

THESE DE DOCTORAT DE L'UNIVERSITE PARIS DIDEROT (PARIS 7)

Ecole doctorale 393 Pierre Louis de santé publique

Santé publique : Epidémiologie et Sciences de l'Information Biomédicale

DOCTORAT

Epidémiologie

Présenté par

Aïda BAFETA

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE PARIS DIDEROT – PARIS VII

Epidémiologie et méthodologie des méta-analyses en réseau

Thèse dirigée par le Pr Philippe RAVAUD

Présentée et soutenue publiquement

Le 16 janvier 2015

Devant le jury composé de :

Monsieur le Professeur Pascal Auquier	Rapporteur
Monsieur le Professeur Bruno Falissard	Rapporteur
Monsieur le Professeur Stefan Michiels	Examinateur
Madame le Professeur Florence Tubach	Examinateur
Monsieur le Professeur Philippe Ravaud	Directeur de thèse

Remerciements

En tout premier lieu, mes plus sincères remerciements vont à Monsieur le Professeur Philippe Ravaud, mon directeur de thèse. Je voudrais le remercier d'avoir cru en moi et mes capacités. Il m'a fait bénéficier durant mon master et mon doctorat de ses conseils avisés, de son grand savoir et ses qualités scientifiques. Je suis très heureuse qu'il ait accepté de guider mes premiers pas en recherche, et qu'il m'ait permis de travailler sur des sujets passionnants.

Je souhaite exprimer toute ma gratitude envers les membres du jury qui ont accepté de juger mon travail. J'adresse mes sincères remerciements à Monsieur le Professeur Pascal Auquier et à Monsieur le Professeur Bruno Falissard qui ont bien voulu être rapporteurs de ce travail, et à Madame le Professeur Florence Tubach et Monsieur le Professeur Stefan Michiels qui ont accepté de participer à ce jury de thèse.

Je remercie les membres de l'équipe doctorale (ED 393) m'ayant accompagnée pendant cette thèse, notamment Madame le Professeur France Mentré, Evelyne Guilloux-Pelmar et Lydie Martorana.

Je souhaite remercier particulièrement Monsieur Ludovic Trinquart pour son soutien et son aide dans la réalisation des travaux de cette thèse, ainsi que Mademoiselle Raphaèle Seror.

Je tiens aussi à remercier Mademoiselle Josefin Blomkvist et Mademoiselle Elise Diard pour leur support technique durant les travaux de cette thèse.

Je remercie particulièrement Monsieur Viet Thi Tran pour sa relecture et ses commentaires.

Je tiens aussi à remercier tous les membres du Centre d'Epidémiologie clinique, de l'unité INSERM U1153 pour les bons moments passés ensemble en particulier Isabelle P, Caroline,

*Aurélie, Marjorie, Carolina, Nizar, Ali, Laurence, Elodie, Gabriel, Marie, Adeline, Sara V,
Jean-Philippe, Isabelle B, Raphaël, Mickaël, Florence, Ahmed, Elis, Valérie, Valéria, Audrey,
Magalie, Timor, Aymar.*

*J'ai une pensée pour les nombreux thésards et camarades de Master avec qui j'ai partagé cette
aventure, notamment Anne, Amélie, Youri, Stéphanie E, Clémence, Idriss, Cécile, Céline,*

Stéphanie S et Romana.

*Enfin, je tiens à remercier en premier lieu ma mère ma fierté qui est une source d'inspiration
pour moi et mon père pour son éducation. A ma sœur mon âme sœur, à mes frères pour leur
soutien, à ma tante Tene, Kalif, ma tante Rokia, à ma marraine pour leurs encouragements, à
mes beaux-parents pour leur aide au quotidien, à toute ma grande famille. A mes amis fidèles
Cédric, Daniella et Thierry, Oumar, Ousmane, Katia et Richard. Et tous les autres qui ont su me
soutenir pendant les périodes difficiles.*

A Marc et Sara ma force, mon bonheur et mes joies de vivre.

Résumé

La méta-analyse en réseau permet d'évaluer toutes les comparaisons deux à deux possibles entre interventions même en absence de certains essais. Depuis les années 2000, l'utilisation des méta-analyses en réseau a considérablement augmenté. La méta-analyse en réseau est une méta-analyse, elle est probablement soumise aux mêmes types de biais et doit se réaliser dans le cadre rigoureux d'une revue systématique. Cependant, aucune étude ne s'était intéressée aux éléments essentiels de la réalisation d'une revue systématique dans les méta-analyses en réseau comme la recherche de la littérature, l'évaluation du risque de biais ou la présentation des résultats. Nous avons évalué si les méta-analyses en réseau suivent les recommandations essentielles à la conduite et la présentation des revues systématiques. Nous avons montré que certains éléments essentiels à la conduite et la présentation des revues systématiques ne sont pas rapportés. Notamment, les étapes concernant la recherche bibliographique, l'évaluation de la qualité méthodologique des essais inclus, le biais de publication sont mal rapportées. Nous avons évalué comment sont présentés les résultats dans les méta-analyses en réseau. Nous avons mis en évidence que la présentation des résultats dans les méta-analyses en réseau est très hétérogène. Les auteurs des méta-analyses en réseau ne présentent pas ou rapportent de manière inadéquate les éléments essentiels permettant d'évaluer l'ensemble des preuves disponibles (les interventions incluses, les comparaisons directes incluses, le nombre d'essais et de patients inclus dans chaque comparaison), les estimations deux à deux provenant des comparaisons directes, indirectes ou mixtes. Nos résultats montrent la nécessité de développer et de promouvoir des recommandations pour la conduite et la présentation de ce type d'étude, dans le but d'améliorer la transparence et la reproductibilité.

