in log-transformed coughs/h. Percent reduction compared with placebo is calculated as $100 \times (e^{\text{diff}} - 1)$, where diff is the mean difference in log-transformed coughs/h. Size of data markers indicates the weight of each study in the analysis.

Comparative Effectiveness and Safety of Cognitive Behavioral Therapy and Pharmacotherapy for Childhood Anxiety Disorders

A Systematic Review and Meta-analysis

Standardized mean difference

• Cohen's *d*:

$$d = \frac{\bar{X}_1 - \bar{X}_0}{S_p}$$

$$S_p = \sqrt{\frac{(n_0 - 1)S_0^2 + (n_1 - 1)S_1^2}{n_0 + n_1 - 2}}$$

$$SE_d = \sqrt{\frac{1}{n_0} + \frac{1}{n_1} + \frac{d^2}{2(n_0 + n_1 - 2)}}$$

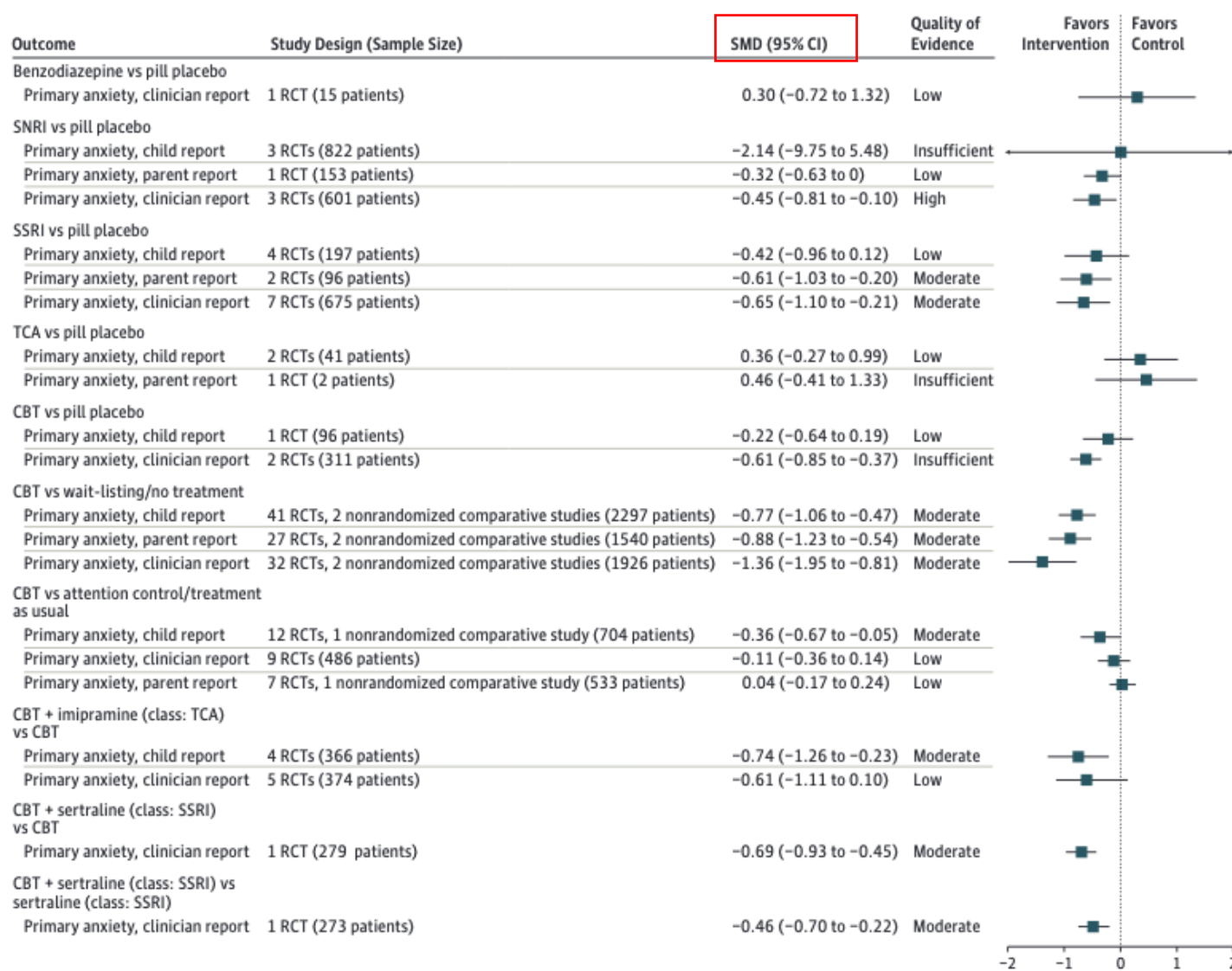
• Hedges' *g*:

$$J = 1 - \frac{3}{4(n_0 + n_1) - 9}$$

$$g = J \times d$$

$$SE_g = J \times SE_d$$

Figure 1. Pooled Effect Size and Quality of Evidence for Primary Anxiety Symptoms (Measured by Child, Parent, and Clinician Report)



CBT indicates cognitive behavioral therapy; RCT, randomized clinical trial; SMD, standardized mean difference; SNRI, serotonin-norepinephrine reuptake inhibitor; SSRI, selective serotonin reuptake inhibitor; TCA, tricyclic antidepressants.

Murad, *BMJ* 2019

Wang *JAMA Ped*, 2017

OPEN ACCESS **The prevalence of loneliness across 113 countries: systematic review and meta-analysis**

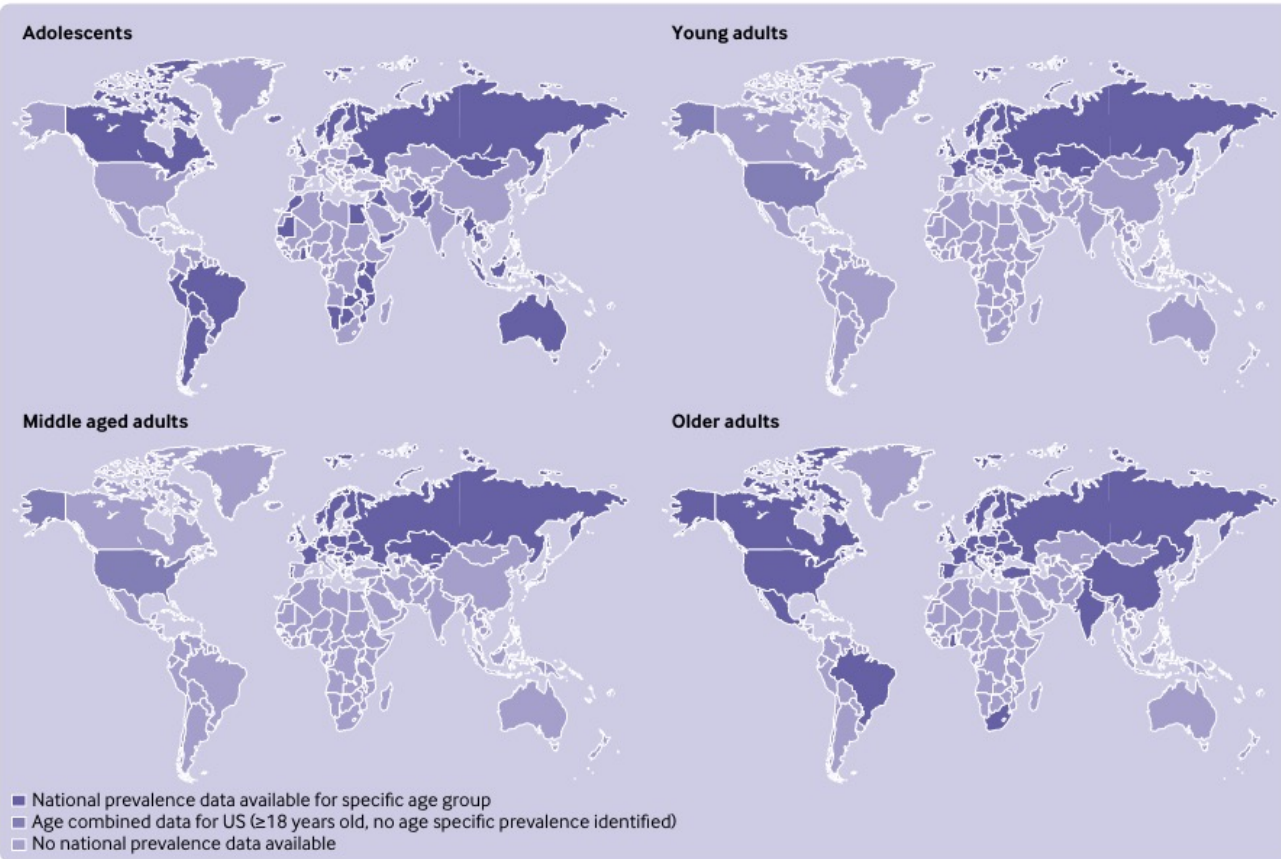


Fig 1 | International estimates for prevalence of loneliness by age groups

The placebo response rate in pharmacological trials in patients with irritable bowel syndrome: a systematic review and meta-analysis

	Number of trials with this variable	Placebo	
		Response rate (95% CI)	P
Global improvement responder endpoint	60	27.3% (24.3–30.9)	90.7%
IBS-D	20	27.2% (22.4–32.7)	89.4%
IBS-C	22	23.3% (19.3–27.9)	93.2%
Abdominal pain responder endpoint	48	34.4% (31.2–37.8)	89.4%
IBS-D	23	33.6% (30.1–37.4)	81.6%
IBS-C	13	31.5% (25.1–38.7)	94.6%
Combined responder endpoint	73	31.3% (28.2–34.5)	91.6%
US FDA endpoint responder			
Composite	20	17.9% (15.2–21.0)	82.8%
IBS-D	9	21.1% (16.6–26.4)	84.0%
IBS-C	10	15.7% (12.6–19.4)	80.9%
Abdominal pain	19	35.1% (30.6–39.9)	89.1%
IBS-D	8	37.0% (29.5–45.1)	90.3%
IBS-C	9	34.6% (28.4–41.5)	90.9%
Bowel symptom	19	28.3% (23.4–33.8)	92.7%
IBS-D	9	29.2% (20.9–39.3)	94.2%
IBS-C	10	27.4% (21.6–34.1)	91.9%

US FDA=US Food and Drug Administration. IBS-C=irritable bowel syndrome with constipation. IBS-D=irritable bowel syndrome with diarrhea. IBS-D=irritable bowel syndrome with diarrhea specific study (either global improvement, improvement in abdominal pain, or the US FDA endpoint).

Table 1: Placebo response rate, intervention response rate, and the therapeutic gain of the p

A guide to systematic review and meta-analysis of prediction model performance

Assessment of C-Reactive Protein Diagnostic Test Accuracy for Late-Onset Infection in Newborn Infants A Systematic Review and Meta-analysis

Figure 4. Coupled Forest Plot Showing Sensitivity and Specificity of C-Reactive Protein (CRP) for Diagnosing Late-Onset Infection

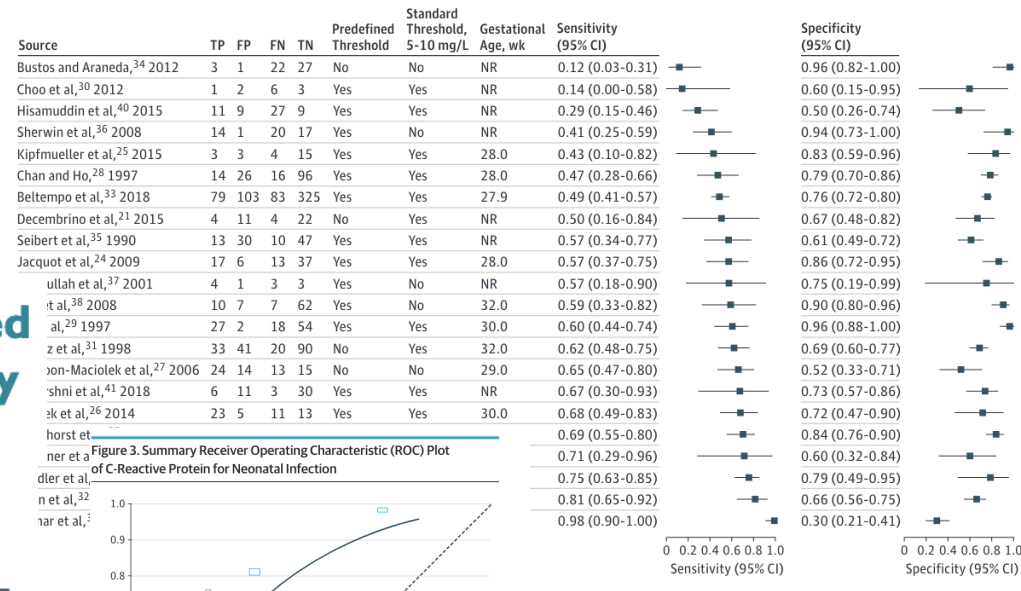
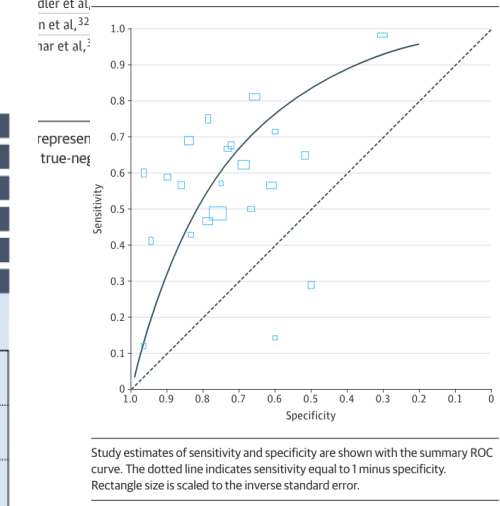


Figure 3. Summary Receiver Operating Characteristic (ROC) Plot of C-Reactive Protein for Neonatal Infection



gestational age not reported; TP, true-positive results; FP, false-positive results; FN, false-negative results; TN, true-negative results; CI, confidence interval.

Cardiovascular immunotoxicities associated with immune checkpoint inhibitors: a safety meta-analysis

Diseases & Risk factors

Peto OR (95%CI)

Number needed to harm (95%CI)

Incidence (per 1000pts)

Dyslipidemia	2,981	74 (37-5,824)	19.3
Myocarditis	9,455	462 (286-1,202)	3.2
Pericardial	16,136	294 (207-509)	8.3
HF	19,802	260 (187-427)	8.7
Ischemic stroke	21,039	330 (209-773)	8.8
MI	21,578	484 (291-1,445)	7.4
Cardiac death or shock	20,065		
Hypertension	15,691		
Endocardial	4,499		
VTE	23,722		
Supraventricular arr.	21,888		
Ventricular arr.	5,544		
TdP / QT prol	2,811		
Pulm. Hypertens.	4,406		
Valvular	3,132		
Conduction blocks	6,259		

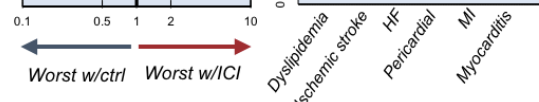


Fig 4 | Forest plots of extracted performance statistics of the additional mortality at 30 days in patients undergoing coronary artery bypass grafting