Mots-clefs

Revue systématique, méta-analyse, méta-analyse en réseau, présentation, réalisation, qualité

Abstract

Network meta-analyses allow for estimating all possible pair-wise comparisons between interventions, even some trials are lacking. In the past few years, network meta-analyses have been increasingly used for comparing healthcare interventions. Network meta-analyses are primarily meta-analyses; they are probably subject to the same types of bias and should therefore be performed according to the explicit and rigorous methods used in systematic reviews. However, did not evaluate the presence of essential methodological components of the systematic review process such as conducting a literature search and assessing the risk of bias of individual studies, or the presentation of findings in reports of network meta-analyses.

We assessed whether network meta-analyses follow the key methodological recommendations for reporting and conduct of systematic reviews. We show that the essential methodological components of the systematic review process—conducting a literature search and assessing risk of bias of individual studies—are frequently lacking in reports of network meta-analyses. We assessed how the results of network meta-analyses are reported. We show that the results of network meta-analyses are heterogeneously reported. The author inadequately reported essential component to judge the amount of evidence (that is, the interventions included in the network, direct comparisons between interventions, number of randomised trials and patients for each comparison) and the relative effect sizes between interventions from direct evidence, indirect evidence, and the network meta-analysis. Development of reporting guidelines to assist authors in writing and readers in critically appraising reports of network meta-analyses is timely.

Keywords

Systematic review, meta-analysis, network meta-analysis, reporting, conduct, quality,

Laboratoire d'accueil

Centre de recherche Epidémiologies et Biostatistique INSERM U1153,

Équipe : Méthodes en évaluation thérapeutique des maladies chroniques

Centre d'épidémiologie clinique Hôpital Hôtel Dieu

1, place du parvis Notre-Dame

75181 Paris, France

Publications liées à la thèse

Ce travail de doctorat s'appuie sur les articles suivants :

Aïda Bafeta, Ludovic Trinquart, Raphaèle Seror, Philippe Ravaud.

Analysis of the systematic reviews process in reports of network meta-analyses: methodological systematic review. *British Medical Journal* 2013;347:f3675

Aïda Bafeta, Ludovic Trinquart, Raphaèle Seror, Philippe Ravaud.

Reporting of results from network meta-analyses: methodological systematic review.

British Medical Journal 2014;348:g1741

Table des matières

1. Chapitre : Introduction et Objectifs	13
1.1. Les revues systématiques et les méta-analyses	14
1.1.1. L'évaluation thérapeutique	14
1.1.2. Nécessité d'une synthèse de l'information médicale	14
1.1.3. Principes des revues systématiques et des méta-analyses.....	16
1.1.4. Les étapes essentielles des revues systématiques et des méta-analyses.....	18
1.1.4.1. La recherche bibliographique	18
1.1.4.2. Evaluation de la qualité méthodologique des essais inclus	20
1.1.5. Les différents biais dans les revues systématiques et les méta-analyses.....	23
1.1.5.1. Le biais de publication.....	23
1.1.6. Les recommandations pour rapporter et évaluer les revues systématiques.....	26
1.1.7. Limites des méta-analyses classiques dans l'évaluation thérapeutique	35
1.2. Méthodologie des méta-analyses en réseau.....	39
1.2.1. L'introduction des méta-analyses en réseau	39
1.2.2. Définition d'un réseau d'essais	40
1.2.3. Les comparaisons indirectes ajustées.....	44
1.2.4. Les comparaisons mixtes	45
1.2.5. Puissance dans une méta-analyse en réseau.....	48
1.2.6. Les différentes approches statistiques.....	49
1.2.7. L'approche Bayésienne décrite par Lu et Ades	49
1.2.8. Problèmes spécifiques aux essais multi-bras	51
1.2.9. Les hypothèses sous-jacentes de la méta-analyse en réseau	52
1.2.9.1. Evaluation des hypothèses.....	52
1.2.10. Présentation et évaluation des hypothèses dans les méta-analyses en réseau	57
1.2.11. Limites de l'évaluation des hypothèses dans les méta-analyses en réseau	58
1.2.12. Présentation des résultats des méta-analyses en réseau	60
1.2.12.1. Les formats pour présenter les résultats d'une méta-analyse en réseau	60
1.2.12.2. Classement des interventions	62
1.2.13. Les différents biais dans la méta-analyse en réseau.....	65
1.2.14. Les recommandations pour rapporter et évaluer les méta-analyses en réseau.....	66
1.3. Justification de la démarche de la thèse	71

2. Chapitre : Evaluation des composantes méthodologiques des revues systématiques dans les publications de méta-analyse en réseau.....	73
2.1. Justification de l'étude	74
2.2. Objectifs	74
2.3. Méthodes	74
2.4. Résultats	76
2.5. Discussion	78
Article	80
3. Chapitre : Présentation des résultats dans les méta-analyses en réseau	93
3.1. Justification de l'étude	94
3.2. Objectif.....	94
3.3. Méthodes	94
3.4. Résultats	95
3.5. Discussion	98
Article	100
4. Chapitre: Discussion et Perspectives	110
4.1. Discussion	111
4.1.1. Principaux résultats.....	111
4.1.2. Implication	112
4.1.2.1. La recherche bibliographique dans les méta-analyses en réseau.....	112
4.1.2.2. L'évaluation du risque de biais dans les méta-analyses en réseau	113
4.1.2.3. La présentation des résultats dans les méta-analyses en réseau	114
4.1.3. Conclusion	119
4.2. Perspectives.....	122
Références	125
Annexe	135