Debray, BMJ 2017

Dolladille, Eur Heart J 2021

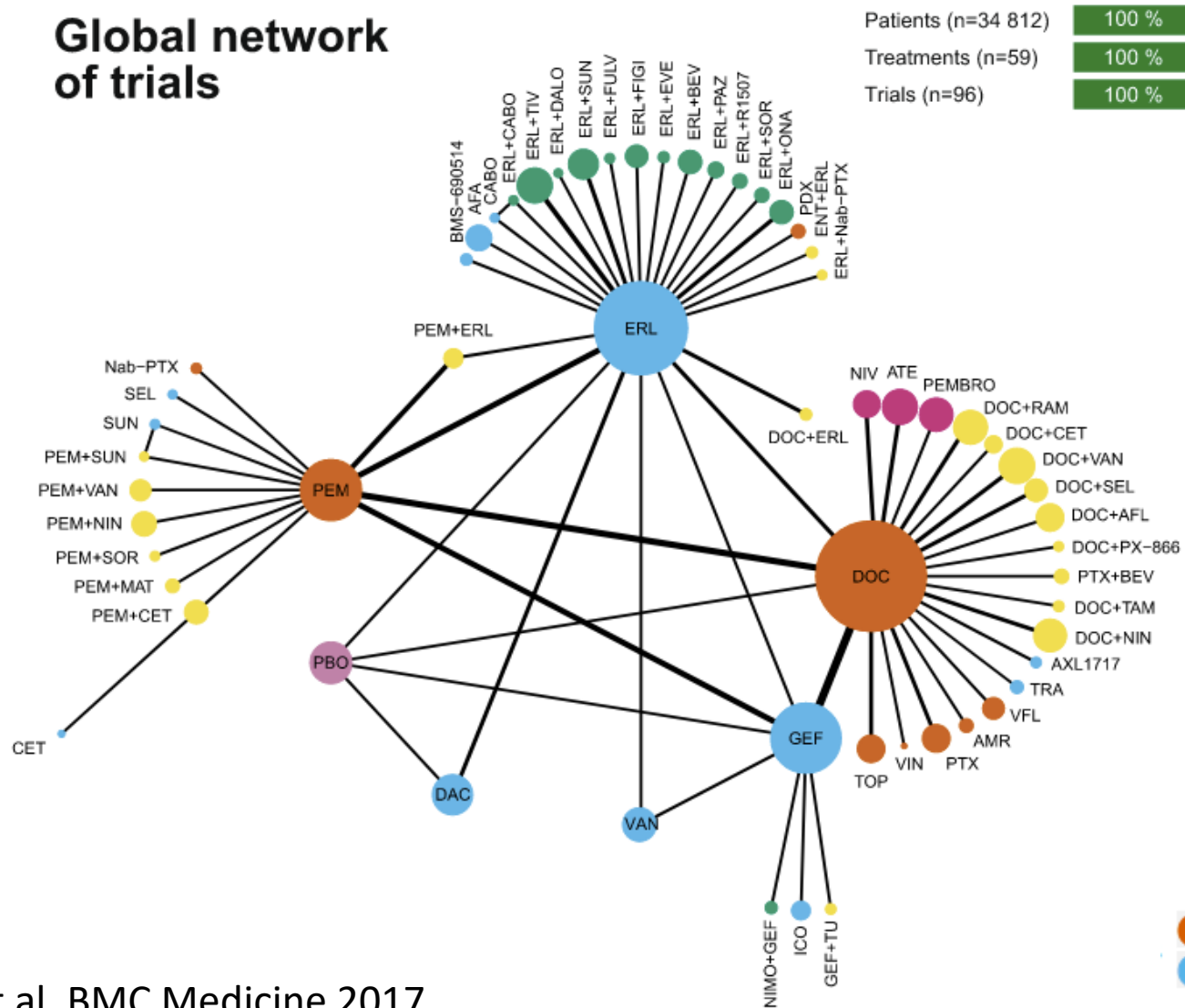
Dolladille, Alexandre

Brown, JAMA Pediatr 2020



Mais des questions parfois trop restreintes

Global network of trials

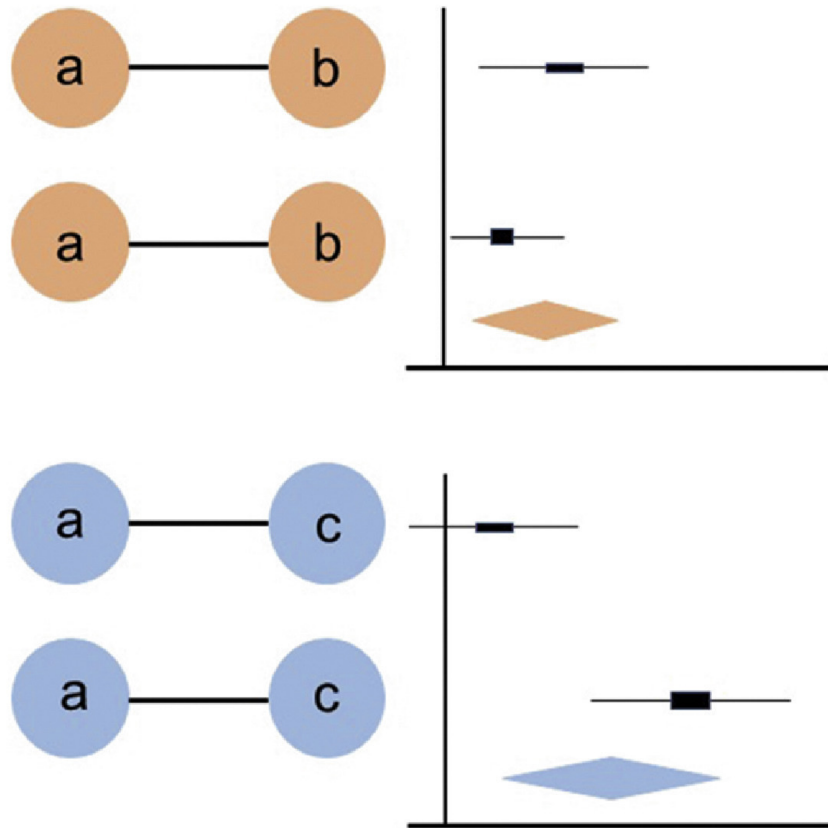


Traitements du cancer du poumon en 2nde ligne

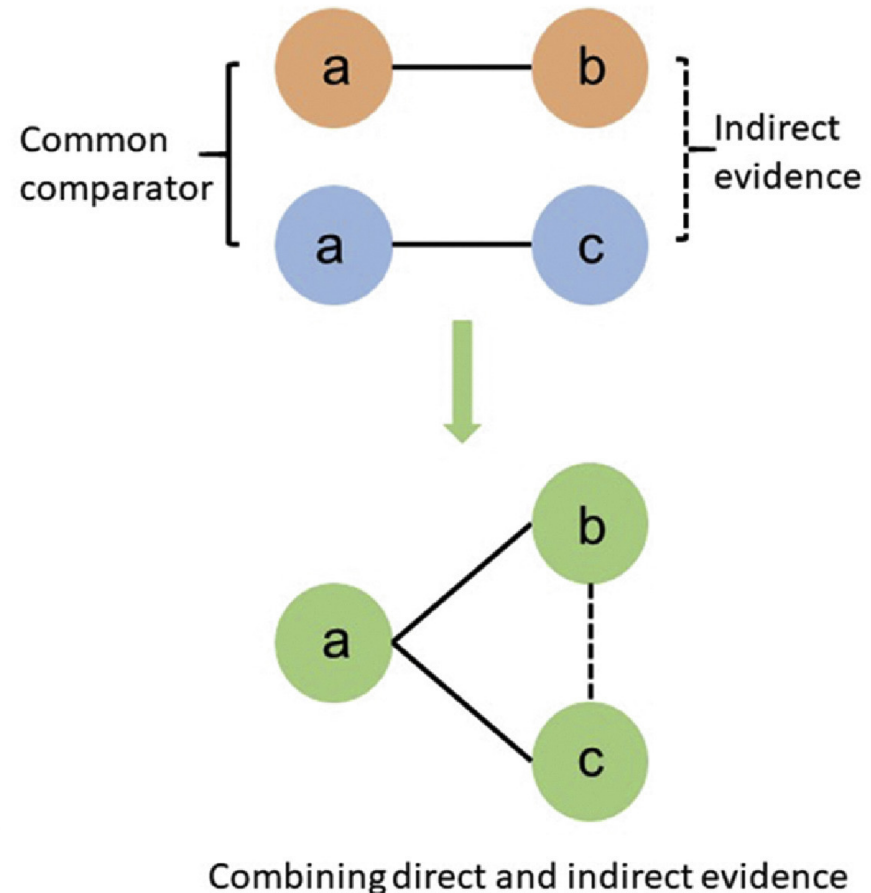
Méta-analyses en réseaux

Permet de comparer l'ensemble des traitements disponibles dans une condition donnée

A

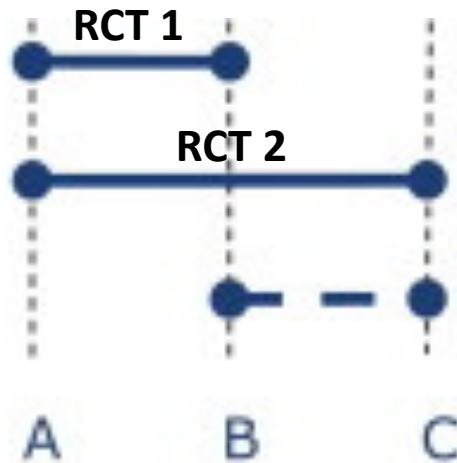
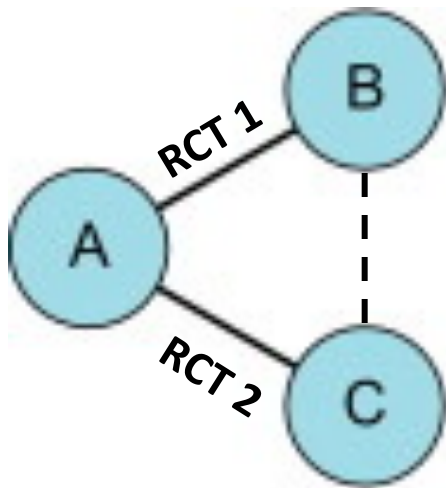


B



Principe de l'extrapolation

Hypothèse de transitivité (échangeabilité des effets)



Direct comparison



Indirect comparison



Hypothèse de transitivité

- Bien entendu jamais vraie !!
- Mais...les effets traitements sont relativement stables

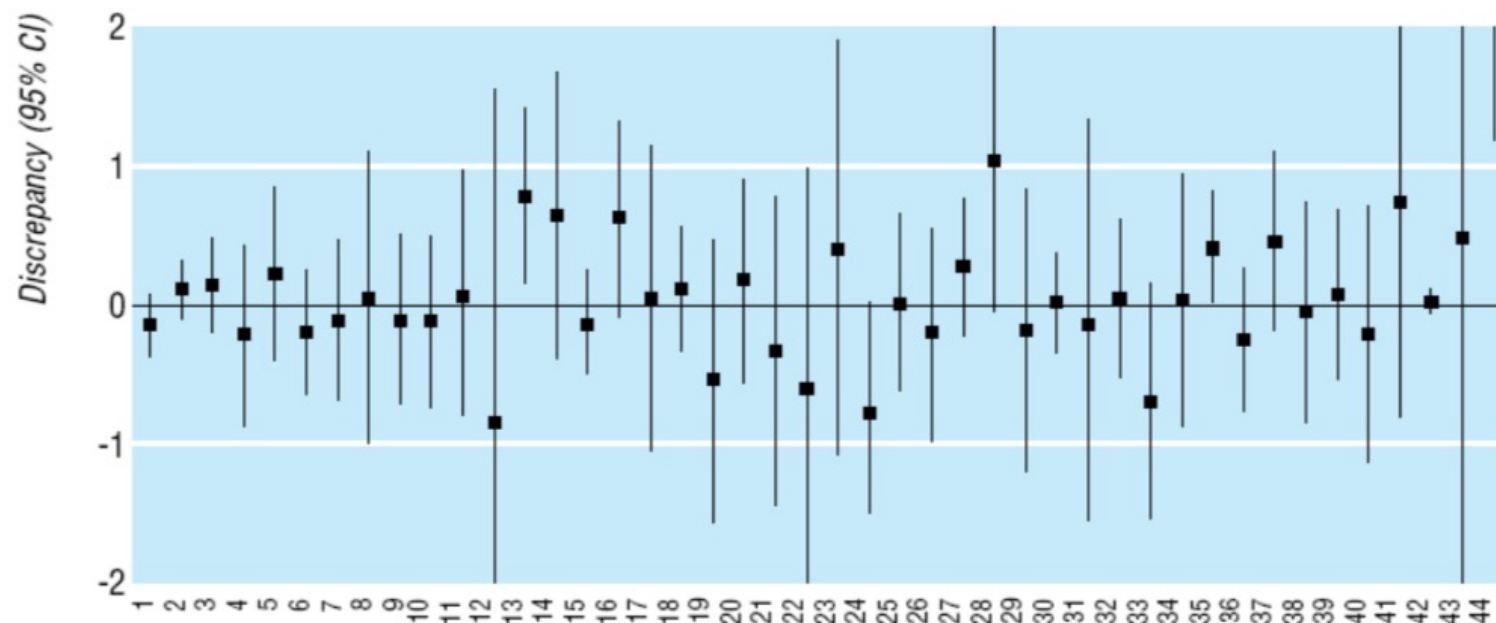
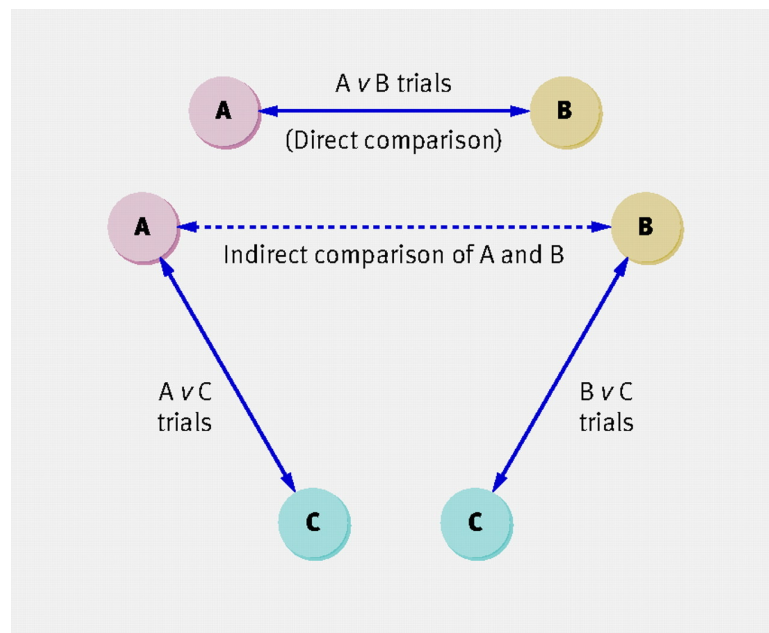


Fig 1 Discrepancy between direct and adjusted indirect comparison defined as difference in estimated log relative risk (meta-analyses 1-39) or difference in estimated standardised mean difference (meta-analysis 40) or difference in estimated mean difference (meta-analyses 41-44): empirical evidence from 44 published meta-analyses (see webextra table A)

Meta-analyses

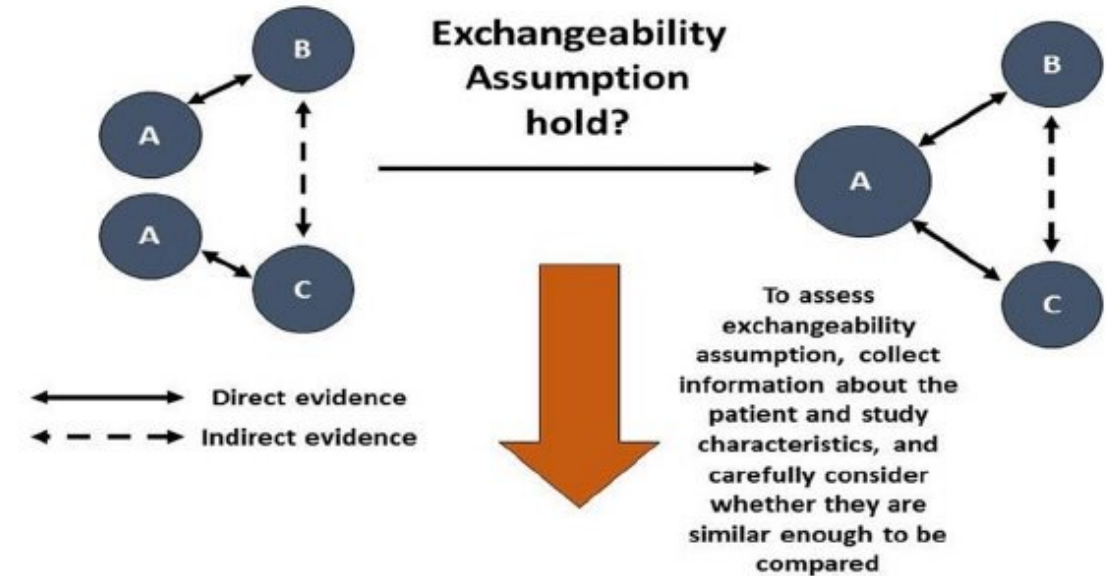
Hypothèse de transitivité

Intestable

Mais on peut évaluer sa **plausibilité**

- Par le raisonnement
- Par analyse de sensibilité
- Par ajustement (modificateurs de l'effet)
- Populations similaires en moyenne entre les traitements
- Traitements et design des essais cliniques homogènes pour être regroupés

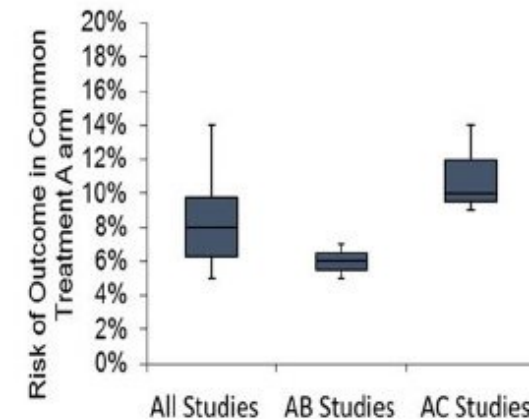
a Network meta-analysis and assessment of exchangeability assumption



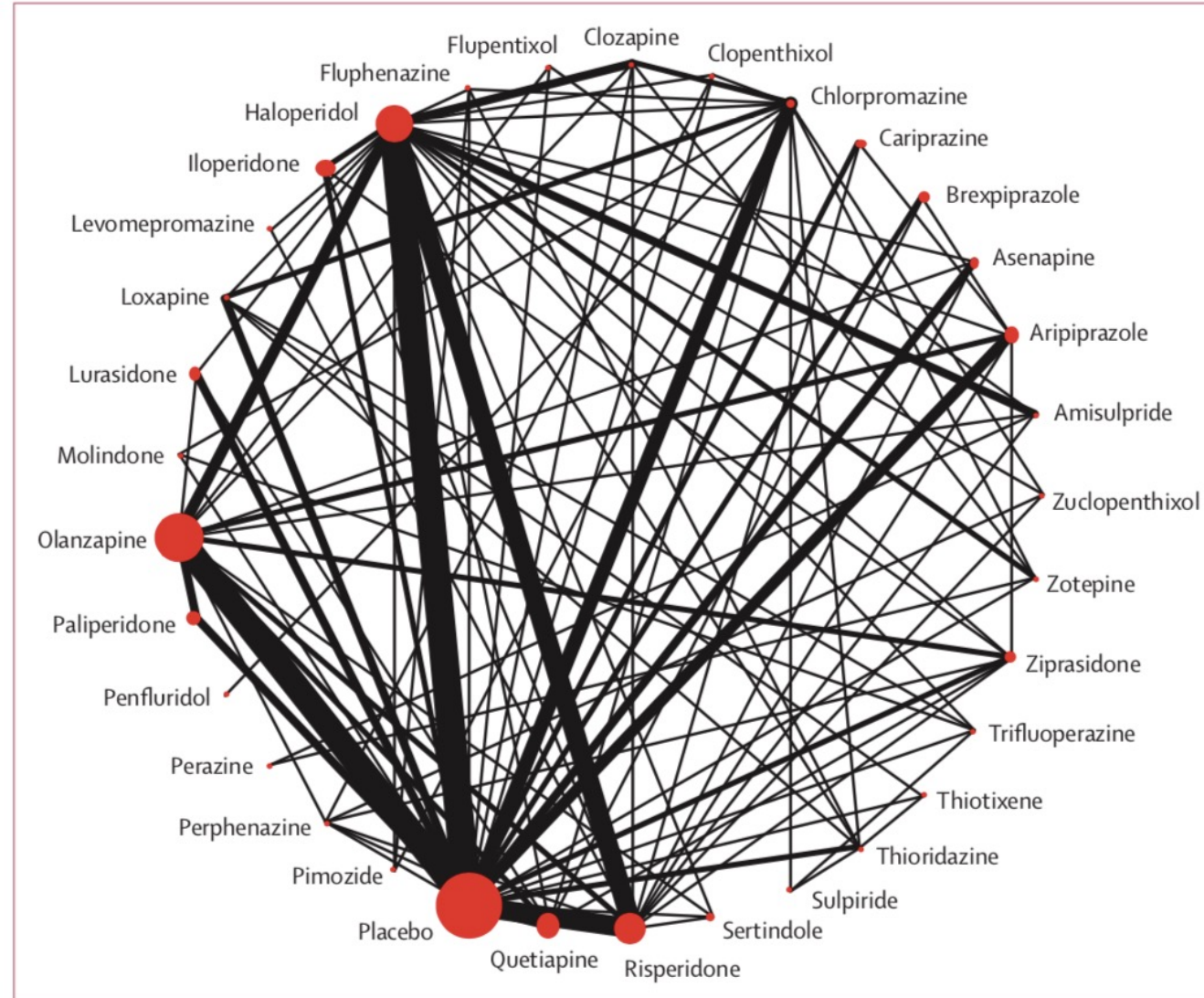
b Summary of patient and study characteristics

	All studies	Drug A versus B	...	Drug A versus C
Age	68	62		72
Duration of diabetes	15 years	10 years		20 years
.....
Concomitant medication	15%	10%		20%

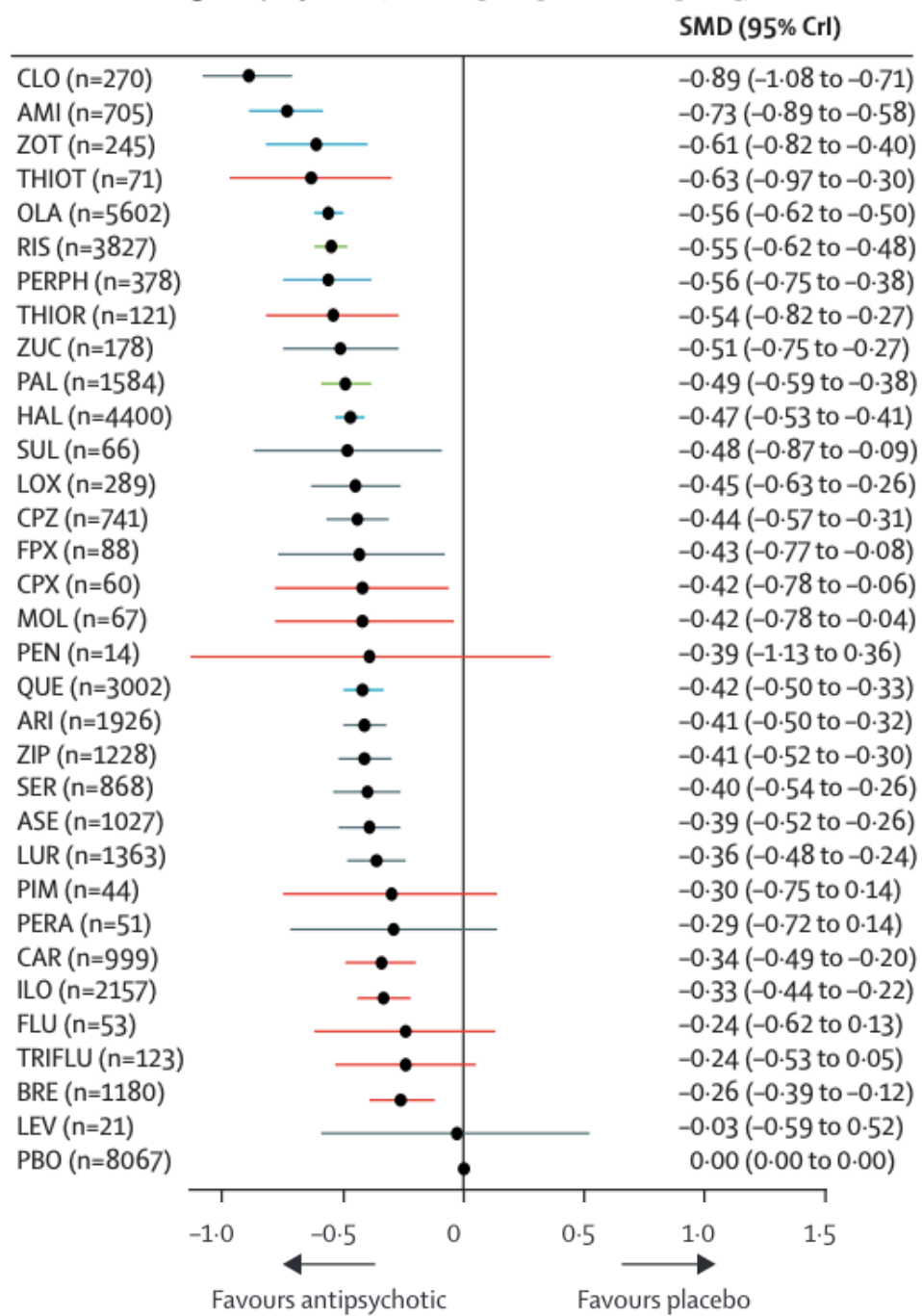
c Box Plot on risk of outcome in common arm



Comparative efficacy and tolerability of 32 oral antipsychotics for the acute treatment of adults with multi-episode schizophrenia: a systematic review and network meta-analysis



A Overall change in symptoms (N_t=218 [54%], n_r=40815 [76%])



CLO	NA	-0.08 (-0.68 to 0.52)	NA	-0.23 (-0.98 to 0.53)	-0.17 (-0.63 to 0.28)
-0.16 (-0.39 to 0.07)	AMI	NA	NA	-0.15 (-0.43 to 0.14)	-0.08 (-0.38 to 0.23)
-0.28 (-0.54 to -0.03)	-0.13 (-0.37 to 0.12)	ZOT	NA	NA	NA
-0.26 (-0.65 to 0.13)	-0.10 (-0.47 to 0.27)	0.02 (-0.37 to 0.43)	THIOT	NA	NA
-0.33 (-0.52 to -0.15)	-0.18 (-0.33 to -0.02)	-0.05 (-0.26 to 0.16)	-0.07 (-0.42 to 0.27)	OLA	-0.15 (-0.35 to 0.04)
-0.34 (-0.53 to -0.16)	-0.18 (-0.34 to -0.03)	-0.06 (-0.27 to 0.16)	-0.08 (-0.43 to 0.26)	-0.01 (-0.09 to 0.08)	RIS

Comparative effects of 18 antipsychotics on metabolic function in patients with schizophrenia, predictors of metabolic dysregulation, and association with psychopathology: a systematic review and network meta-analysis

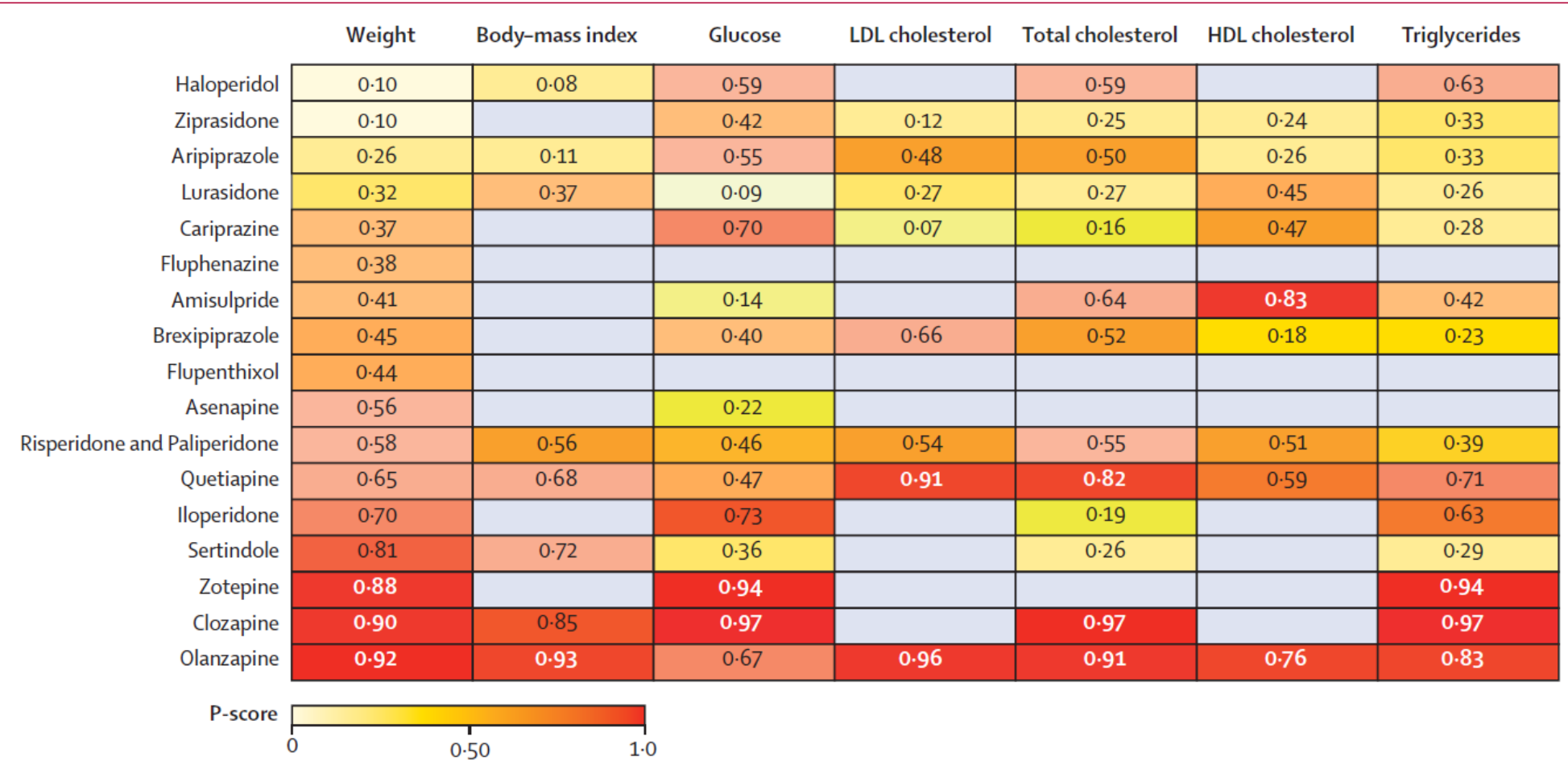


Figure 3: Heat map of antipsychotic drugs ranked according to associated degree of alteration in bodyweight, body-mass index, and metabolic parameters

Ranking en NMA

A Markov chain approach for ranking treatments in network meta-analysis

Anna Chaimani^{1,2} | Raphaël Porcher¹ | Émilie Sbidian^{3,4} | Dimitris Mavridis^{1,5}

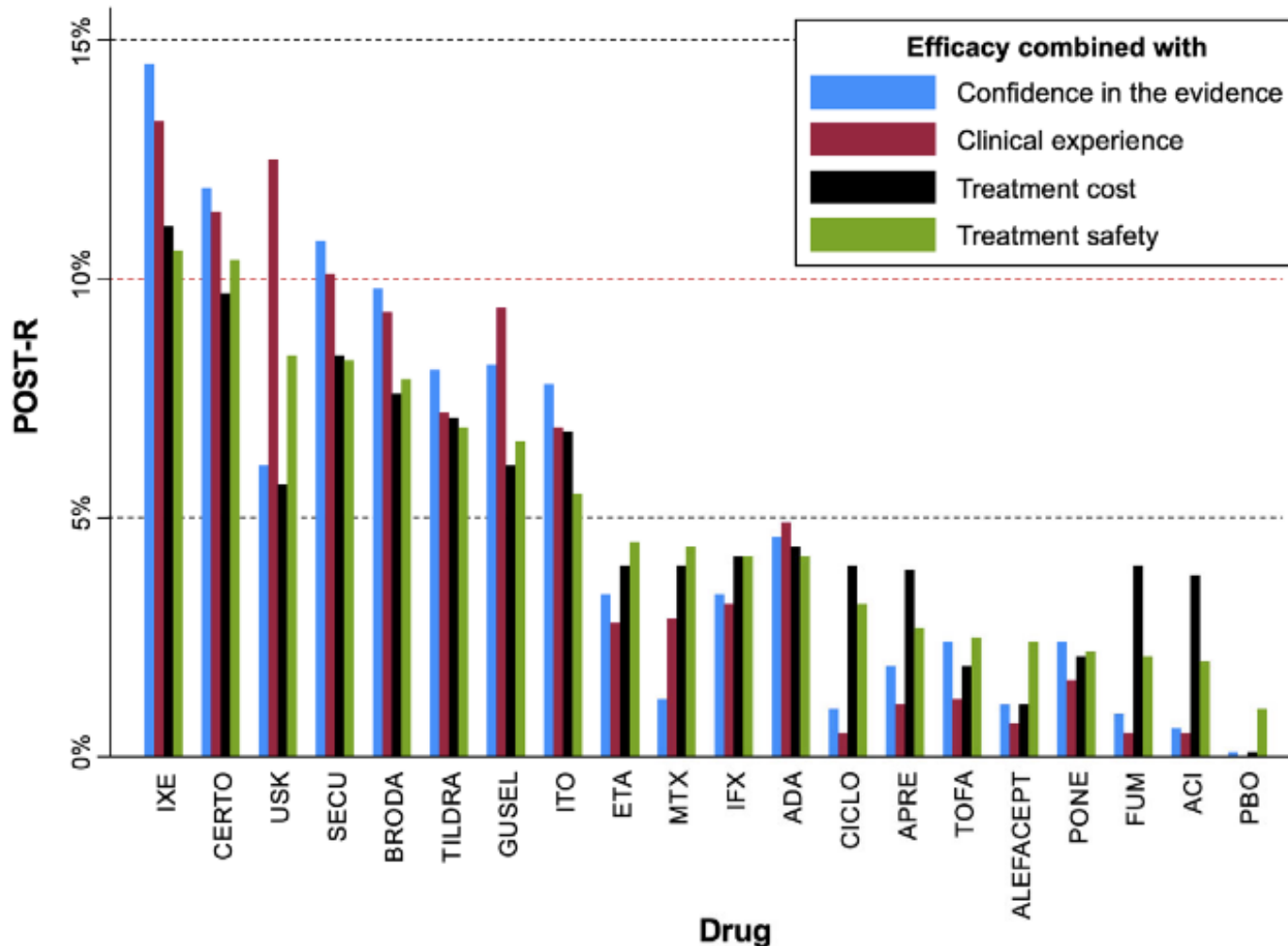


FIGURE 4 Ranking results of the psoriasis network using the POST-R measure and considering the relative effects for efficacy in the transition probabilities and different characteristics in the initial probability distribution as indicated by the legend. Drugs have been ordered according to the ranking for efficacy and safety (green bars) [Colour figure can be viewed at wileyonlinelibrary.com]

Méta-analyse en composante

Interventions complexes

“Complex interventions in health care, whether therapeutic or preventative, comprise a number of separate elements which seem essential to the proper functioning of the intervention although the ‘active ingredient’ of the intervention that is effective is difficult to specify...**The greater the difficulty in defining precisely what, exactly, are the ‘active ingredients’ of an intervention and how they relate to each other, the greater the likelihood that you are dealing with a complex intervention.**”

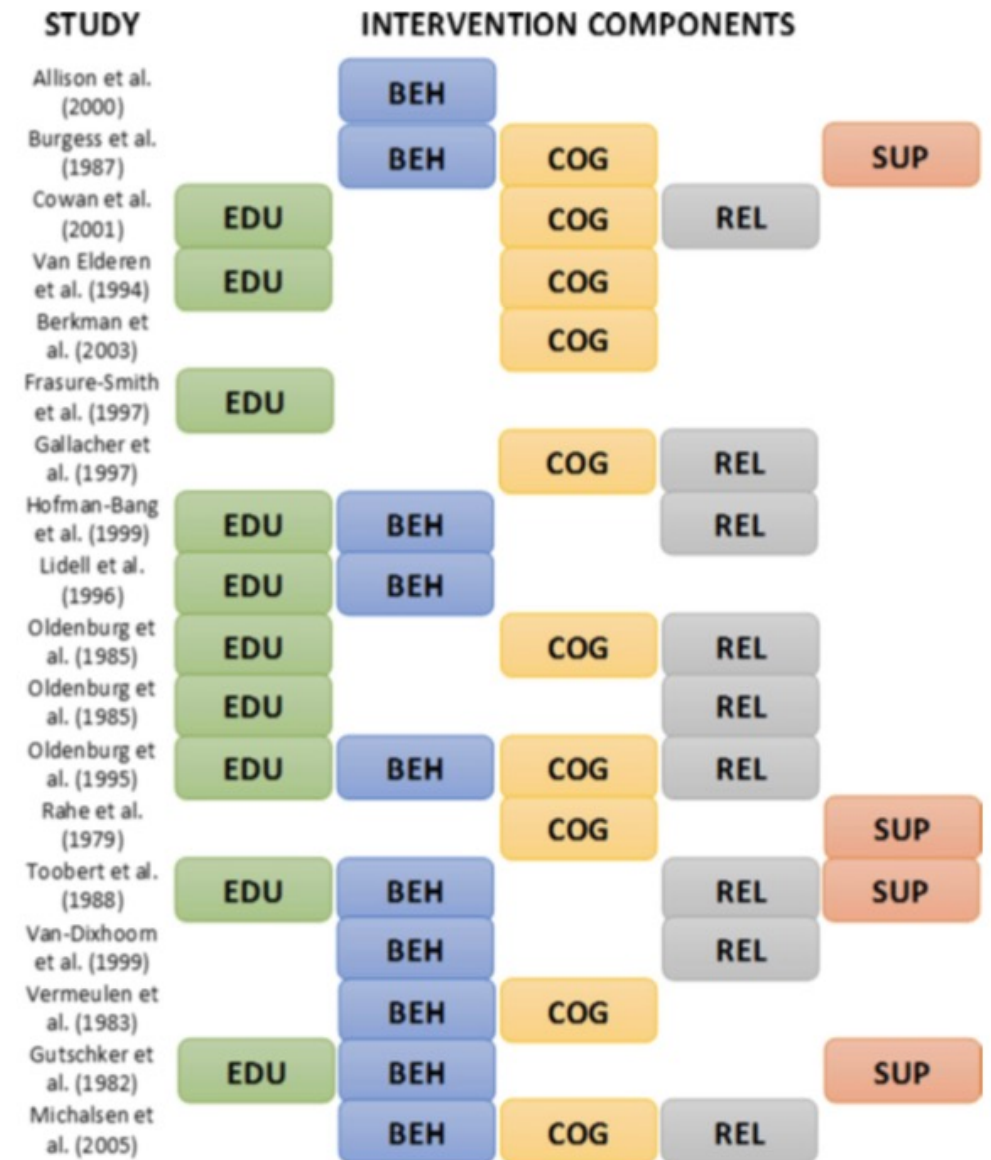
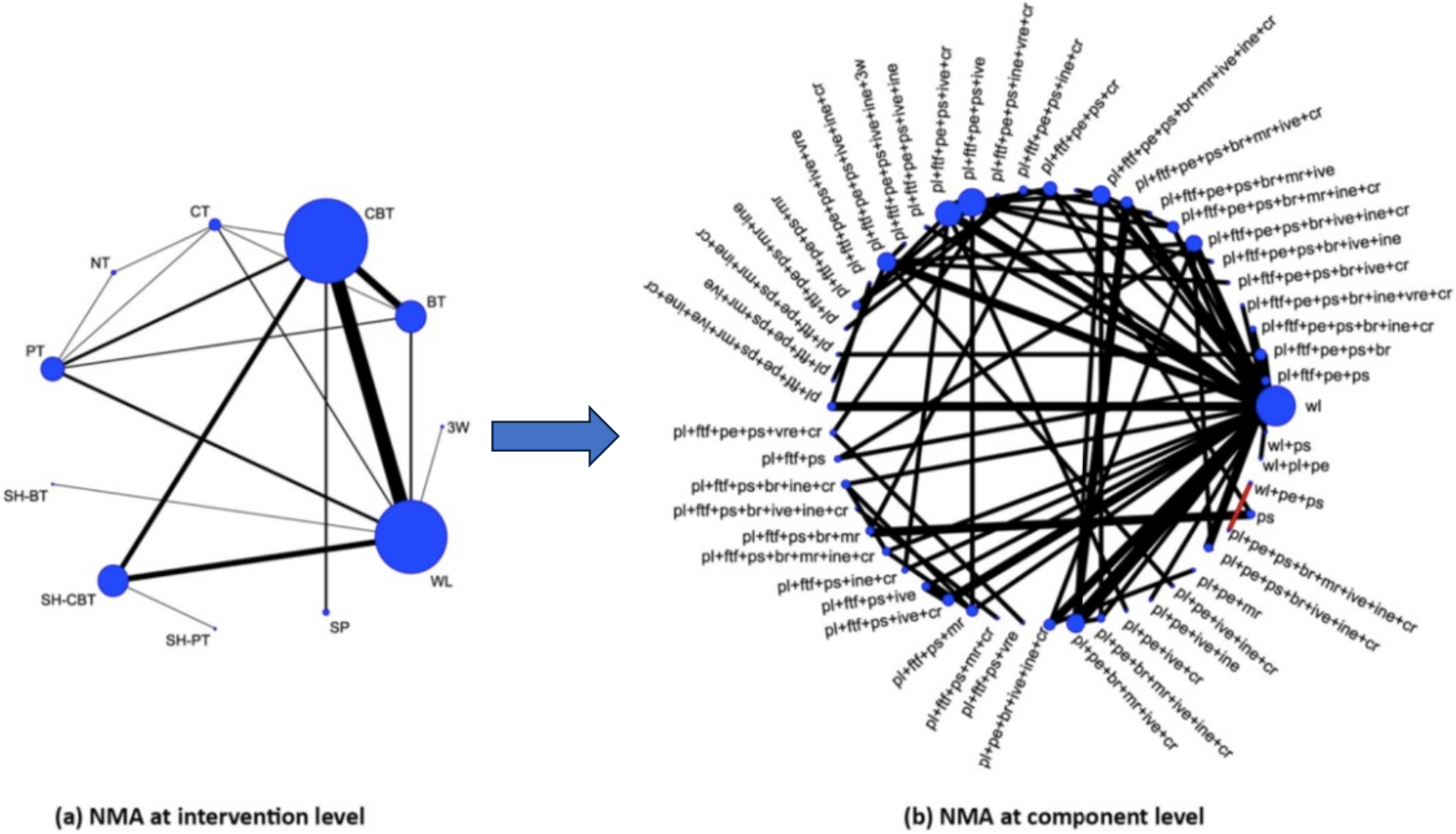


Figure 2 Intervention components in the studies integrated by Welton *et al* (a sample of 18 from 56 active treatment arms). EDU, educational component; BEH, behavioural component; COG, cognitive component; REL, relaxation component; SUP, psychosocial support component.

Network plot for psychological interventions for panic disorder at intervention and component level



Reconnecter les réseaux séparés

- Composantes:
- Additivité
 - Interactions...