Liste des figures

Figure 1 : Exemple de graphique en forêt de 5 essais contrôlés randomisés comparant les effets d'une ventilation protectrice à ceux d'une ventilation conventionnelle. Adaptée de la source ²²	18
Figure 2 : Estimation de l'effet traitement de la poly-chimiothérapie comparé à la monothérapie dans le cancer ovarien dans les essais publiés et non publiés. Adaptée de la source ⁴⁹	25
Figure 3 : Evolution du nombre de publications ayant pour thématique la méta-analyse en réseau. Adaptée de la source ¹⁰³	40
Figure 4 : Evolution du nombre de références indexées dans Pubmed ayant pour thématique la méta-analyse en réseau	40
Figure 5 : Exemple des différentes géométries de réseau possible.....	42
Figure 6 : Exemples de réseaux existants dans la littérature. Adaptée de la source ¹⁰⁵	42
Figure 7 : Exemple de comparaison mixte. Adaptées de la source ¹²⁰ :	46
Figure 8 : Représentation graphique d'un réseau de 4 interventions.....	47
Figure 9 Représentation des résultats sur le schéma de la géométrie du réseau Adaptée de la source ¹⁵⁵ ...	61
Figure 10 : Courbe des probabilités des rangs. Adaptée de la source ¹⁵⁷	63
Figure 11 : Courbe des probabilités de rang cumulées. Adaptée de la source ¹⁰³	64
Figure 12 : Les différents domaines de l'évaluation de la validité d'une méta-analyse en réseau selon ISPOR. Adaptée de la source ¹⁶⁴	69
Figure 13 : L'approche GRADE pour évaluer la qualité des méta-analyses en réseau. Adapté de la source ¹⁶⁵	70
Figure 14 : Nombre de publications de méta-analyse en réseau par an.....	76
Figure 15 : Présentation des éléments relatifs aux résultats dans les méta-analyses en réseau.	97
Figure 16 : Représentation de l'ensemble des Forest plots sous le format d'une matrice dans l'exemple des traitements tocolytiques dans l'accouchement prématuré. Adaptée de la source ¹⁸²	116
Figure 17 : Représentation de l'ensemble des Forest plots sous le format d'un tableau dans l'exemple des traitements tocolytiques dans l'accouchement prématuré. Adaptée de la source ¹⁸²	117

Figure 18 : Représentation de l'ensemble des rangs moyens dans l'exemple des traitements tocolytiques dans l'accouchement prématuré. Adaptée de la source ¹⁸² 118

Figure 19 : Evolution du nombre de citations via Web Of Science faisant référence à un article de méta-analyse en réseau dans la littérature médicale. 120

Liste des tableaux

Tableau 1 : Les éléments devant être dans un rapport de revue systématique selon PRISMA	29
Tableau 2: Les éléments devant être dans un rapport de revue systématique selon MECIR	31
Tableau 3: Les éléments devant être dans un rapport de revue systématique selon AMSTAR	34
Tableau 5 : Matrice des résultats dans une méta-analyse en réseau. Adaptée de la source ¹⁵⁶	61
Tableau 4 : Rangs des différents traitements antiplaquettaire. Adaptée de la source ¹⁵⁷	64
Tableau 6: Les éléments devant figurer dans un rapport de méta-analyse en réseau selon ISPOR.....	68

Abréviations

ECR: Essai contrôlé randomisé

IC 95 % : Intervalle de confiance à 95 %

FDA: Food and Drug Administration

ITT: Intention de traiter

ROB *tool*: Risk Of Bias tool

QUOROM: the Quality of Reporting of Meta-analyses

PRISMA: Preferred Reporting Items of Systematic reviews and Meta-Analyses

MECIR: Methodological Expectations of Cochrane Intervention Reviews

AMSTAR: Assessment of Multiple Systematic Reviews

GRADE: Grading of Recommendations Assessment, Development and Evaluation

MCMC : Monte Carlo des Chaines de Markov

SUCRA: Surface Under the Cumulative Ranking

RCT: Randomized Controlled Trial,

NMA: Network Meta-analysis

ISPOR: the International Society for Pharmacoeconomics and Outcomes Research

1. Chapitre : Introduction et Objectifs

1.1. Les revues systématiques et les méta-analyses

1.1.1. L'évaluation thérapeutique

Les traitements et les interventions médicales doivent avoir une efficacité prouvée scientifiquement, c'est l'évaluation thérapeutique. Actuellement, les preuves les plus solides, les moins entachées de biais (erreur systématique entre une estimation et la véritable valeur du paramètre estimé), proviennent d'essais contrôlés randomisés (ECRs) qui constituent la méthode de référence pour évaluer l'efficacité ou la sécurité d'une intervention médicale¹⁻³ ; par exemple sous réserve de la qualité de la randomisation, du suivi comparable des groupes et de l'évaluation en aveugle des critères de jugement.

1.1.2. Nécessité d'une synthèse de l'information médicale

Pendant longtemps, les ECRs ont été réalisés et interprétés de manière indépendante. Ainsi le nombre de données médicales devant être prises en compte par les professionnels de santé a littéralement explosé. Dans de nombreux domaines, il est devenu tout simplement impossible pour un professionnel de santé de lire, d'évaluer et de synthétiser de façon critique l'état des connaissances actuelles. Il a été estimé que les médecins devraient lire quotidiennement 17 articles, tous les jours de l'année⁴.

Dans les années 1970, l'épidémiologiste anglais Archibald Lemman Cochrane (1909-1988) a mis en évidence l'absence de synthèse régulièrement mise à jour, rapportant l'efficacité des interventions médicales. *"It is surely a great criticism of our profession that we have not organised a critical summary, by specialty or subspecialty, adapted periodically, of all relevant randomised controlled trials"*⁵.

L'absence de synthèse rigoureuse a donné lieu à des retards concernant l'élaboration de recommandations pour ou contre l'utilisation d'interventions en pratique clinique usuelle ; ou bien face à la répétition inutile d'études scientifiques. Par exemple, Antman et ses collègues ont constaté un écart de 15 ans entre d'une part, le moment où l'ensemble des preuves montrait de manière solide l'efficacité de la thrombolyse dans les cas d'infarctus aigu du myocarde et le moment où son utilisation a été largement recommandée, et d'autre part entre le moment où l'ensemble des preuves aurait pu démontrer les dangers possibles de l'utilisation des anti-arythmiques et les recommandations en défaveur de leur utilisation⁶. Dans une autre étude réalisée par Fergusson et ses collègues, il a été constaté qu'un groupe de 7 674 patients ont participé à 55 essais sur l'efficacité de l'aprotinine pendant plus de onze ans, et ce, après que l'efficacité du médicament ait été clairement établie par la méta-analyse^{7,8}.

Il faut attendre les années 1980 pour voir apparaître les premières revues systématiques. Elles font la synthèse critique de toutes les données scientifiques répondant à une question clinique précise. Elles servent à comprendre ou à interpréter les résultats contradictoires des études primaires sur une même intervention et à voir si l'ensemble des preuves scientifiques concordent d'une étude à une autre⁹. Les revues systématiques permettent aussi d'identifier les domaines où les données sont disponibles ou insuffisantes, et d'orienter les recherches futures. Les revues systématiques se sont considérablement multipliées¹⁰. Ainsi les revues systématiques de la Collaboration Cochrane (Organisation internationale indépendante à but non lucratif) dont le but est d'apporter des informations exactes, actualisées sur l'efficacité thérapeutique 1) en recensant l'ensemble des ECRs, 2) en produisant et en diffusant des revues systématiques en médecine, 3) en participant à l'amélioration de la méthodologie des revues systématiques, 4) en aidant à la communication et à la transparence dans la prise de décision en médecine, sont devenues incontournables en éducation thérapeutique.

Les revues systématiques sont devenues des outils essentiels pour la synthèse des connaissances scientifiques, elles participent à la formation continue des professionnels de santé, servent de base pour l'élaboration des recommandations de pratique clinique et de support pour les décideurs de santé^{10-12 13,14}.

1.1.3. Principes des revues systématiques et des méta-analyses

Une revue systématique suit une méthodologie systématique rigoureuse et reproductible en vue de minimiser les biais^{15,16}. Les étapes essentielles de la revue systématique sont les suivantes :

- 1) Définir les objectifs, c'est-à-dire une question clinique claire et précise.
- 2) Définir les critères d'éligibilités c'est-à-dire spécifier les caractéristiques des études (population, critère de jugement, durée du suivie), les critères utilisés pour sélectionner les études (année, langue et statut de la publication).
- 3) Une recherche systématique qui vise à identifier toutes les études qui répondent aux critères d'éligibilités.
- 4) Une évaluation de la validité des résultats des études incluses.
- 5) Une présentation systématique avec la synthèse des caractéristiques et des résultats des études incluses.

Une revue systématique peut inclure ou non une synthèse quantitative des résultats appelée méta-analyse. Les méta-analyses synthétisent l'ensemble des résultats des études répondant à une question thérapeutique donnée, dans le but d'obtenir une estimation moyenne (combinée) de l'effet traitement^{17,18}. L'estimation de l'effet traitement la moins biaisée provient de méta-analyse réalisée dans le cadre d'une revue systématique¹⁹.

Les méta-analyses permettent d'augmenter la puissance statistique, la précision de l'estimation de la taille de l'effet traitement et la généralisation des résultats en prenant en compte plusieurs essais qui augmentent le nombre de patients et apportent une plus grande variabilité dans les caractéristiques de base des patients étudiés¹⁶⁻¹⁹. Elles permettent aussi d'expliquer la variabilité des résultats entre essais, de réaliser des analyses en sous-groupes (en augmentant le nombre des effectifs dans les sous-groupes), de répondre à une question non initialement posée par les essais²⁰. Il existe deux types de méta-analyse : Les méta-analyses réalisées sur données résumées extraites à partir des articles rapportant les résultats des études incluses; les méta-analyses sur données individuelles réalisées à partir des données individuelles des patients. Les méta-analyses calculent une moyenne pondérée de l'estimation de l'effet traitement issue de chaque étude sélectionnée. Dans l'analyse, les études ayant une variance de l'estimation de l'effet traitement la plus petite ont plus de poids dans la méta-analyse et donc contribuent davantage à l'estimation de l'effet traitement moyen¹⁹.

Les deux principaux modèles statistiques qui existent pour combiner les résultats de plusieurs études sont²¹:

- Un modèle à effet fixe : on considère que l'effet traitement est une constante. Chaque étude mesure cette même constante, ainsi la série d'études est une série de mesures de ce même effet traitement soumis à des fluctuations d'échantillonnages.
- Un modèle à effet aléatoire : émet l'hypothèse que le résultat d'une étude est composé d'une partie fixe, commune à l'ensemble des études, et d'une partie propre à cette étude donc variable.

Il est commode de présenter les résultats d'une méta-analyse graphiquement par un graphique en forêt (Forest plot).