Adjunctive Psychotherapy for Bipolar Disorder

A Systematic Review and Component Network Meta-analysis

Figure 2. Geometry of Networks for Treatment-Level Comparisons

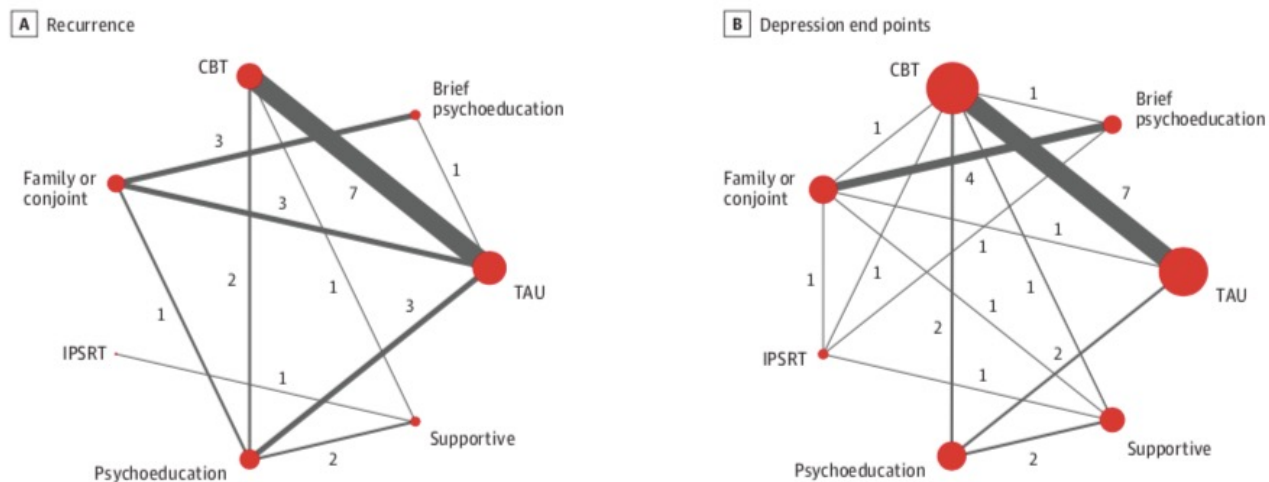


Table 1. Pairwise and Network Meta-analysis for Treatment-Level Comparisons of Recurrence (Odds Ratios)^a

Brief psychoeducation	NA	1.15 (0.60-2.20)	NA	NA	NA	0.34 (0.08-1.49)
0.66 (0.28-1.55)	CBT	NA	NA	0.89 (0.41-1.95)	0.81 (0.26-2.52)	0.52 (0.32-0.83)
1.15 (0.63-2.09)	1.75 (0.87-3.49)	Family or conjoint therapy	NA	0.51 (0.13-1.92)	NA	0.31 (0.16-0.62)
0.22 (0.04-1.15)	0.34 (0.08-1.46)	0.20 (0.04-0.92)	IPSRT	NA	1.76 (0.49-6.35)	NA
0.66 (0.28-1.58)	1.01 (0.61-1.68)	0.58 (0.29-1.16)	2.96 (0.72-12.11)	Psychoeducation	0.54 (0.28-1.03)	0.51 (0.26-0.99)
0.39 (0.14-1.08)	0.60 (0.30-1.19)	0.34 (0.14-0.82)	1.76 (0.49-6.35)	0.59 (0.33-1.06)	Supportive	NA
0.34 (0.16-0.74)	0.52 (0.34-0.79)	0.30 (0.17-0.53)	1.53 (0.35-6.62)	0.52 (0.32-0.84)	0.87 (0.43-1.76)	TAU

Table 3. Data on Incremental Effectiveness of Treatment Components Regarding Outcome Variables^a

Component	Definition	Groups with component, No. ^b	iOR recurrence (95% CrI)	iSMD depression (95% CrI)	iSMD mania (95% CrI)	iOR dropout (95% CrI)
Psychoeducation: information only	Information giving without skill practice	7-11	0.95 (0.30 to 3.22)	0.25 (-0.57 to 1.06)	0.14 (-1.07 to 1.42)	1.27 (0.65 to 2.86)
Psychoeducation: information plus active skill practice	Information giving with skills development and practice	23-42	0.93 (0.27 to 3.46)	0.01 (-0.94 to 0.87)	-0.03 (-1.58 to 1.56)	1.09 (0.53 to 2.44)
Self-monitoring assignments	Patient tracks moods, prodromal symptoms, sleep, or thoughts	20-33	0.22 (0.04 to 1.35)	-0.13 (-0.71 to 0.58)	-0.17 (-1.25 to 0.95)	1.27 (0.55 to 2.94)
Self-management assignments	Patient learns to implement preventive strategies for managing early warning signs	17-30	1.70 (0.45 to 6.69)	0.38 (-0.36 to 1.07)	0.30 (-0.70 to 1.22)	1.34 (0.61 to 3.39)
Cognitive restructuring	Guided practice in challenging self-defeating thoughts and rehearsing adaptive thinking	10-16	0.64 (0.20 to 2.10)	-1.26 (-2.10 to -0.35)	-1.00 (-2.15 to 0.16)	1.08 (0.35 to 2.66)
Maintaining regular daily rhythms	Patient coached to regulate daily activities and sleep and wake schedules	12-21	1.63 (0.55 to 4.10)	-0.78 (-1.28 to -0.24)	-0.42 (-1.08 to 0.28)	1.15 (0.66 to 1.95)
Behavioral activation assignments	Patient coached to plan pleasurable activities to increase or modulate engagement with environment	3-6	0.69 (0.18 to 2.72)	0.92 (0.11 to 1.71)	0.98 (-0.10 to 2.03)	0.81 (0.30 to 2.18)
Behavioral problem solving	Patient coached on how to identify and define problems and generate, evaluate, and implement solutions	16-28	1.17 (0.39 to 3.22)	0.30 (-0.26 to 0.83)	0.10 (-0.64 to 0.88)	0.88 (0.43 to 1.86)
Interpersonal problem-solving	Patient encouraged to identify interpersonal habits and consider alternative behaviors	7-12	0.81 (0.08 to 8.94)	0.55 (-0.37 to 1.39)	0.73 (-0.47 to 1.90)	0.55 (0.17 to 1.72)
Communication training	Coaching families, couples, or groups on effective speaking and listening skills with in-session rehearsal	5-8	2.46 (0.07 to 74.4)	-0.84 (-1.81 to 0.23)	-0.80 (-2.23 to 0.63)	1.86 (0.46 to 6.49)
Group format	Sessions occur with other patients; includes multifamily groups	10-18	0.94 (0.18 to 4.76)	-0.12 (-0.99 to 0.78)	-0.05 (-1.42 to 1.29)	1.07 (0.49 to 2.16)
Family format	Sessions occur with patient's family members or spouse	11-15	0.16 (0.02 to 1.22)	0.93 (-0.34 to 2.15)	0.76 (-0.96 to 2.45)	0.39 (0.15 to 1.11)
Individual format	Sessions are limited to the individual patient	17-26	1.27 (0.23 to 5.21)	1.01 (-0.12 to 2.07)	0.41 (-1.17 to 1.95)	0.73 (0.29 to 1.95)

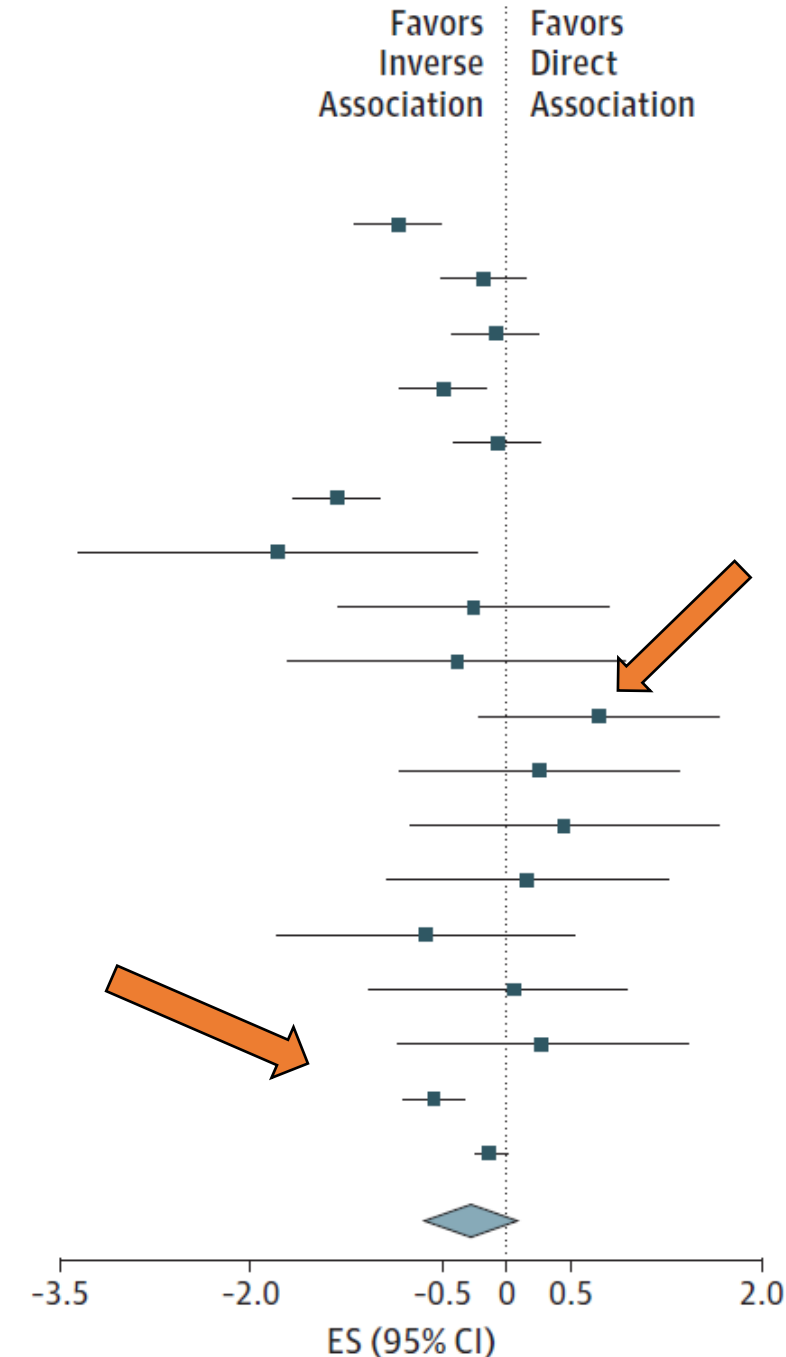
Une méta-analyse, pourquoi faire ?

1/ Estimer l'effet moyen d'une intervention

- Intervention: traitement, test diagnostic, biomarqueur...
- Interprétation globale des informations disponibles
- Augmenter la puissance et la précision

2/ Evaluer et comprendre l'hétérogénéité de l'effet d'une intervention

- Qu'est ce qui explique la différence d'effet entre plusieurs études ?



L'hétérogénéité en méta-analyse

Variation d'effet observé entre les résultats des études qui ne peut être expliquée par la chance (variabilité d'échantillonnage)

Différentes métriques

- Q-test for heterogeneity
- I-squared statistic
- Tau-squared and Tau
- Prediction interval

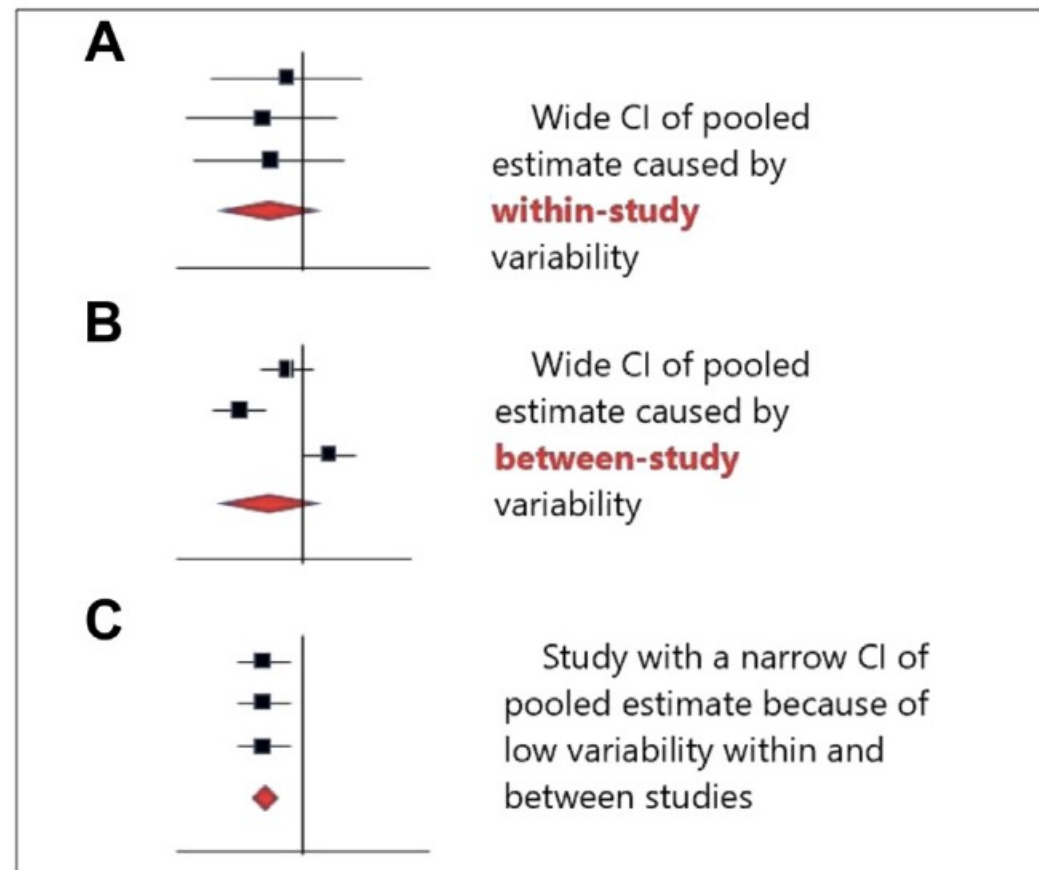


Fig. 1. Effect of within-study and between-study variability on the Confidence Interval of pooled estimates.

Qu'est ce qui entraine de l'hétérogénéité ?

- **Design des études** : randomisation, aveugle, parallèle-crossover...
- **Populations**: sévérité de la pathologie, caractéristiques des patients, critères d'exclusions...
- **Intervention**: type, traitement seul ou en association, dose, durée...
- **Comparateur**: placebo, traitement actif...
- **Critères de jugements**: type, échelle, temps...

Explorer l'hétérogénéité

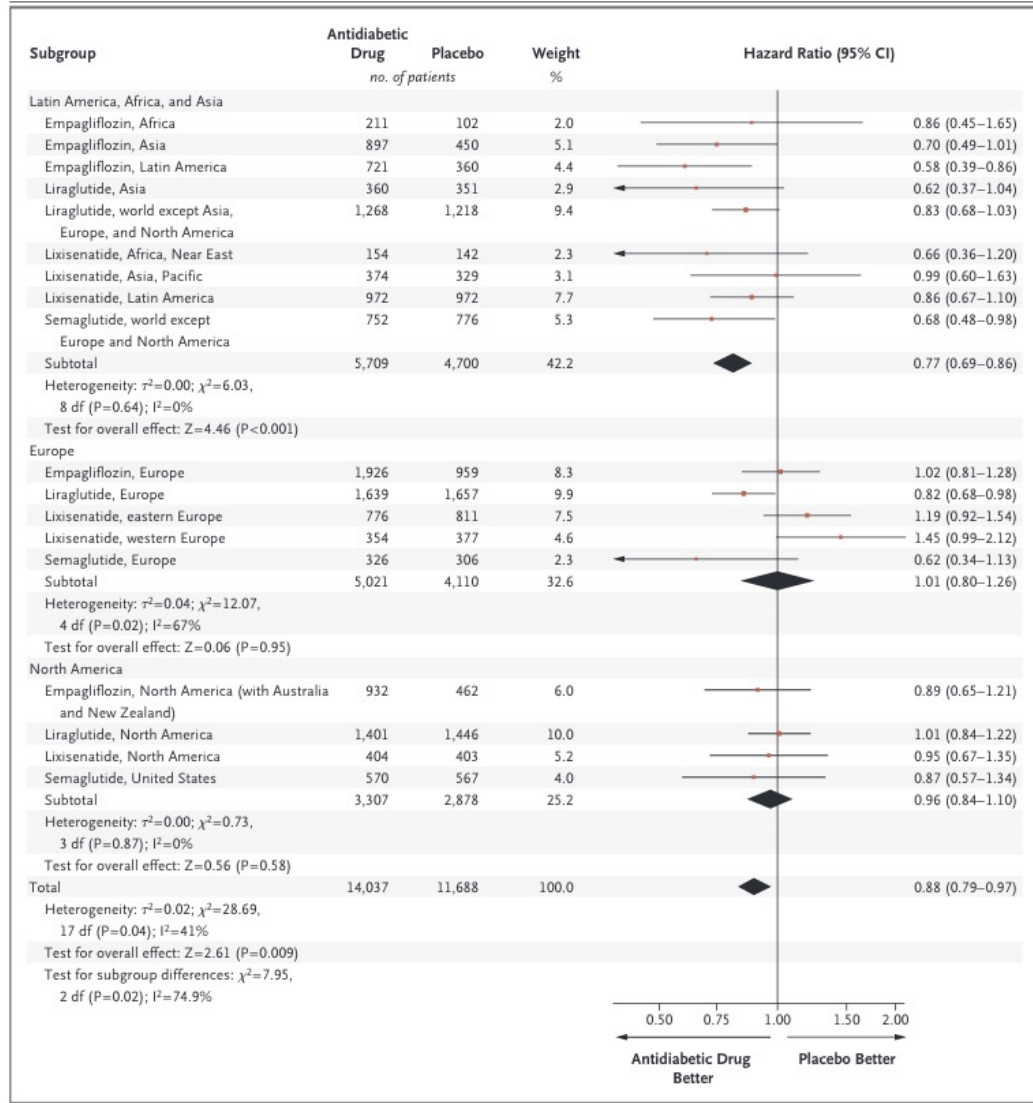


Figure 2. Meta-regression Analysis of All-cause Mortality by Baseline LDL-C Level (34 RCTs)

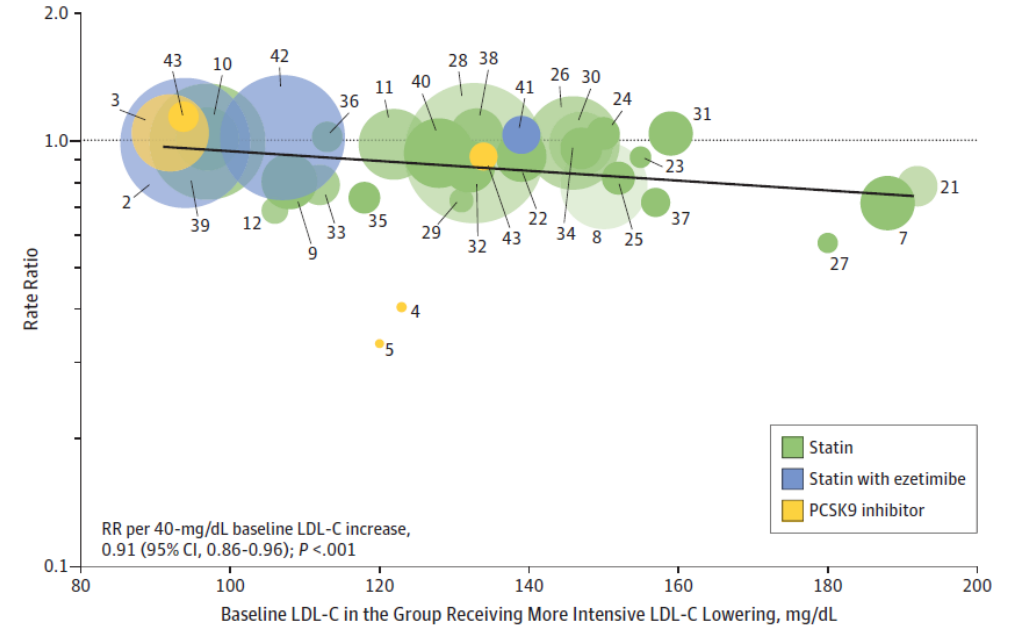
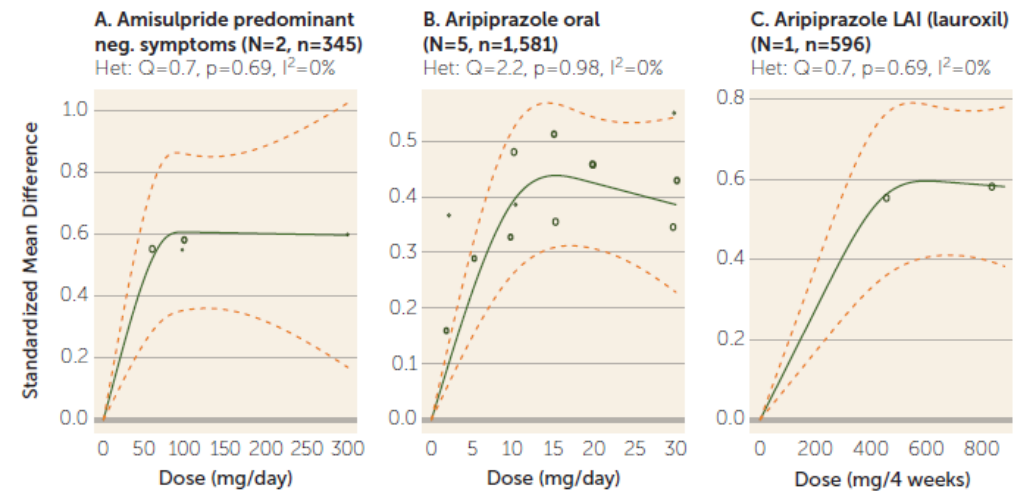