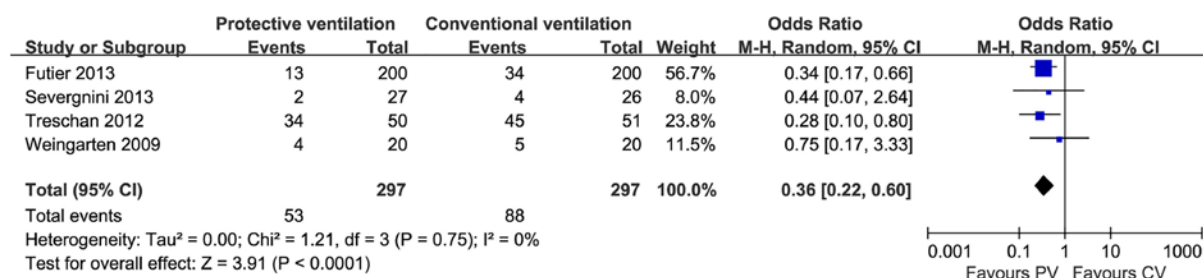


Figure1 : Exemple de graphique en forêt de 5 essais contrôlés randomisés comparant les effets d'une ventilation protectrice à ceux d'une ventilation conventionnelle . Adaptée de la source ²²

A gauche se trouve la description des études incluses, le nombre d'événements et de patients, et leur poids alloué selon la taille de l'effectif. A droite, le résultat de chaque étude est représenté sous forme d'un carré dont la taille est proportionnelle au poids de l'essai dans la méta-analyse et un trait horizontal de part et d'autre qui représente l'intervalle de confiance. L'effet traitement combiné est représenté sous la forme d'un losange. Le milieu du losange représente la valeur estimée ponctuelle et les extrémités sont les bornes de l'intervalle de confiance. Cet exemple montre que la ventilation protectrice chez les patients subissant une anesthésie générale permettait de diminuer le nombre d'infection pulmonaire par rapport à la ventilation conventionnelle avec un Odds ratio combiné à 0.36 (IC 95%.0.22 à0.60).

1.1.4. Les étapes essentielles des revues systématiques et des méta-analyses

1.1.4.1. La recherche bibliographique

La recherche bibliographique est une étape cruciale dans l'élaboration d'une revue systématique¹⁹. Elle vise à repérer l'ensemble des publications scientifiques pertinentes répondant à la question posée. Le processus de recherche doit se faire de façon rigoureuse et exhaustive afin de réduire au maximum le risque de biais, et doit être rapportée de façon transparente dans le but de permettre l'évaluation par le lecteur et d'être reproductible^{23,24}. Il est

recommandé de faire appel au soutien d'un documentaliste spécialisé²⁵. La recherche bibliographique doit combiner plusieurs approches afin de s'assurer de repérer l'ensemble des études pertinentes répondant au critère d'éligibilité.

L'approche traditionnelle pour identifier des études est de conduire une nouvelle recherche de la littérature définie comme une « *de novo search* ».

Une seconde option est de fonder sa recherche sur les précédentes revues systématiques et méta-analyses, cette approche est appelée « *overview* ».

Enfin une troisième possibilité est de rechercher initialement des revues systématiques et méta-analyses, puis de les mettre à jour avec des études non couvertes par la période de recherche de ces revues systématiques et méta-analyses. On parle de « *mixed search* ».

La Collaboration Cochrane considère que la recherche doit se faire dans au moins deux bases de données électroniques. Elle estime que les bases de données CENTRAL, MEDLINE et EMBASE devraient être interrogées, d'autres bases spécifiques au sujet peuvent être consultées outre MEDLINE, par exemple : PsycINFO, CINAHL ou LILACS^{19,26}.

La recherche bibliographique doit être complétée par la recherche d'études publiées ou non, qui ne sont pas indexées dans des bases de données électroniques²⁷. Plusieurs moyens peuvent être employés :

- Recherche des revues systématiques déjà existantes à l'aide de bases de données spécifiques aux revues systématiques et aux méta-analyses (Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects) ou le registre PROSPERO (<http://www.crd.york.ac.uk/PROSPERO/>).
- Recherche de la « littérature grise » – la littérature grise comprend notamment des études qui ne sont pas publiées officiellement dans des livres ou des périodiques, et peut inclure les résumés, les actes d'un colloque, des dissertations, des thèses, des rapports

de projet et des documents du gouvernement – par exemple en utilisant le moteur de recherche Google Scholar, ou en contactant les experts.

- Examen des listes de références des études sélectionnées et des tables des matières des périodiques clés : cela permet de repérer des publications récentes qui n'ont pas encore été indexées.
- Recherche des études en cours dans les registres essais (ClinicalTrials.gov (<http://www.clinicaltrials.gov/>) ou Clinical Trials Registry Platform.
- Rechercher des citations à partir d'articles clés par exemple en utilisant Science Citation Index.

1.1.4.2. Evaluation de la qualité méthodologique des essais inclus

Les caractéristiques des essais inclus dans les revues systématiques et les méta-analyses ont un impact dans l'estimation des résultats²⁸. Si les études sources sont potentiellement biaisées, les résultats de la méta-analyse risquent de l'être aussi. La méta-analyse n'améliore pas la qualité des études prises en compte ; comme disent les Anglo-saxons, la méta-analyse est un processus GIGO « garbage in, garbage out »²⁹. Un exemple illustratif est le cas extrême d'une étude réalisée sur l'utilisation des fluides synthétiques (l'hydroxyéthylamidon) utilisés en réanimation³⁰. Les auteurs ont réalisé une revue systématique avec deux méta-analyses, l'une incluait et l'autre excluait 7 études réalisées par l'anesthésiste allemand Joachim Boldt, dont 89 articles ont été invalidés en 2011, avec des preuves que des données avaient été inventées pour certaines (ce sont les 7 études qui n'ont pas été rétractées car publiées avant 1999, alors que les investigations ont été faites sur les publications de 1999 et après.) Les études de Boldt favorisaient l'emploi de ces produits. En excluant les études de Boldt, l'hydroxyéthylamidon était responsable d'insuffisance rénale (RR=1.27, IC 95% 1.09 à 1.47), et de mortalité augmentée (RR=1.09, IC 95% 1.02 à 1.17). Ce cas illustre bien l'influence des caractéristiques

des études sur les résultats et les conclusions de la méta-analyse. Si cette dernière inclut des études biaisées, son résultat risque d'être biaisé ce qui peut conduire à des conclusions erronées²⁸. Par ailleurs, des études méta-épidémiologiques, en examinant de nombreuses méta-analyses sur divers sujets, ont identifié certains facteurs pouvant influencer les résultats des méta-analyses avec une surestimation de l'effet traitement comme l'absence d'assignation secrète dans les essais³¹⁻³⁹. C'est pourquoi, l'évaluation de la validité des résultats des études est une étape essentielle de la méta-analyse. Cette évaluation est basée sur deux concepts de validité interne et validité externe¹⁹.

1.1.4.2.1. La validité interne d'un essai contrôlé randomisé

La validité interne permet de s'assurer que le résultat obtenu reflète bien la réalité et qu'il n'est pas dû à un biais ou au hasard. Selon la Collaboration Cochrane¹⁹, il existe 4 types de biais qui peuvent affecter les résultats d'un ECR :

- Le biais de sélection survient lorsque les 2 groupes de l'ECR ne sont pas comparables, c'est-à-dire qu'il existe des différences systématiques entre les caractéristiques initiales des groupes à comparer. Pour limiter ce biais, il faut réaliser une génération adéquate de la séquence de randomisation, s'assurer le respect de l'assignation secrète.
- Le biais de performance (biais de suivi ou de réalisation) survient lorsque les 2 groupes ne sont pas suivis de la même manière, c'est-à-dire qu'il existe des différences systématiques dans l'administration des soins entre les 2 groupes de l'ECR. Par exemple le nombre de co-interventions est différent dans l'un des groupes de l'ECR. Pour limiter ce biais, une solution est de réaliser l'essai en double aveugle, c'est-à-dire de s'assurer que l'administration de l'intervention se fait à l'insu des patients et des thérapeutes.

- Le biais de détection (biais d'évaluation) survient lorsqu'il existe des différences systématiques dans la mesure du critère de jugement entre les groupes. Il peut être limité par l'insu des évaluateurs vis-à-vis du bras de traitement.
- Le biais d'attrition survient quand des patients randomisés sont écartés de l'analyse. Les patients sont des perdus de vue, ou sortent de l'étude, ou sont des non-observants. Ce biais peut être corrigé en limitant le nombre de perdus de vue et par une analyse en intention de traiter (ITT) avec une gestion adéquate des données manquantes.

1.1.4.2.2. La validité externe d'un essai contrôlé randomisé

La validité externe se rapporte à la généralisation des résultats⁴⁰⁻⁴³. La validité externe permet d'évaluer si les résultats sont extrapolable aux patients à traiter en pratique courante en évaluant les caractéristiques des patients (âge, sexe, facteurs de risque, sévérité de la maladie, comorbidités), les caractéristiques des traitements (doses, durée, fréquence d'administration, traitements concomitants), les caractéristiques des centres et des thérapeutes, le contexte de soin, le pays.

1.1.4.2.3. Evaluation du risque de biais

De nombreux outils ont été proposés pour évaluer la qualité des études dans le cadre d'une revue systématique. La plupart de ces outils sont des échelles qui donnent un score à chaque item et produisent un score global⁴⁴. La Collaboration Cochrane recommande de ne plus utiliser ce type d'échelle mais plutôt d'évaluer le risque de biais en utilisant un outil spécifique appelé *Risk of Bias Tool* (ROB tool)¹⁹. Celui-ci comprend des domaines liés à la validité interne: biais de sélection, biais de performance, biais de détection, les biais d'attrition, présentation sélective des critères de jugement et d'autres sources potentielles de biais.

1.1.4.2.4. Prise en compte du risque de biais des essais dans les revues systématiques et méta-analyses

Deux méthodes sont actuellement recommandées par la Collaboration Cochrane¹⁹: soit restreindre la méta-analyse aux essais à faible risque de biais, soit présenter des méta-analyses stratifiées selon le risque de biais. Une troisième méthode consiste à faire la méta-analyse en incluant toutes les études identifiées et à décrire le risque de biais dans chacune d'elles.

1.1.5. Les différents biais dans les revues systématiques et les méta-analyses

Les revues systématiques et les méta-analyses sont soumises à deux grands types de biais, ceux liés aux essais inclus et ceux liés à la réalisation de celles-ci appelés les méta-biais qui se manifestent seulement lorsque l'on examine une collection d'essais⁴⁵.

Toute synthèse d'information est sujette à un biais qui lui est propre : les biais de dissémination qui désignent des différences systématiques entre les résultats rapportés et non rapportés influencés par la nature et la direction des résultats.

Plusieurs types de biais de dissémination sont possibles : le biais de publication⁴⁶⁻⁵³ ; le biais du délai de publication (un délai de publication plus rapide pour des résultats favorables)^{54,55} ; le biais de publication multiple (multiplication des publications)⁵⁶⁻⁵⁸, les biais de langage et de citation (plus souvent publiée en anglais et citée)⁵⁹⁻⁶² ; le biais dans le recueil des données. Enfin le biais de présentation sélective des résultats (« selective outcome reporting bias » : sélectionner une partie seulement des résultats sur la base de leur significativité statistique)⁶³⁻⁶⁵.