FIGURE 2. Dose-response curves of individual antipsychotic drugs^a



Limites analyses en sous-groupes et régressions

Faible puissance
Outliers
Multiplicité des tests
Biais écologique

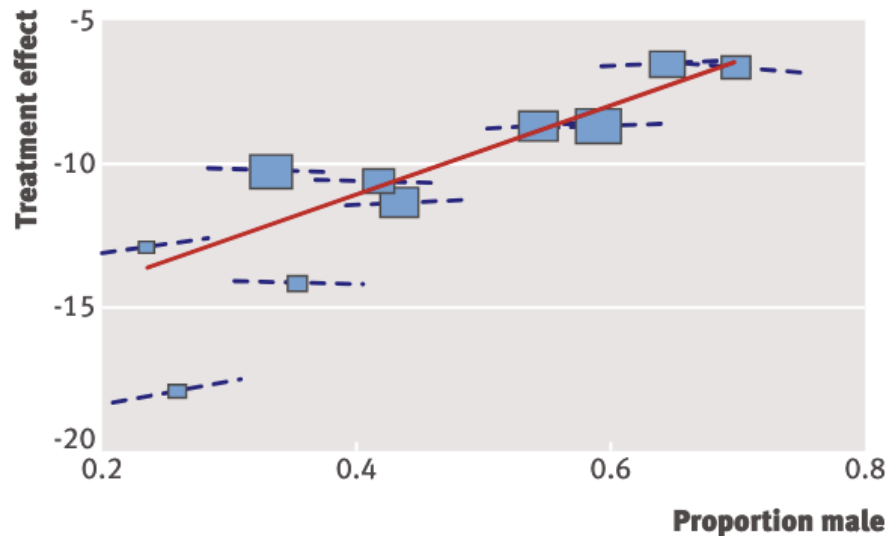
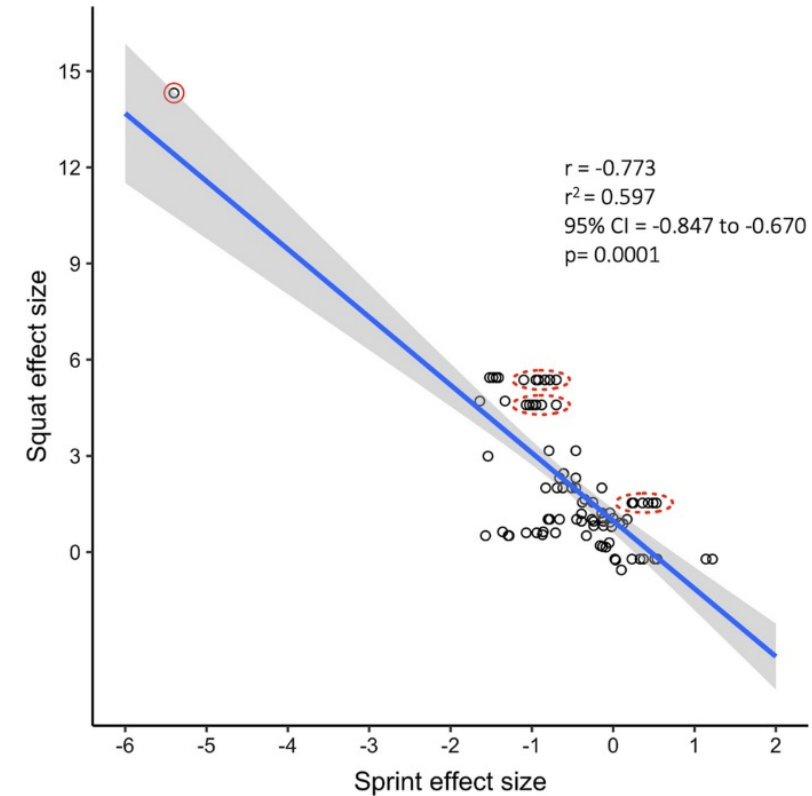


Fig 3 | An example of ecological bias within an aggregate data meta-analysis

Riley, BMJ 2010



Kadlec, Sport Medicine 2022

Development of the Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in randomized controlled trials and meta-analyses

Table 1: Comparison of the core questions of the 2 versions of the Instrument for assessing the Credibility of Effect Modification Analyses

Core question	Version; question no.*	
	Randomized controlled trials	Meta-analyses
Is the analysis of effect modification based on comparison within rather than between trials?	-	1
For within-trial comparisons, is the effect modification similar from trial to trial?	-	2
For between-trial comparisons, is the number of trials large?	-	3
Was the direction of effect modification correctly hypothesized a priori?	1	4
Was the effect modification supported by prior evidence?	2	-
Does a test for interaction suggest that chance is an unlikely explanation of the apparent effect modification?	3	5
Did the authors test only a small number of effect modifiers or consider the number in their statistical analysis?	4	6
Did the authors use a random-effects model?	-	7
If the effect modifier is a continuous variable, were arbitrary cut points avoided?	5	8

NA = not applicable.

*Numbers reflect order of appearance within the full instrument (see Appendices 5 and 6).

Méta-analyse sur données individuelles

- Réanalyse homogène de tous les essais (outcomes, données manquantes, follow-up...)
- Plus grande puissance modificateur d'effets
- Analyses de RCT non publiés
- Sous groupes spécifiques
- Ajustement sur facteurs pronostics

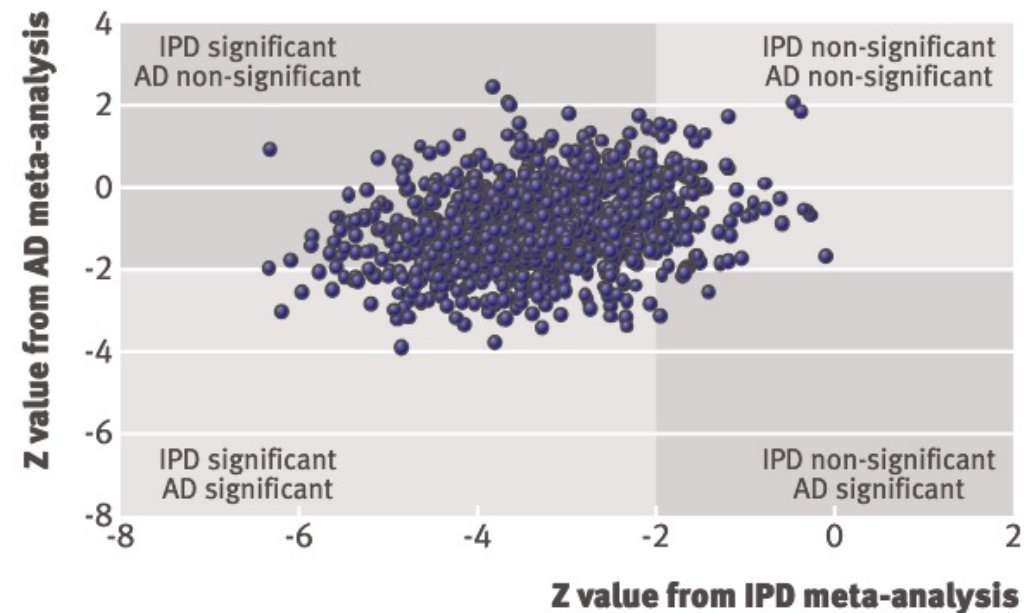
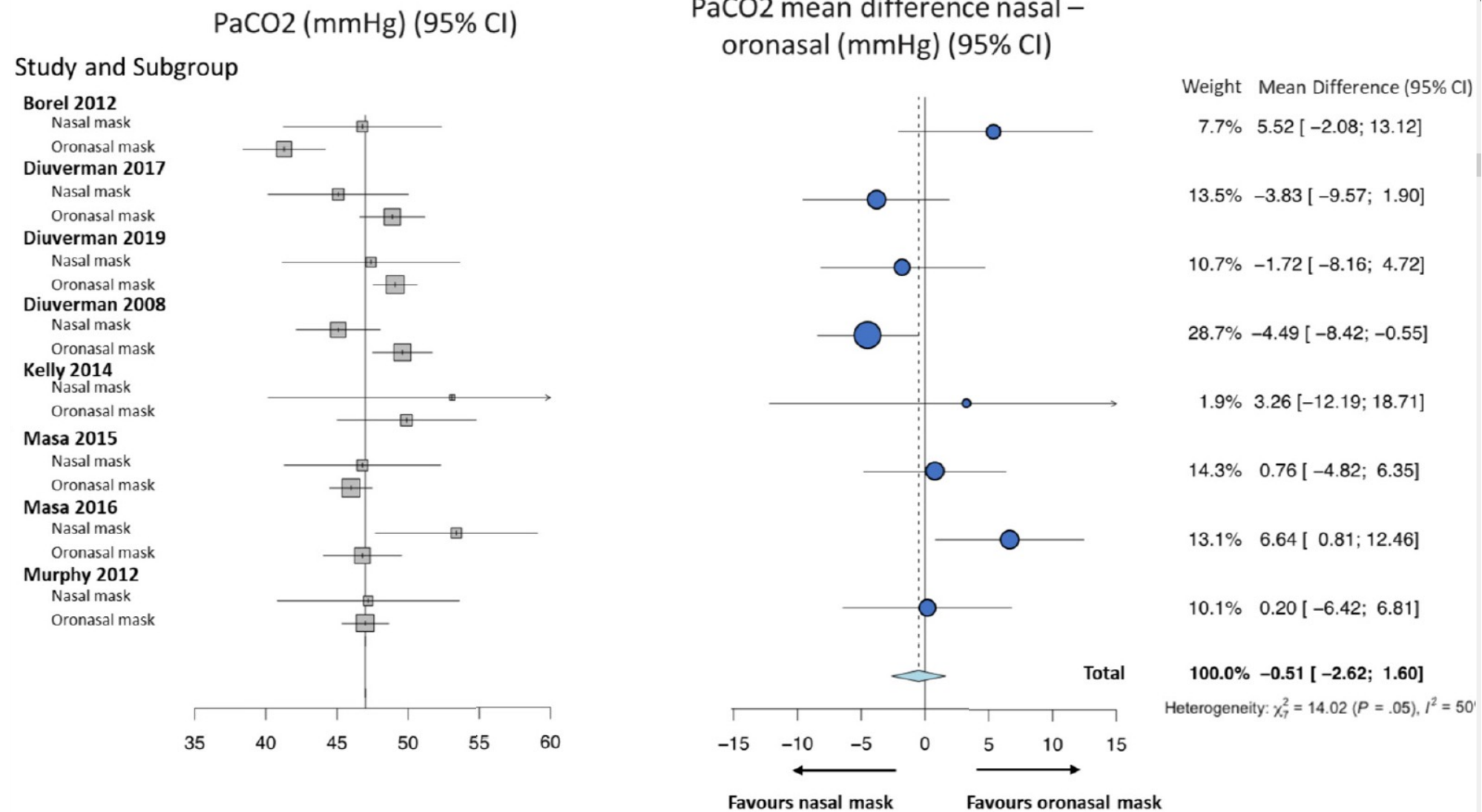


Fig 2 | Comparison of the power of meta-analyses to detect a differential treatment effect across two groups of patients when individual participant data (IPD) or aggregate data (AD) are used. Adapted from Lambert et al¹⁹

Nasal versus oronasal masks for home non-invasive ventilation in patients with chronic hypercapnia: a systematic review and individual participant data meta-analysis



Méta-analyse sur données individuelles

- Parfois très très long...
- Biais de sélection
 - méthodes pour combiner IPD et données agrégées RCT
- Nécessité souvent d'inclure les PI des RCT

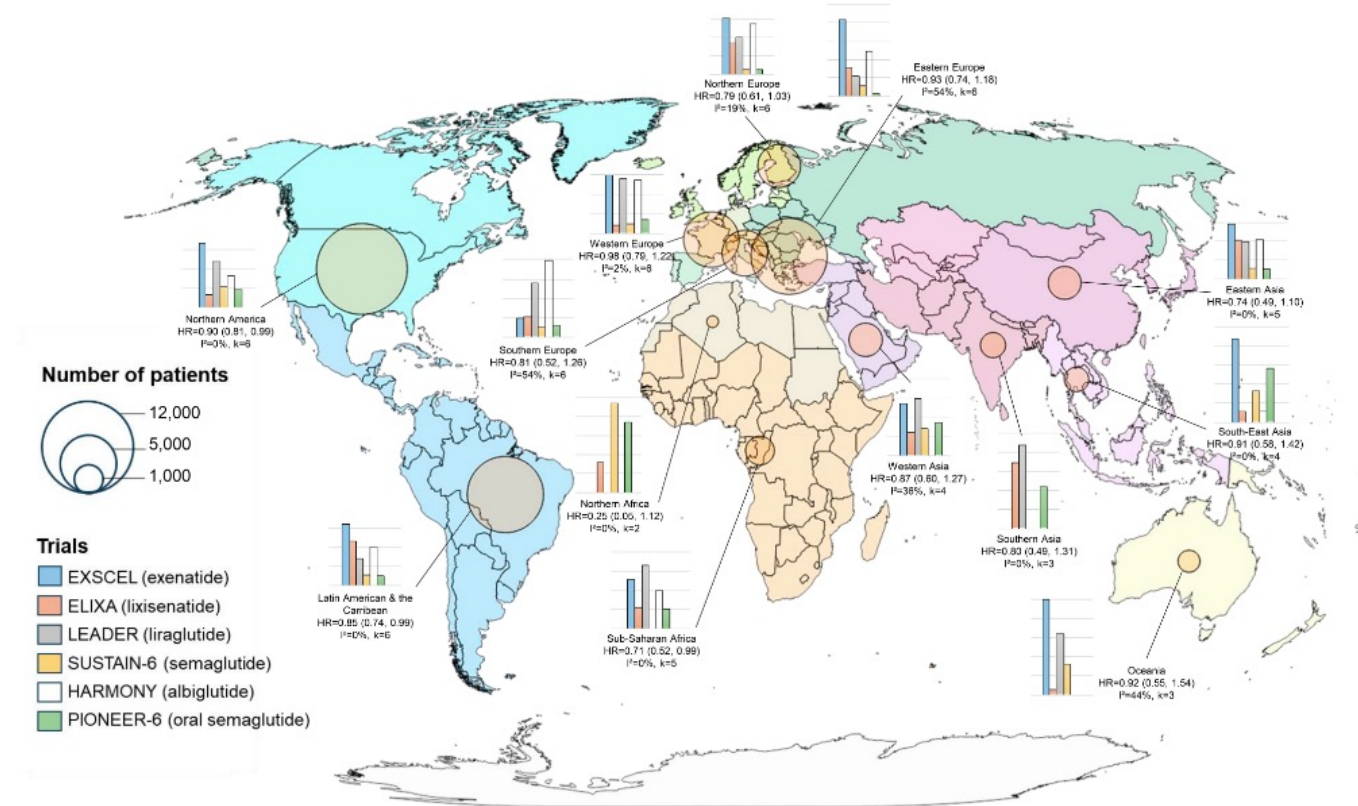
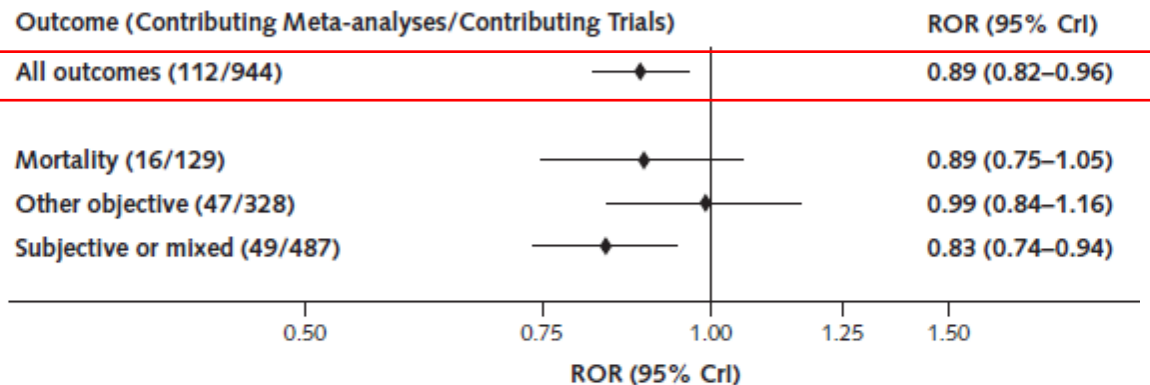


Figure 2—Effect of GLP-1RAs on MACE across the various UN regions. The size of the circles is proportional to the number of patients included per region. Histograms represent the proportion of patients from each trial in each region. Efficacy is expressed as the HR with its 95% CI. Heterogeneity within each region is expressed with the I^2 statistic, and k represents the number of trials per region included in the meta-analyses.

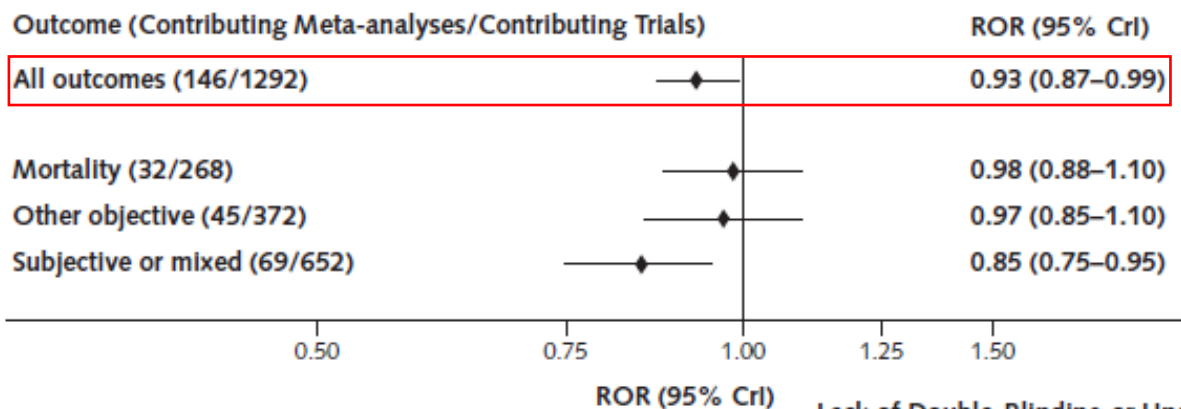
Interaction test p-values: $p=0.70$ for HR and $p=0.51$ for ARD

Biais dans les études incluses

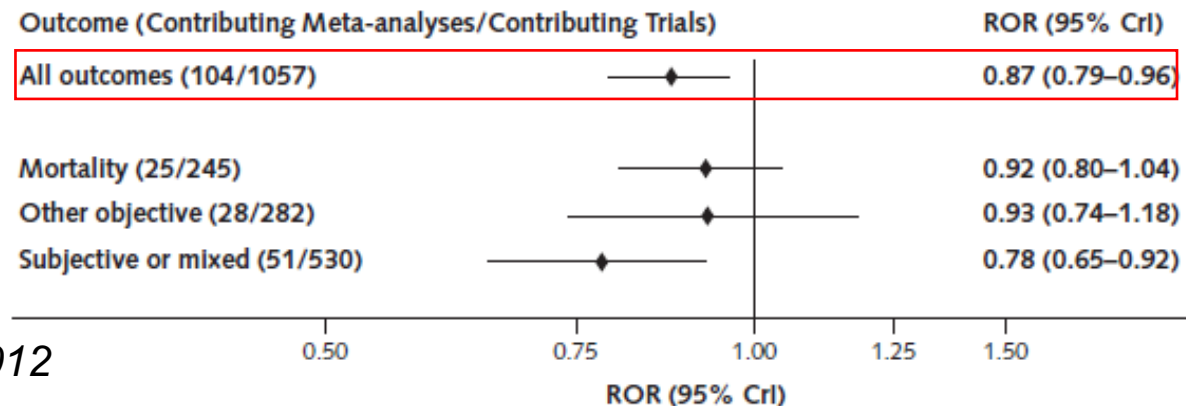
Inadequate or Unclear Generation of Randomization Sequence (vs. Adequate)



Inadequate or Unclear Allocation Concealment (vs. Adequate)



Lack of Double-Blinding or Unclear Double-Blinding (vs. Double-Blind)



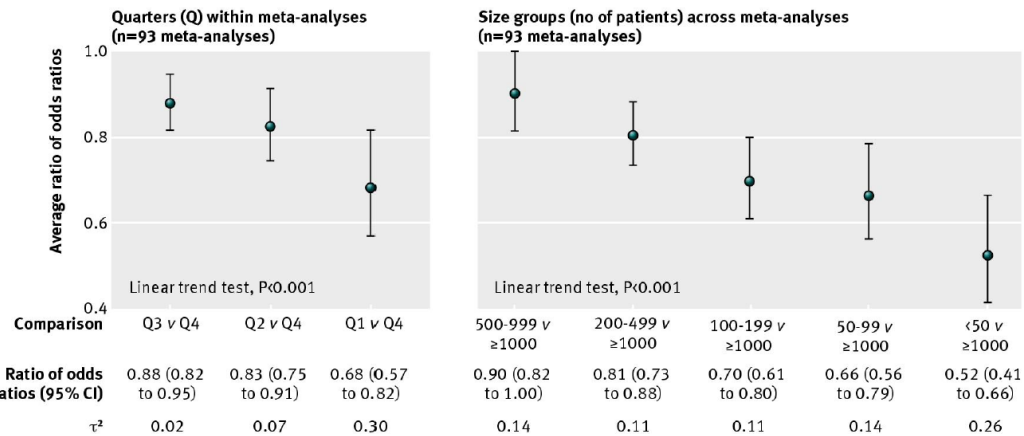
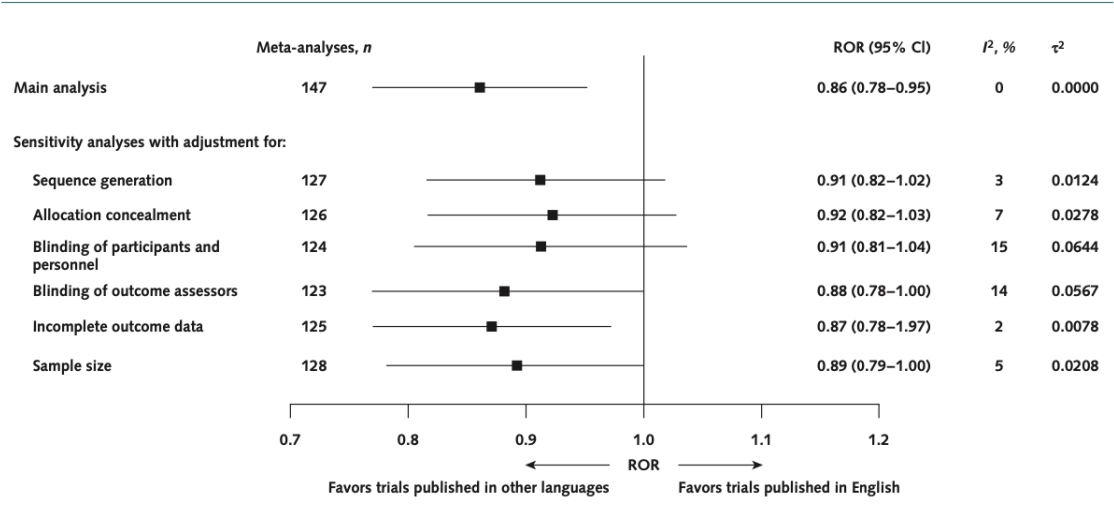


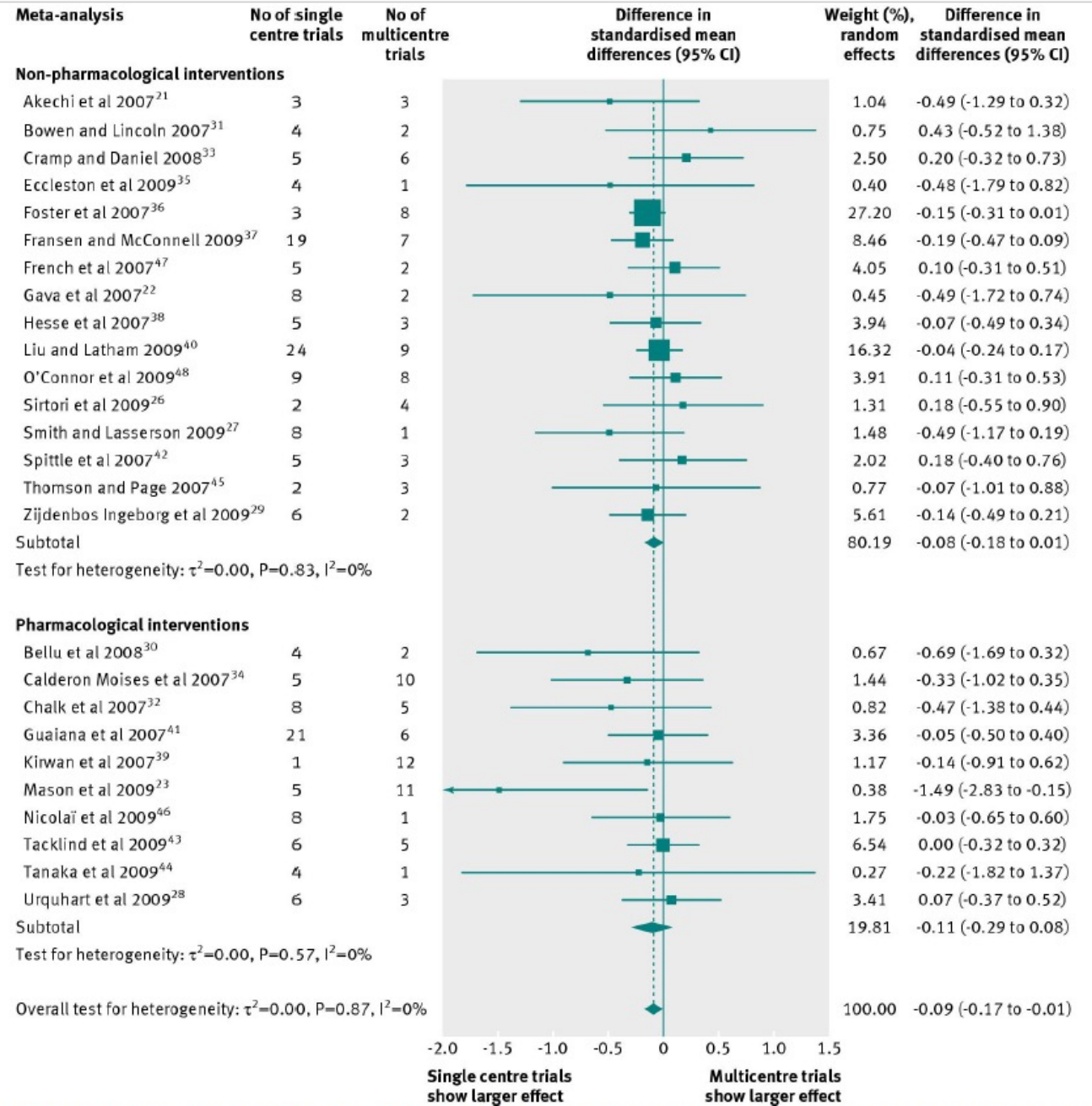
Fig 1 Comparison of treatment effect estimates between trial sample sizes, grouped by quarters (from quarter 1 with the smallest trials, to quarter 4 with the largest trials) and by size groups (from <50 patients to ≥1000 patients). Both analyses included all 93 individual meta-analyses (735 randomised controlled trials)

Dechartres, BMJ 2013

Figure 2. Comparison of treatment effects between trials published in a language other than English and English in 147 meta-analyses from Cochrane reviews (unadjusted and adjusted for risk of bias and sample size).



Dechartres, Ann Intern Med 2018



Bafeta, BMJ 2012

PROBAST: A Tool to Assess the Risk of Bias and Applicability Prediction Model Studies

Robert F. Wolff, MD*; Karel G.M. Moons, PhD*; Richard D. Riley, PhD; Penny F. Whiting, PhD; Marie Westw...

Clinical prediction models combine multiple predictors to estimate risk for the presence of a particular condition (diagnostic models) or the occurrence of a certain event in the future (prognostic models).

RESEARCH METHODS AND REPORTING

Check for updates

Check for updates

RESEARCH METHODS AND REPORTING

ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions

Jonathan AC Sterne, Miguel A Hernan, Barnaby C Reeves, Jelena Savovic, Nancy D Berkman, Meera Viswanathan...

For numbered affiliations see end of article. Correspondence to: J A C Sterne jonathan.sterne@bristol.ac.uk

Additional material is published online only. To view please visit the journal online. Cite this as: BMJ 2016;355:i4999

http://dx.doi.org/10.1136/bmj.i4999

Non-randomised studies of the effects of interventions are critical to many areas of health care evaluation, but their results may be biased.

Non-randomised studies of the effects of interventions (NRSI) are critical to many areas of health care evaluation. Designs of NRSI that can be used to evaluate the effects of interventions include observational studies

SUMMARY POINTS

- Non-randomised studies of the effects of interventions are critical to many areas of healthcare evaluation but are subject to confounding and a range of other potential biases. We developed, piloted, and refined a new tool, ROBINS-I, to assess "Risk Of Bias In Non-randomised Studies - of Interventions". The tool views each study as an attempt to emulate (mimic) a hypothetical

RoB 2: a revised tool for assessing risk of bias in randomised trials

Jonathan A C Sterne, Jelena Savovic, Matthew J Page, Roy G Elbers, Natalie S Blencowe, Isabelle Boutron, Christopher J Cates, Hung-Yuan Cheng, Mark S Corbett, Sandra M Eldridge, Jonathan R Emberson, Miguel A Hernan, Sally Hopewell, Asbjorn Hrobjartsson, Daniela R Junqueira, Peter Juni, Jamie J Kirkham, Toby Lasserson, Tianjing Li, Alexandra McAleenan, Barnaby C Reeves, Sasha Shepperd, Ian Shrier, Lesley A Stewart, Kate Tilling, Ian R White, Penny F Whiting, Julian PT Higgins

Hooijmans et al. BMC Medical Research Methodology 2014, 14:43 http://www.biomedcentral.com/1471-2288/14/43

RESEARCH METHODS AND REPORTING

RESEARCH ARTICLE

Open Access

SYRCLE's risk of bias tool for animal studies

Carlijn R Hooijmans, Maroeska M Rovers, Rob BM de Vries, Marlies Leenaars, Merel Ritskes-Hoitinga, and Miranda W Langendam

Abstract Background: Systematic Reviews (SRs) of the merits of conducting such SRs is steeply clinical trials (RCT) in many aspects, the met for animal intervention studies. The Cochran consistency and avoid discrepancies in asse in the field of animal experimentation.

Methods: We provide an RoB tool for animal Cochrane RoB tool and has been adjusted to enhance transparency and applicability.

Results: The resulting RoB tool for animal s performance bias, detection bias, attrition b with the items in the Cochrane RoB tool. M design between RCTs and animal studies. S design of animal studies compared to clin

Conclusions: SYRCLE's RoB tool is an adapt implementation of this tool will facilitate an subsequently enhance the efficiency of tran the necessity of improving the methodolog

Keywords: Risk of bias, Methodological qua

Background: The use of systematic reviews (SRs) for dence-based decisions on healthcare is com in the clinical setting. Although most exper studies aim to test safety and or efficacy of t be used for human healthcare, summarizing evidence in an SR is far less common in laboratory animal experiments. Fortunately, s ential commentary was published in the Lanc first setting out the scientific rationale for S studies, awareness of the merits of SRs of

Development: We followed th of the origina for assessing dies of inter development collaborators, and adminis Preliminary original tool review and 1 studies of en

essential component of a n; the most commonly used tool he Cochrane risk-of-bias tool,

l-bias tool were identified on nce and feedback, approaches ipments in estimation of

essing risk of bias in

in each domain, users of RoB 2 nswers lead to judgments of of bias"

verall risk-of-bias judgment for ivers of RoB 2 to stratify meta-

such as cohort studies and case-control studies in which intervention groups are allocated during the course of usual treatment decisions, and quasi-randomised studies in which the method of allocation falls short of full randomisation. Non-randomised studies can provide evidence additional to that available from randomised trials about long term outcomes, rare events, adverse effects and populations that are typical of real world practice.

Many tools to assess the methodological quality of observational studies in the context of a systematic review have been proposed. The Newcastle-Ottawa and Downs-Black tools have been two of the most popular: both were on a shortlist of methodologically sound tools, but each includes items relating to external as well as internal validity and a lack of comprehensive manuals means that instructions may be interpreted differently by different users.

In the past decade, major developments have been made in tools to assess study validity. A shift in focus from methodological quality to risk of bias has been accompanied by a move from checklists and numeric scores towards domain-based assessments in which different types of bias are considered in turn. Examples are the Cochrane Risk of Bias tool for randomised trials, the QUADAS 2 tool for diagnostic test accuracy

risk of bias is regarded component of a w on the effects of an e most commonly used ised trials is the f-bias tool. We updated ond to developments in how bias arises in ils, and to address user d limitations of the

the effect of ir large random not well defini (such as perfor not inherently RoB tool consi a trial (known the basis of t consideration supported by (eg, trial proto by justification) After nearly tool, potentia A formal eva be confusing to incomplete of outcomes confusion ove should auton risk of bias, of-bias assess conclusions is and user prac Cochrane syst RoB tool in noi as at low risk - bias are comm moderate reli We develop tool to addre in assessment developed tool in estimation trials.

After nearly tool, potentia A formal eva be confusing to incomplete of outcomes confusion ove should auton risk of bias, of-bias assess conclusions is and user prac Cochrane syst RoB tool in noi as at low risk - bias are comm moderate reli We develop tool to addre in assessment developed tool in estimation trials.



ROBIS: A new tool to assess risk of bias in systematic reviews was developed

Penny Whiting, Jelena Savovic, Julian P.T. Higgins, Deborah M. Caldwell, Barnaby C. Reeves, Beverley Shea, Philippa Davies, Jos Kleijnen, Rachel Churchill, the ROBIS group

School of Social and Community Medicine, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol BS8 2PS, UK The National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care West at University Hospitals Bristol NHS Foundation Trust, 9th Floor Whatley Road, Bristol BS1 3NF



RESEARCH ARTICLE

Open Access

SYRCLE's risk of bias tool for animal studies

Carlijn R Hooijmans, Maroeska M Rovers, Rob BM de Vries, Marlies Leenaars, Merel Ritskes-Hoitinga, and Miranda W Langendam

Abstract Background: Systematic Reviews (SRs) of the merits of conducting such SRs is steeply clinical trials (RCT) in many aspects, the met for animal intervention studies. The Cochran consistency and avoid discrepancies in asse in the field of animal experimentation.

Methods: We provide an RoB tool for animal Cochrane RoB tool and has been adjusted to enhance transparency and applicability.

Results: The resulting RoB tool for animal s performance bias, detection bias, attrition b with the items in the Cochrane RoB tool. M design between RCTs and animal studies. S design of animal studies compared to clin

Conclusions: SYRCLE's RoB tool is an adapt implementation of this tool will facilitate an subsequently enhance the efficiency of tran the necessity of improving the methodolog

Keywords: Risk of bias, Methodological qua

Background: The use of systematic reviews (SRs) for dence-based decisions on healthcare is com in the clinical setting. Although most exper studies aim to test safety and or efficacy of t be used for human healthcare, summarizing evidence in an SR is far less common in laboratory animal experiments. Fortunately, s ential commentary was published in the Lanc first setting out the scientific rationale for S studies, awareness of the merits of SRs of

essential component of a n; the most commonly used tool he Cochrane risk-of-bias tool,

l-bias tool were identified on nce and feedback, approaches ipments in estimation of

essing risk of bias in

in each domain, users of RoB 2 nswers lead to judgments of of bias"

verall risk-of-bias judgment for ivers of RoB 2 to stratify meta-

such as cohort studies and case-control studies in which intervention groups are allocated during the course of usual treatment decisions, and quasi-randomised studies in which the method of allocation falls short of full randomisation. Non-randomised studies can provide evidence additional to that available from randomised trials about long term outcomes, rare events, adverse effects and populations that are typical of real world practice.

Many tools to assess the methodological quality of observational studies in the context of a systematic review have been proposed. The Newcastle-Ottawa and Downs-Black tools have been two of the most popular: both were on a shortlist of methodologically sound tools, but each includes items relating to external as well as internal validity and a lack of comprehensive manuals means that instructions may be interpreted differently by different users.

In the past decade, major developments have been made in tools to assess study validity. A shift in focus from methodological quality to risk of bias has been accompanied by a move from checklists and numeric scores towards domain-based assessments in which different types of bias are considered in turn. Examples are the Cochrane Risk of Bias tool for randomised trials, the QUADAS 2 tool for diagnostic test accuracy

risk of bias is regarded component of a w on the effects of an e most commonly used ised trials is the f-bias tool. We updated ond to developments in how bias arises in ils, and to address user d limitations of the

the effect of ir large random not well defini (such as perfor not inherently RoB tool consi a trial (known the basis of t consideration supported by (eg, trial proto by justification) After nearly tool, potentia A formal eva be confusing to incomplete of outcomes confusion ove should auton risk of bias, of-bias assess conclusions is and user prac Cochrane syst RoB tool in noi as at low risk - bias are comm moderate reli We develop tool to addre in assessment developed tool in estimation trials.

After nearly tool, potentia A formal eva be confusing to incomplete of outcomes confusion ove should auton risk of bias, of-bias assess conclusions is and user prac Cochrane syst RoB tool in noi as at low risk - bias are comm moderate reli We develop tool to addre in assessment developed tool in estimation trials.

After nearly tool, potentia A formal eva be confusing to incomplete of outcomes confusion ove should auton risk of bias, of-bias assess conclusions is and user prac Cochrane syst RoB tool in noi as at low risk - bias are comm moderate reli We develop tool to addre in assessment developed tool in estimation trials.

ber than in primary studies). v the evidence base, hold a face-to- settings: interventions, diagnosis, reviews of systematic reviews ("revs. The tool is completed in three ik of bias. Phase 2 covers four do- ation and selection of studies; data se interpretation of review findings re included to help judge concerns

QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies

Penny F. Whiting, PhD; Anne W.S. Rutjes, PhD; Marie E. Westwood, PhD; Susan Mallett, PhD; Jonathan J. Deeks, PhD; Johannes B. Reitsma, MD, PhD; Mariska M.G. Leeflang, PhD; Jonathan A.C. Sterne, PhD; Patrick M.M. Bossuyt, PhD; and the QUADAS-2 Group*

In 2003, the QUADAS tool for systematic reviews of diagnostic accuracy studies was developed. Experience, anecdotal reports, and feedback suggested areas for improvement; therefore, QUADAS-2 was developed. This tool comprises 4 domains: patient selection, index test, reference standard, and flow and timing. Each domain is assessed in terms of risk of bias, and the first 3 domains are also assessed in terms of concerns regarding applicability. Signalling questions are included to help judge risk of bias.

The QUADAS-2 tool is applied in 4 phases: summarize the review question, tailor the tool and produce review-specific guidance, construct a flow diagram for the primary study, and judge bias and applicability. This tool will allow for more transparent rating of bias and applicability of primary diagnostic accuracy studies.

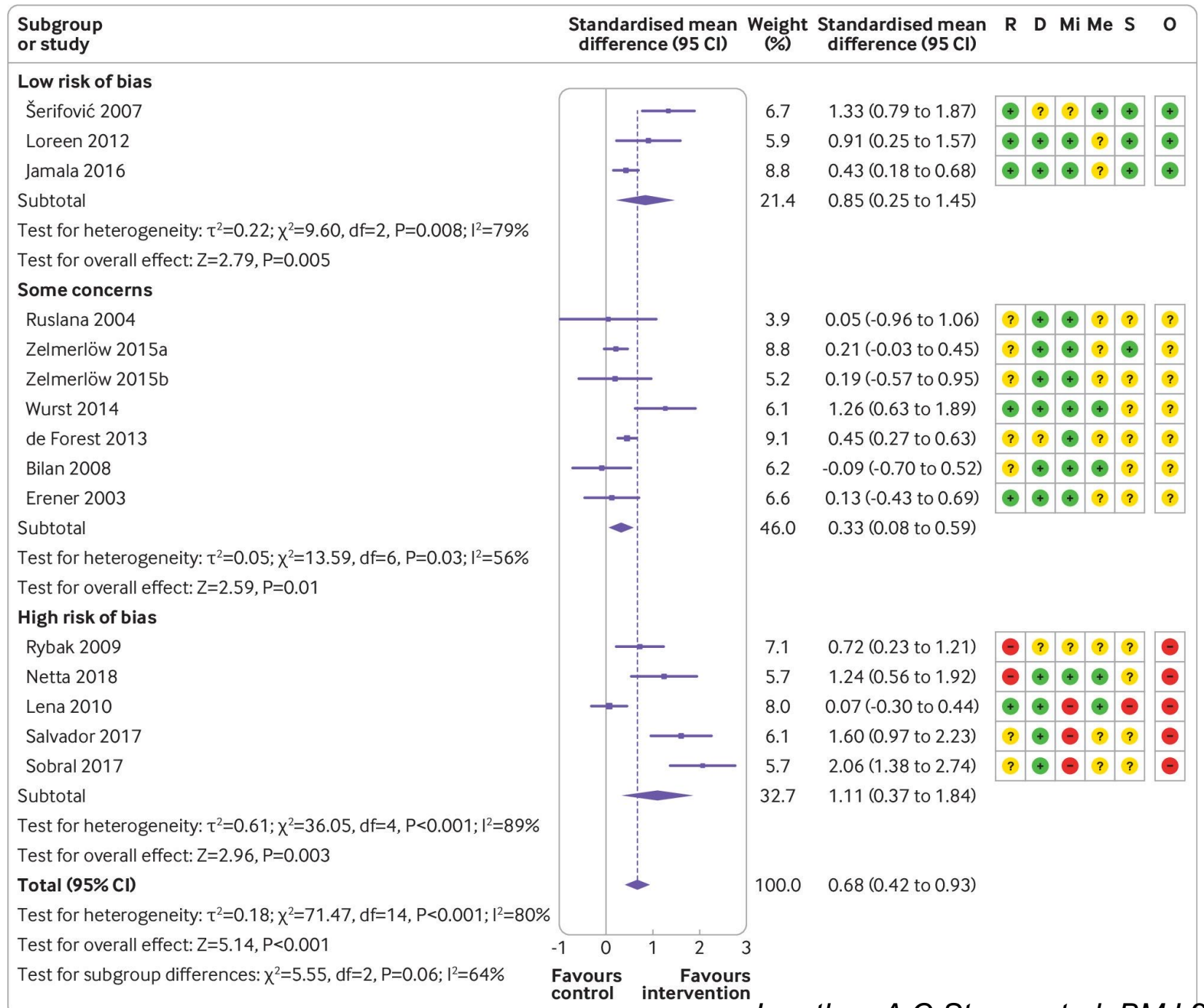
Ann Intern Med 2011;155:529-536. For author affiliations, see end of text. * For members of the QUADAS-2 Group, see the Appendix (available at www.annals.org).