1.1.5.1. Le biais de publication

Il est connu qu'un certain nombre de travaux ne font jamais l'objet d'une publication⁶⁶. En effet il a été mis en évidence que les études non significatives sont moins publiées que les études

positives et que plus les études sont de petite taille, plus le risque est grand que seules les études très positives soient publiées, alors que les grandes études sont généralement publiées quels que soient leurs résultats⁴⁶⁻⁵³. Les mécanismes de cette censure sont multiples et peuvent provenir soit des comités de relecture, soit des firmes finançant l'étude, mais aussi d'une autocensure que s'inflige spontanément les auteurs^{48,67}.

Dans une méta-analyse, si aucune recherche poussée d'essais non publiés n'est entreprise, le risque couru est d'inclure principalement des essais positifs, ce qui peut conduire à une surestimation de l'effet traitement⁶⁸. Par exemple dans un article publié dans le *New England Journal*⁶⁹ Turner et ses collègues ont choisi douze antidépresseurs pour lesquels ils ont comparé les données publiées dans la littérature médicale à celles de la Food and Drug Administration (l'agence américaine d'évaluation du médicament [FDA]). Ils ont mis en évidence que la majorité des études positives ont été publiées, et que la majorité des études négatives n'ont pas été publiées. Parmi ces 36 études négatives, 11 sont publiées avec des résultats présentés comme étant positifs. Les auteurs ont également découvert que l'efficacité de ces médicaments est moins grande que ce que l'on pourrait en conclure en examinant seulement la littérature publiée. L'effet traitement des médicaments est de 11% à 69% plus important dans les données publiées que dans les analyses de la FDA. Un autre exemple sur l'impact du biais de publication sur les résultats et les conclusions des méta-analyses provient d'une méta-analyse rapportant une réduction de la mortalité chez les patients avec un infarctus aigu du myocarde traités par du magnésium intraveineux⁷⁰. Ce résultat a été contredit par un essai de grande taille (ISIS trial)⁷¹. L'analyse des données révèle un probable biais de publication. C'est pourquoi les auteurs des revues systématiques et des méta-analyses ont intérêt à rechercher ce biais de publication éventuel et à le corriger si c'est possible.

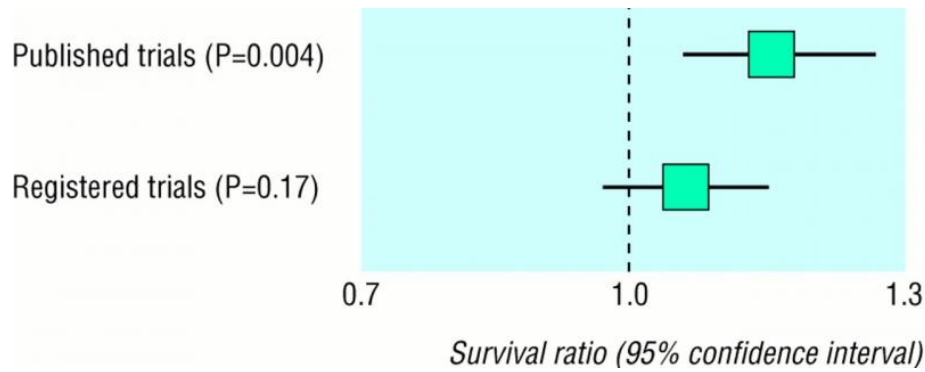


Figure 2 : Estimation de l'effet traitement de la poly-chimiothérapie comparé à la monothérapie dans le cancer ovarien dans les essais publiés et non publiés. Adaptée de la source⁴⁹

Ces résultats montraient que la poly-chimiothérapie améliorait significativement la survie dans les études publiées. Par contre, si l'on prenait les études non publiées (recensées dans le registre), la quantité d'effet était réduite, voire non statistiquement significative.

Les solutions au biais de publication commencent à être apportées –notamment par les registres d'essais⁷². Par ailleurs, tout initiateur d'essai ayant un centre situé aux Etats-Unis a l'obligation d'enregistrer son étude et un protocole abrégé avant l'inclusion du premier patient. Ces registres prospectifs d'essais permettent alors de retrouver la trace de tous les essais entrepris qu'ils soient ou non publiés. De plus en septembre 2004, l'International Committee of Medical Journal Editors demande aux journaux de ne plus publier à partir du 1er juillet 2005 les ECRs qui n'auraient pas été enregistrés dans un registre d'essai lors de leur mise en place⁷³.

1.1.5.1.1. L'évaluation du biais de publication

L'évaluation du biais de publication peut être réalisée de différentes façons. La plus courante est la réalisation d'un funnel-plot (graphique en entonnoir inversé): ce graphique représente pour chaque essai la taille de l'échantillon en ordonnée d'une part, la taille de l'effet traitement en abscisse d'autre part^{74,75}. On fait l'hypothèse que la répartition des points est symétrique autour

de la vraie valeur de l'effet traitement en absence de biais. Le nuage de points a la forme d'un entonnoir inversé d'où le nom du graphique. L'existence d'un funnel-plot asymétrique permettait de suspecter un biais de publication. D'autres causes d'asymétrie liées à des défauts méthodologiques sont possibles. Autrement dit, la mise en évidence d'une asymétrie doit inviter à interpréter les résultats d'une méta-analyse avec prudence. Des tests statistiques pour rechercher un biais ne sont cependant guère puissants, surtout à cause du nombre faible d'études incluses dans la plupart des méta-analyses⁷⁶. Il est donc souvent difficile d'affirmer formellement un biais de publication et encore plus difficile de le récuser. C'est pourquoi la recherche bibliographique d'une revue systématique et d'une méta-analyse se doit d'être la plus exhaustive en recherchant tous les essais publiés et non publiés dans le domaine clinique en question pour limiter ce biais.