Define the Scope

We established a steering group of 9 experts in the area of diagnostic research, most of whom participated in developing the original QUADAS tool. This group agreed on key features of the desired scope of QUADAS-2. The main decision was to separate "quality" into "risk of bias" and "concerns regarding applicability." We defined quality as "both the risk of bias and applicability of a study; 1) the degree to which estimates of diagnostic accuracy avoided risk of bias, and 2) the extent to which primary studies are applicable to the reviewer's research question." Bias occurs if systematic flaws or limitations in the design or conduct of a study distort the results. Evidence from a primary study may have limited applicability to the review if, compared with the review question, the study was conducted in a patient group with different demographic or clinical features, the index test was applied or interpreted differently, or the definition of the target condition differed.

Other decisions included limiting QUADAS-2 to a

Example forest plot showing results of a risk-of-bias assessment in a systematic review of randomised trials, using version 2 of the Cochrane risk-of-bias tool.



Risk of bias legend

- R** Bias arising from the randomisation process
- D** Bias due to deviations from intended interventions
- Mi** Bias due to missing outcome data
- Me** Bias in measurement of the outcome
- S** Bias in selection of the reported result
- O** Overall risk of bias

Méthodes d'ajustement sur les biais

1. Analyses hiérarchiques/modèles mixtes en prenant en compte le design
2. Pondération négative des études biaisées (et/ou observationnelles)
 - Inflation de la variance
 - Facteur de biais

Synthesizing cross-design evidence and cross-format data using network meta-regression

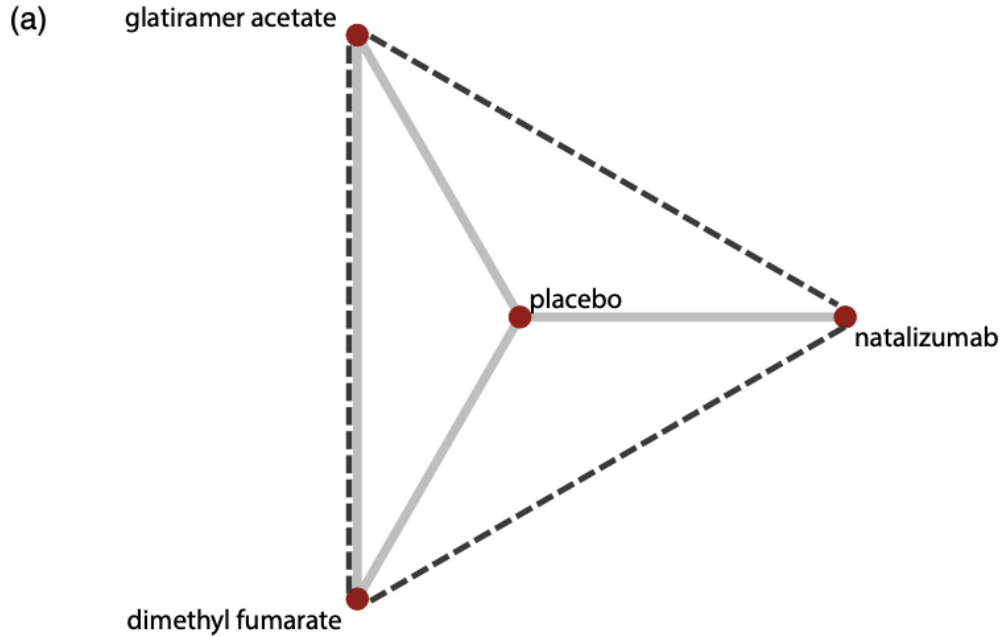


FIGURE 1 Network plots of (a) treatments for patients with relapsing–remitting multiple sclerosis compared in randomised controlled trials (solid, grey edges) and in the Swiss Multiple Sclerosis Cohort (dashed, black edges). The outcome is relapse in 2 years

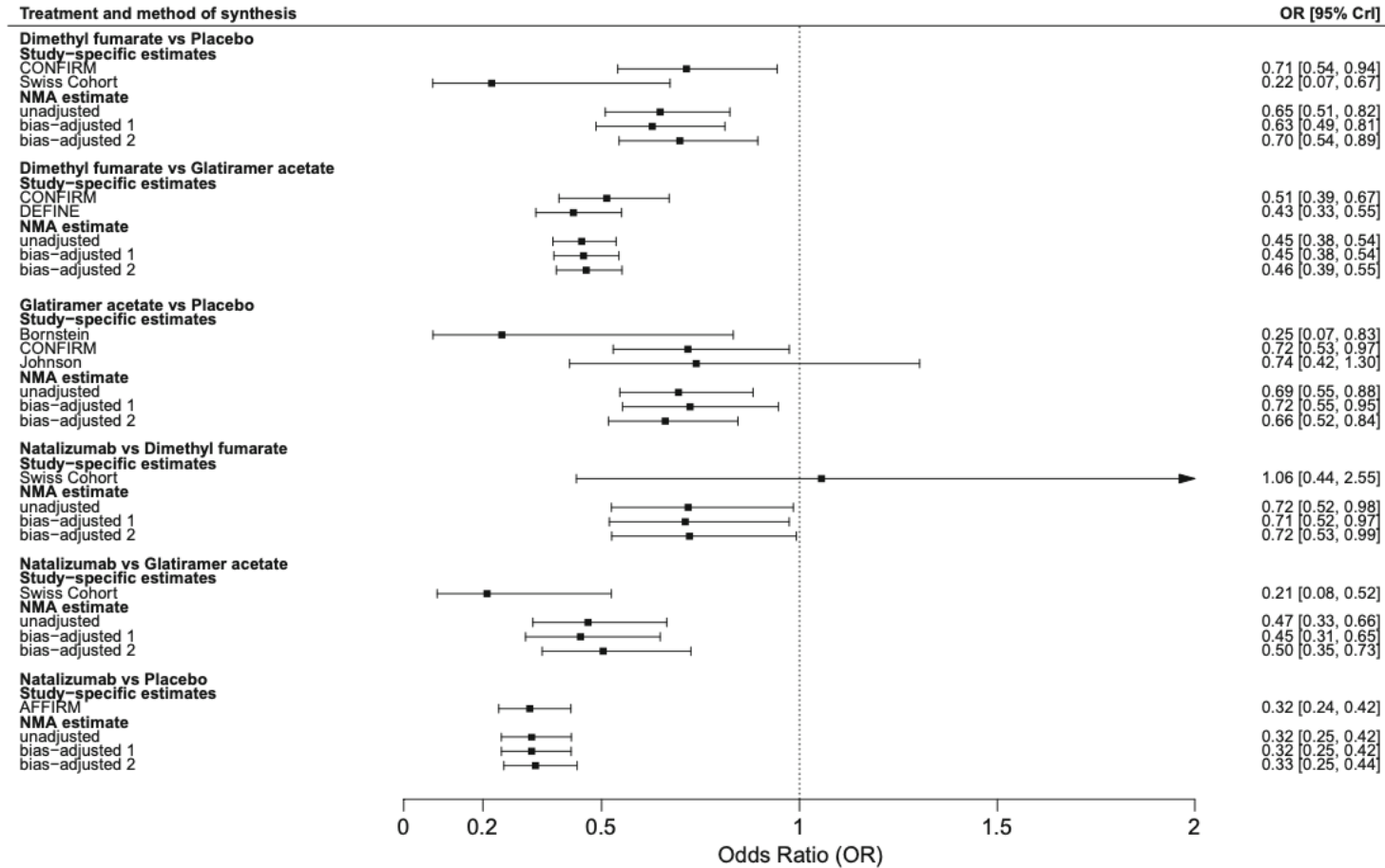


FIGURE 2 Relapse odds ratios with 95% credible intervals (CrI) of all comparisons of treatments among patients with relapsing–remitting multiple sclerosis. The estimates are computed by conducting unadjusted analysis and bias-adjusted analyses 1 and 2 in a Bayesian framework of the data in the network of Figure 1a. The study-specific estimates have been computed in a frequentist framework and hence the lines represent confidence intervals. To compute these estimate, we used *glm()* function to analyze IPD studies and *metabin()* function (from *meta* package) to analyze AD studies

Les méta-analyses inutiles

+ de 300,000 méta-analyses réalisées en santé

THE
MILBANK QUARTERLY
A MULTIDISCIPLINARY JOURNAL OF POPULATION HEALTH AND HEALTH POLICY

Original Investigation

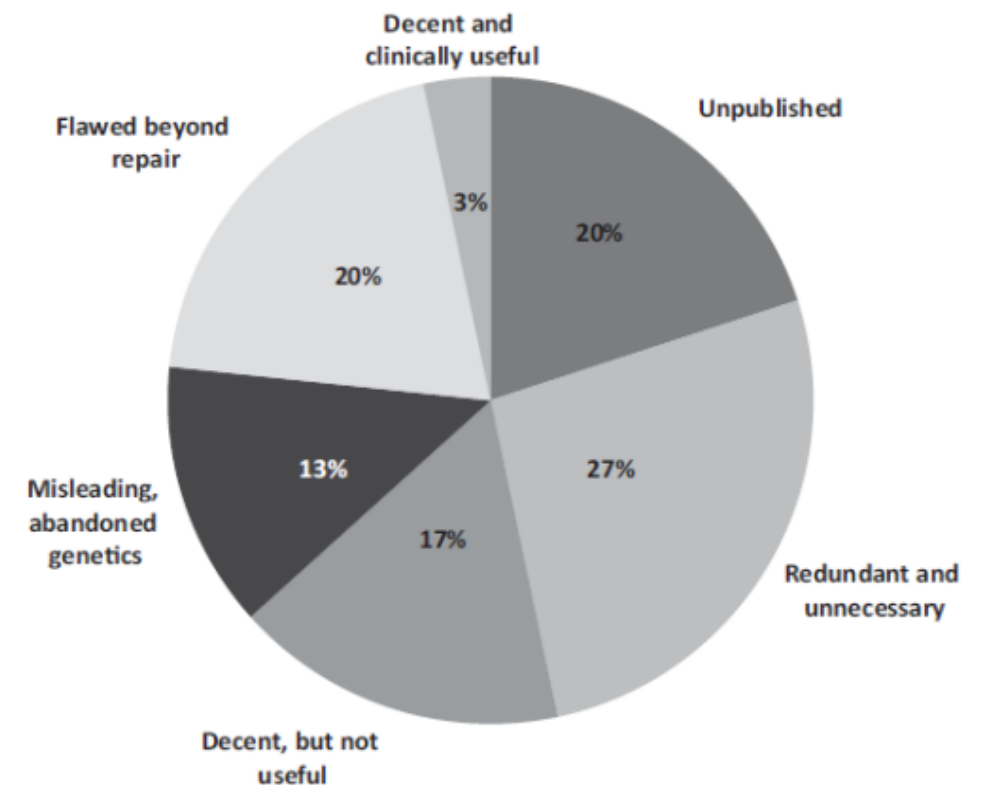
The Mass Production of Redundant,
Misleading, and Conflicted Systematic
Reviews and Meta-analyses

JOHN P.A. IOANNIDIS

*Stanford University School of Medicine; Stanford University School of
Humanities and Sciences; Meta-Research Innovation Center at Stanford
(METRICS), Stanford University*

Conclusions: The production of systematic reviews and meta-analyses has reached epidemic proportions. Possibly, the large majority of produced systematic reviews and meta-analyses are unnecessary, misleading, and/or conflicted.

Figure 4. A Summary Overview of Currently Produced Meta-analyses



L'échec des revues systématiques

Différents critères de sélection, dates de recherches, qualité méthodologique..

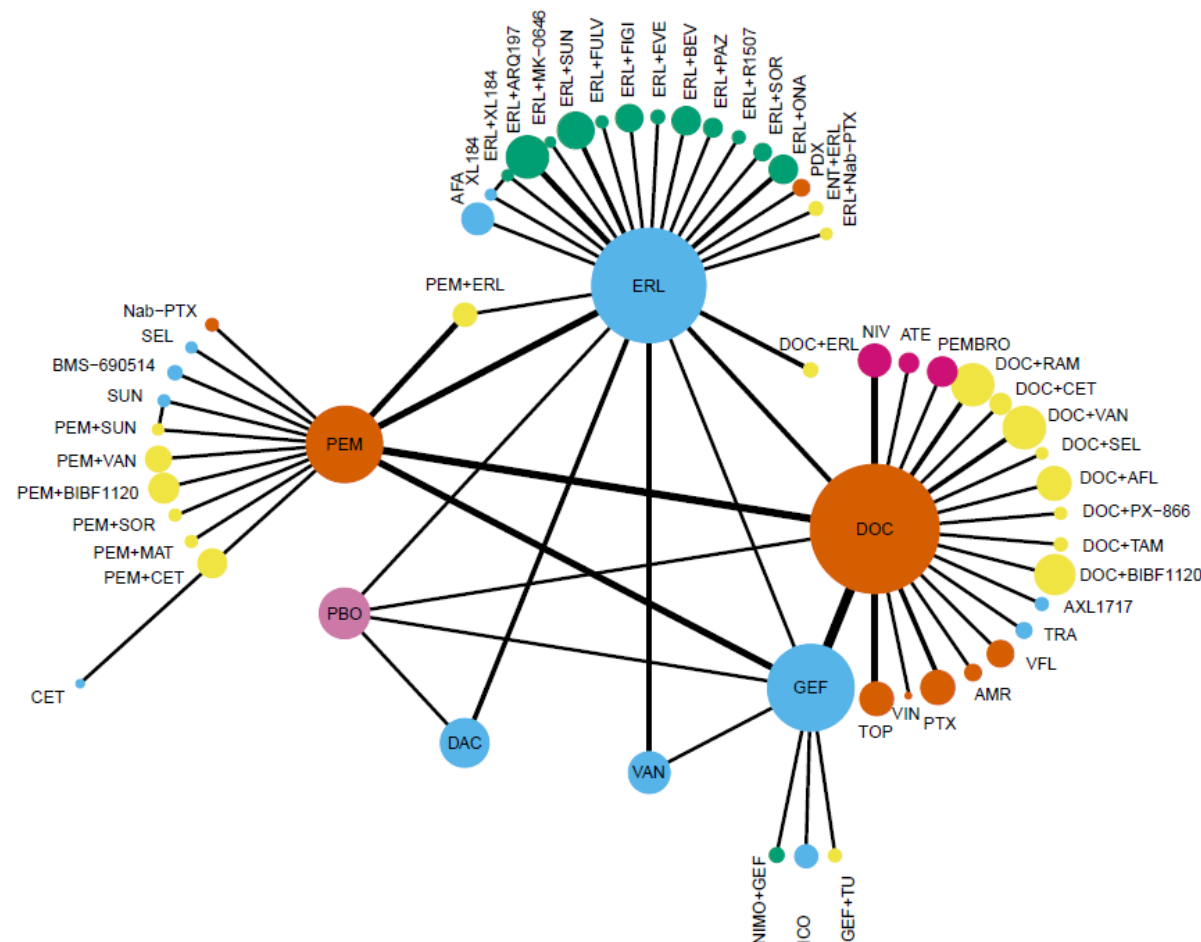
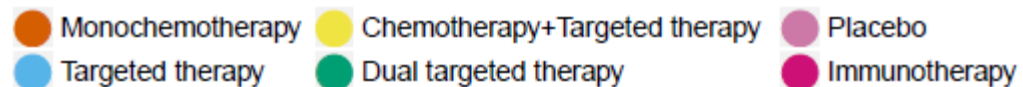
Traitements du cancer du poumon en 2nde ligne

Traitements (n=59)

RCT (n=96)

Patients (n=34 812)

29 Revues systématiques



L'échec des revues systématiques

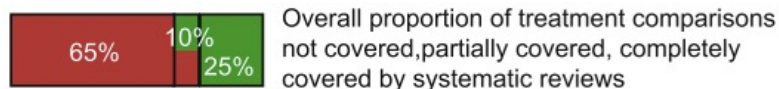
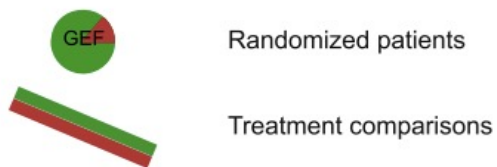
Même collectivement toutes les revues systématiques ne prennent pas en compte tous les RCT disponibles

40% à 66% de traitements manquants

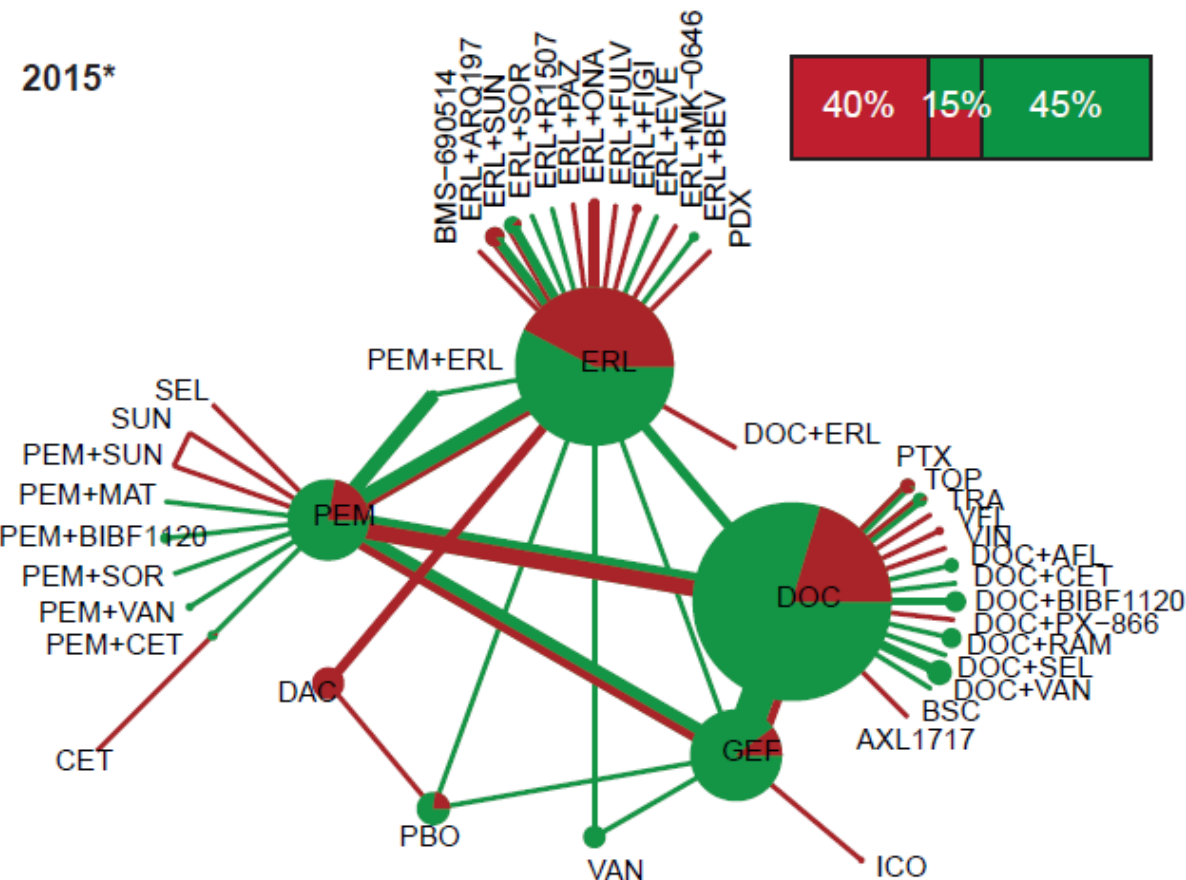
45% à 70% d'essais manquants

30% à 58% de patients manquants

- Not covered by any systematic reviews
- Completely covered by systematic reviews
- Partially covered by systematic reviews



Crequit, BMC Medicine 2016

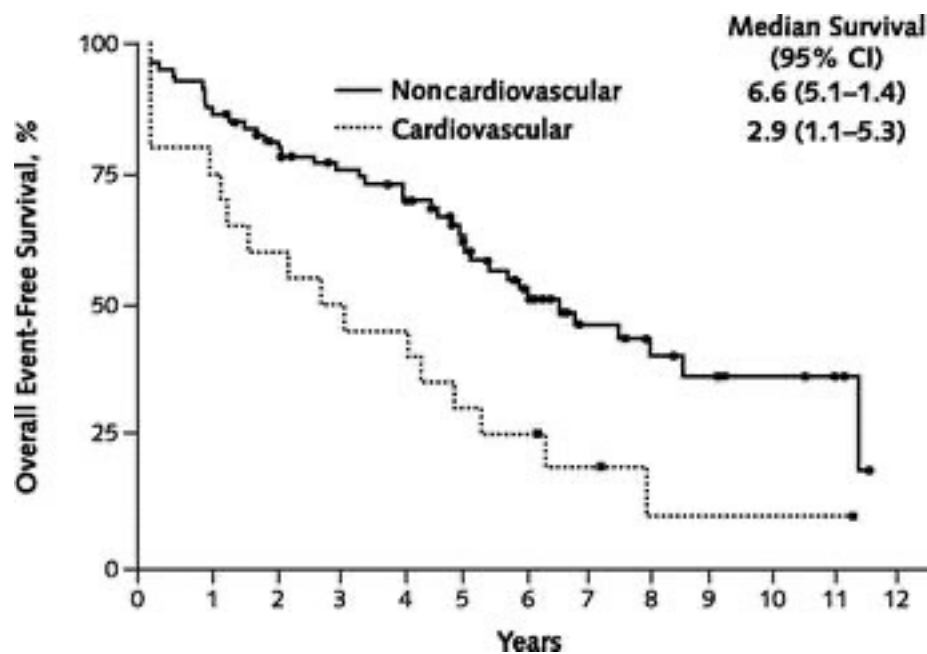
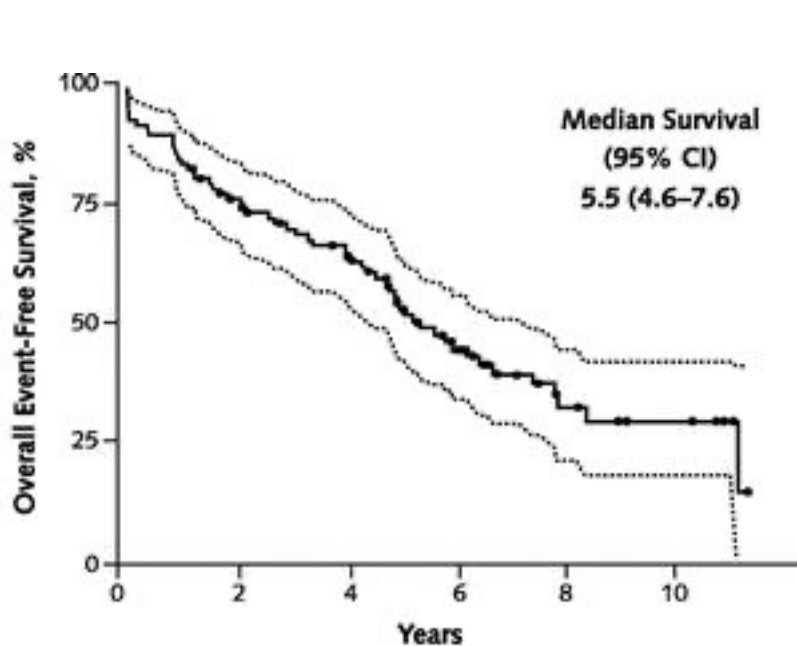


Seventy-Five Trials and Eleven Systematic Reviews a Day: How Will We Ever Keep Up?

Hilda Bastian^{1*}, Paul Glasziou², Iain Chalmers³

¹ German Institute for Quality and Efficiency in Health Care (IQWiG), Cologne, Germany, ² Centre for Research in Evidence-Based Practice, Faculty of Health Sciences, Bond University, Gold Coast, Australia, ³ James Lind Library, James Lind Initiative, Oxford, United Kingdom

How Quickly Do Systematic Reviews Go Out of Date? A Survival Analysis FREE



Systematic reviews at risk, n

100 73 59 34 14 6

An epidemic of redundant meta-analyses

4 RCT

Direct oral anticoagulants with low molecular weight heparins in cancer patients with venous thromboembolism (VTE).

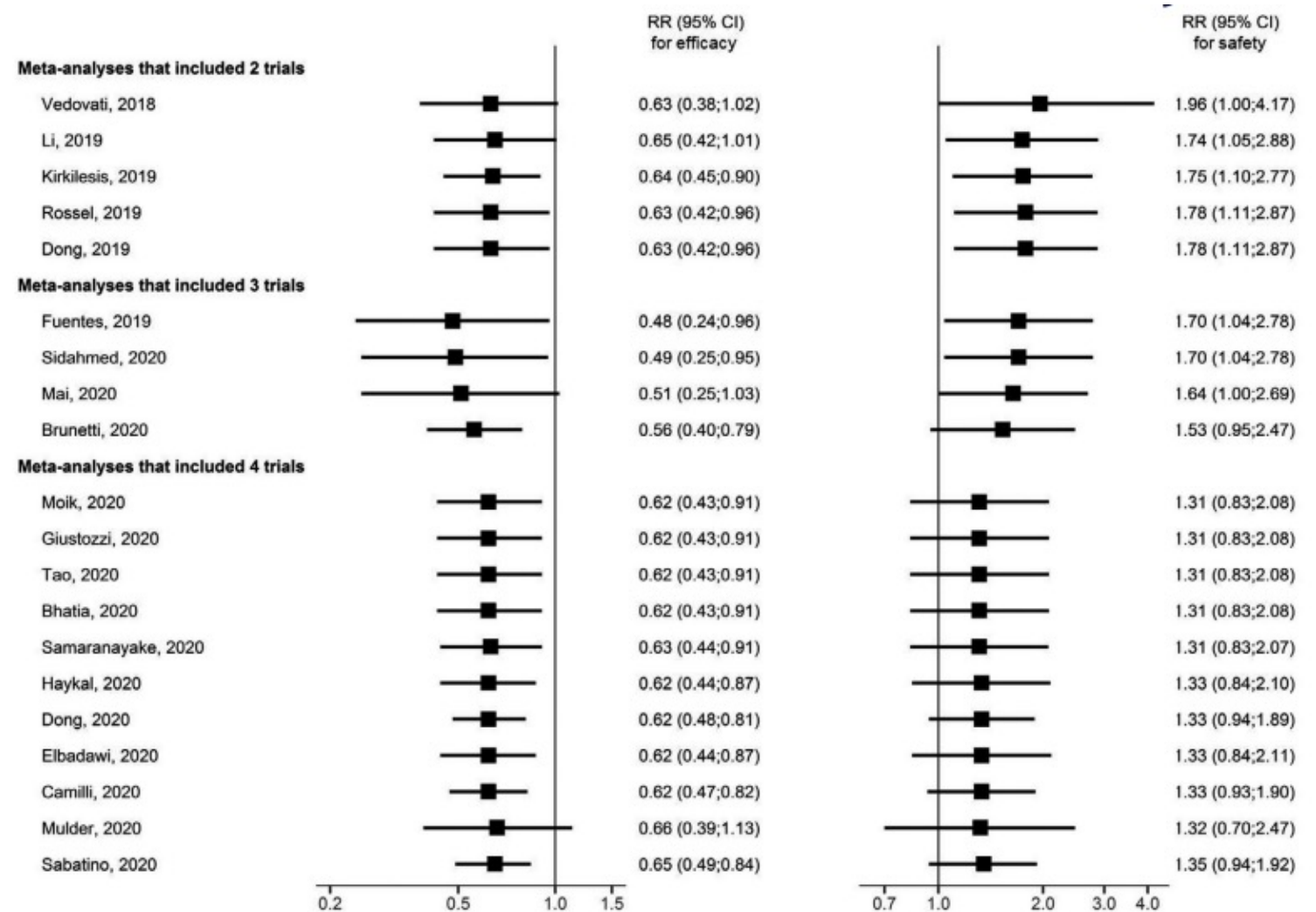


FIGURE 1 Treatment effect estimate by number of studies included in the meta-analyses for efficacy (recurrent venous thromboembolism) on the left and safety (major bleeding) on the right

Vibration de l'effet dans les MA

A meta-epidemiological study found that meta-analyses of the same trials may obtain drastically conflicting results

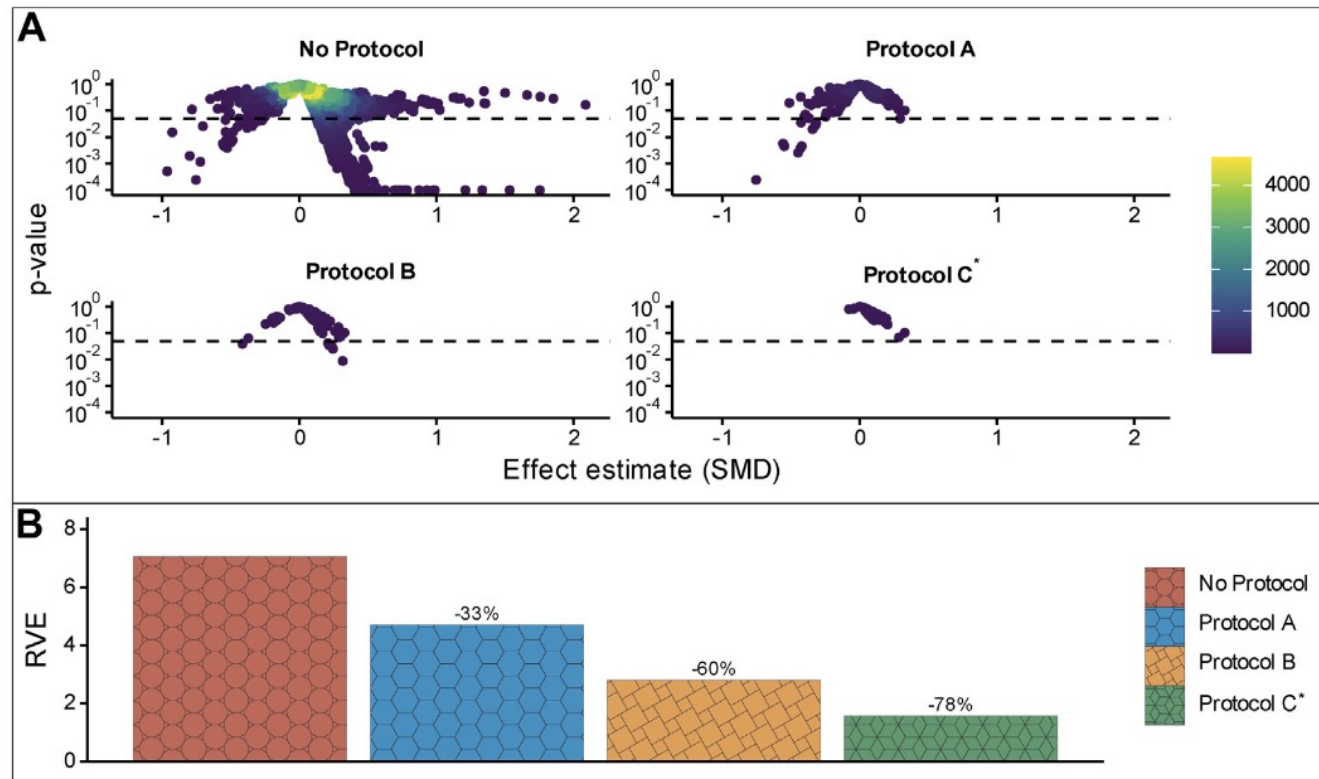


Fig. 2. Distributions of effect estimates and relative vibration of effects for functional outcomes. Panel A: Effect estimates and P values obtained without a protocol and when applying Protocol A, B, and C. A negative effect estimate favors operative treatment, whereas a positive effect estimate favors nonoperative. Effect estimates below the dashed line are statistically significant at the 0.05 level. Panel B: The relative vibration of effects without a protocol and when applying Protocol A, B, and C. *Abbreviations:* SMD, standardized mean difference; RVE, relative vibration of effects. *: Based on PROSPERO registration.

Quelles solutions ?

Si vous avez envie de faire une méta-analyse:

NE LE FAITES PAS 😊

- Revue de la littérature
- Registres prospero...
- Update ?

Garner, BMJ 2016

RESEARCH METHODS AND REPORTING

When and how to update systematic reviews: consensus and checklist

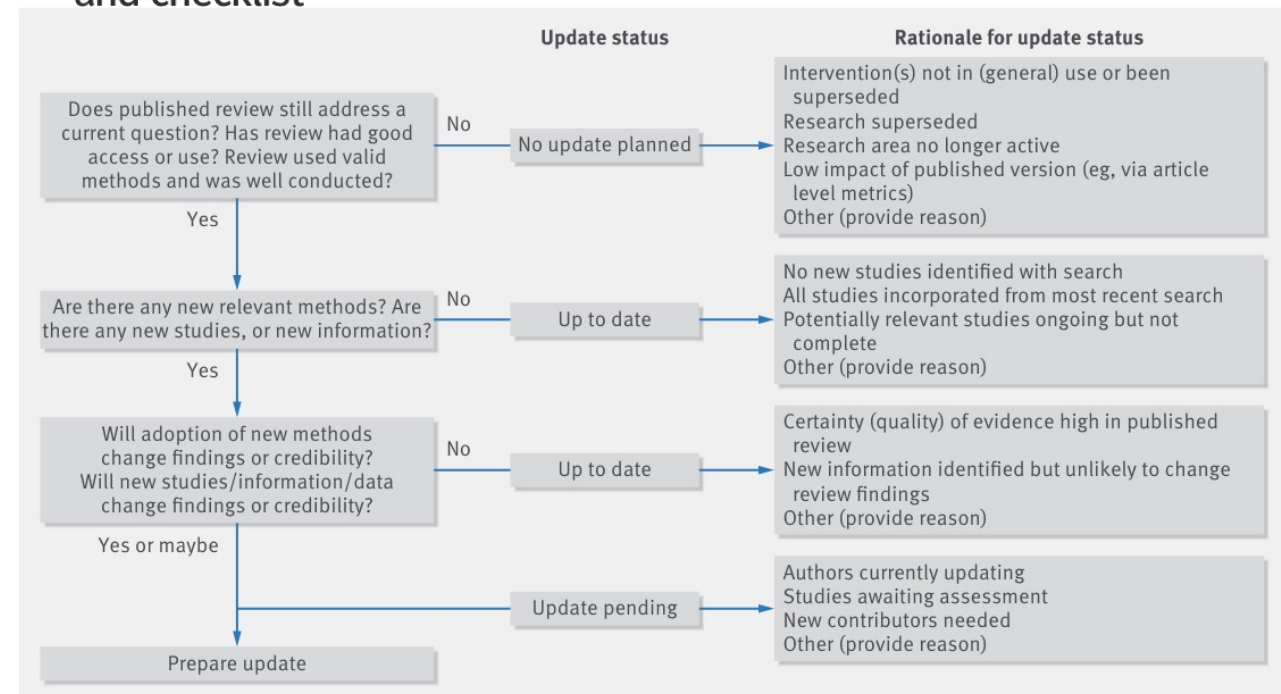
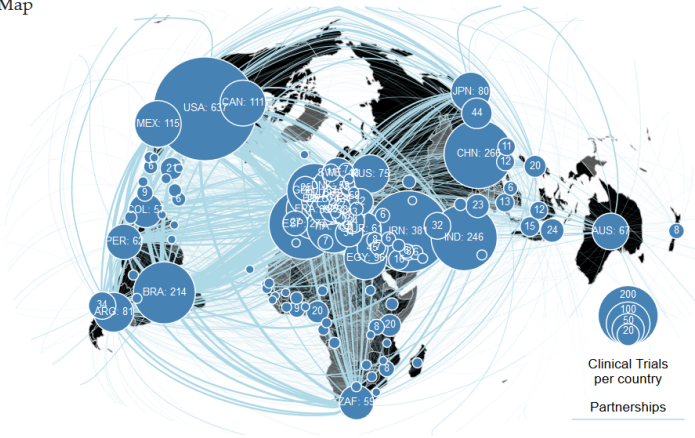


Fig 1 | Decision framework to assess systematic reviews for updating, with standard terms to report such decisions

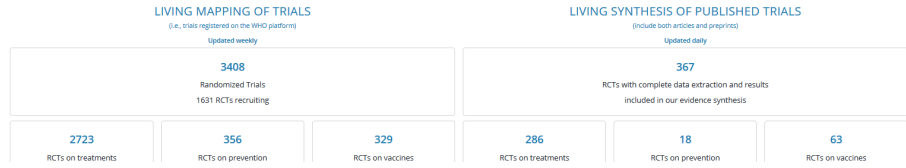
The COVID-NMA Project: Building an Evidence Ecosystem for the COVID-19 Pandemic

▼ Map



The COVID-NMA initiative
A living mapping and living systematic review of Covid-19 trials

COVID-NMA is an international research initiative supported by the WHO and Cochrane.
We provide a living mapping of COVID-19 trials. We are also conducting living evidence synthesis on preventive interventions, treatments and vaccines for COVID-19 to assist decision makers.
See the description of our model [here](#), and our living review protocol [here](#).



metaEvidence - COVID-19

Living meta-analysis and evidence synthesis of therapies for COVID19

results overview latest studies 0 standard queries browse treatments mapping NMA

Meta-analyses vivantes

noteworthy treatments COVID 19 hospitalized all patients Top evidence (RCT only, high risk of bias excluded)

button bars forest bayes pooled summary

Treatment	Trials	deaths	clinical improvement	death or ventilation
corticosteroids	11/22	-8%	+4%	-8%
remdesivir	6/9	-10%	statistically conclusive +30%	-3%
convalescent plasma treatment	16/27	-1%	+3%	-1%
baricitinib	2	-38%	statistically conclusive +16%	-29%
casirivimab and imdevimab (REGN-COV2)	1	statistically conclusive -18%	-	-16%
interferon	10/13	-32%	-	-16%
tocilizumab	12/26	safety concern -12%	statistically conclusive +20%	statistically conclusive -20%
sarilumab	3/6	+9%	+3%	-
anticoagulant, curative dose	3/4	+8%	-16%	-17%
anticoagulation, intermediate prophylactic dose	1	+9%	-	-

● Statistically conclusive: statistically significant with a strict control of overall risk of type 1 error (statistically demonstrated), does not take into account the risk of bias; ● suggested: nominally statistically significant but without a strict control of overall risk of type 1 error; ● inconclusive: not nominally statistically significant

En Fr
Our approach
RSS

Besoin d'automatisation

Annals of Internal Medicine RESEARCH AND REPORTING METHODS

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

Viet-Thi Tran, MD, PhD; Gerald Gartlehner, MD, MPH; Sally Yaacoub, PhD; Isabelle Boutron, MD, PhD; Lukas Schwingshackl, PhD, MSc; Julia Stadelmaier, MSc; Isolde Sommer, PhD; Farzaneh Alebouyeh, MSc; Sivem Afach, PhD; Joerg Meerpohl, MD, PhD; and Philippe Ravaud, MD, PhD

Scoping review task	Method used	Suggested automation tools to increase efficiency or gain	Level of automation	Availability of the tool (open access/proprietary)
Team establishment	Personnel needed for the completion of the review	Crowd sourcing or using task sharing platforms i.e., JBI SUMARI and Covidence. Management teams—Trello	Level 1	Proprietary
Protocol development	Templated reports of some report items	JBI SUMARI and Revman	Level 1	Proprietary
Searches	Running searches on databases	Automatic alerts from databases. Database (i.e., Cochrane, JBI and Campbell, Health Database Advanced Search (HDAS)) Automatic retrieval of full-text papers (e.g., CrossRef), Polyglot Search Translator Litsuggest, eSuRFr	Level 4	CrossRef and Litsuggest (Open access) The rest of the tools are proprietary.
Deduplication	Deleting duplicates	Endnote, Systematic Review Accelerator (SRA De-duplicate), Abstrackr, Rayyan, EPPI-Reviewer, Deduklick, Covidence and DistillerSR.	Level 3	Abstrackr (Open access) Rayyan (Open Access to early career researchers) The remainder of tools are proprietary.
Screening	Selecting studies for inclusion	Crowdsourced inclusion decision and machine learning classifier (Abstrackr, Rayyan, BIBOT, EPPI-Reviewer, ASReview and DistillerSR). Covidence is also used for this step.	Level 2	Abstrackr (Open access) Rayyan (Open Access to early career researchers) The remainder of tools are proprietary.
Data extraction	Extracting information on participants, Concept and context	Machine learning-information extraction systems such as DistillerSR and ExaCT), Tabulizer, UDPipe, Webplotdigitizer. For non-machine learning but wanting a faster extraction process- Covidence	Level 4	UDPipe (open access) and the remainder of the tools are proprietary.
Data charting	Presentation of data	GoogleCharts, Tableau Public, Google Data studio, Drawio, Eppi- reviewer, NVivo.	Level 1	Googlecharts, Google data studio and Drawio (open access). The remainder of the tools are proprietary.
Report writing	Templated reports of some report items	JBI SUMARI, Revman and Robotreviewer	Level 1	Proprietary

Living evidence ecosystem