1.1.6. Les recommandations pour rapporter et évaluer les revues systématiques

Les revues systématiques et les méta-analyses sont considérées comme le plus au niveau de preuve⁷⁷, et reposent sur des méthodes rigoureuses et transparentes. Il est important que la communication de ses résultats soit présentée adéquatement. La présentation complète et précise des résultats de recherches originales dans la littérature est essentielle pour que les professionnels de soins de santé puissent traduire correctement les résultats de recherche en pratique clinique.

En 1987, Mulrow a examiné 50 synthèses publiées dans quatre grands journaux médicaux entre 1985 et 1986. Il a montré que les auteurs ne mentionnaient pas l'utilisation de méthodes rigoureuses pour trouver, évaluer et synthétiser l'information⁷⁸. Une autre étude réalisée par Sacks et ses collègues en 1987, puis 1996 qui évaluaient également la qualité de publications de méta-analyses, est arrivée aux mêmes conclusions^{79,80}. Ces études ont mis en évidence l'absence de directives pour la réalisation et la présentation des revues systématiques et des méta-analyses.

C'est pourquoi en 1996 pour améliorer la transparence, la qualité méthodologique et la présentation des revues systématiques et des méta-analyses; un groupe international a élaboré un guide comportant une liste d'éléments appelé QUOROM (« the QUality Of Reporting Of Meta-analyses »)⁸¹. Il permet de vérifier que les étapes de réalisation d'une méta-analyse ont été correctement conduites et rapportées. Ainsi il permet d'évaluer de façon critique la qualité méthodologique des études et d'examiner les biais possibles. En 2007 Moher et ses collègues ont évalué 300 revues systématiques, ils ont montré que de nombreux aspects liés aux méthodes des revues systématiques n'étaient pas rapportés, cependant ils ont observé une amélioration de la qualité de présentation de ces études au cours du temps¹⁰.

En 2009, ces directives ont été actualisées afin d'inclure les revues systématiques et ont été rebaptisées PRISMA (un acronyme de « Preferred Reporting Items of Systematic reviews and Meta-Analyses »)⁸². PRISMA est un ensemble de points fondés sur des faits probants qui devraient être utilisés pour les publications de revues systématiques et de méta-analyses; il s'agit d'un guide dynamique révisé au fur et à mesure que les preuves changent. PRISMA est une liste de 27 points regroupant les composantes importantes de chaque section de la revue systématique, notamment le titre, le résumé, les méthodes, les résultats, les conclusions et le financement. Le tableau 1 présente les éléments devant être dans un rapport de revue systématique selon PRISMA.

A ce jour, il existe des données probantes encourageantes selon lesquelles les recommandations améliorent la qualité de présentation des recherches publiées⁸³⁻⁸⁶. De plus, un grand nombre de périodiques exigent que les auteurs suivent ces lignes directrices pour rédiger les manuscrits qui leur sont soumis⁸⁷.

Certaines organisations comme la Collaboration Cochrane exigent que les revues systématiques et les méta-analyses soient présentées conformément à une structure normalisée. Pour cela ils ont

édité des recommandations MECIR (un acronyme de « Methodological Expectations of Cochrane Intervention Reviews ») qui sont une liste de 108 points classés en obligatoire et fortement recommandé¹⁹. Le tableau 2 présente les éléments devant être dans un rapport de revue systématique selon MECIR.

Il existe de nombreux systèmes pour évaluer la qualité méthodologique des revues systématiques. AMSTAR (un acronyme de « Assessment of Multiple SysTemAtic Reviews »)^{88,89} est un outil qui se compose de 11 éléments en rapport avec la réalisation et la présentation des revues systématiques. Le tableau 3 présente les éléments devant être dans un rapport de revue systématique selon AMSTAR. Li et ses collègues⁹⁰ proposent d'utiliser quatre critères : la recherche bibliographique, l'évaluation de la qualité méthodologique des essais inclus, l'utilisation d'une méthode statistique adéquate, et la cohérence des conclusions avec les résultats pour classer les résultats d'une revue systématique à faible, incertain, ou à haut risque de biais.

L'approche GRADE (un acronyme de « Grading of Recommendations Assessment, Development and Evaluation »)⁹¹ fournit une méthode pour évaluer la qualité des données, qui s'appuie sur l'examen de la méthodologie de l'étude, du risque de biais au sein de l'étude, de l'uniformité et de la précision des résultats des études, et sur l'évaluation du biais de publication dans le but d'établir des recommandations.

En conclusion, une présentation inadaptée des résultats pourrait influencer l'interprétation et l'application de la recherche publiée, c'est pourquoi les auteurs doivent s'appuyer sur les recommandations comme PRISMA pour améliorer la transparence et la qualité de leur recherche

Tableau 1 : Les éléments devant être dans un rapport de revue systématique selon PRISMA

SECTION DE LA REVUE SYSTEMATIQUE	DEFINITION
TITRE	
Titre	Identifier la publication comme une revue systématique, méta-analyse ou les deux.
RESUME	
Résumé	Fournir un résumé structuré avec le contexte, les objectifs, la provenance des données, les critères d'éligibilité des études, des participants et des interventions, les méthodes pour analyser les données, les résultats, les limites, les conclusions et les implications ; le numéro d'enregistrement de l'étude.
INTRODUCTION	
Contexte	Décrire le rationnel de la revue.
Objectifs	Une formulation claire des questions abordées avec une référence aux participants, aux interventions, aux comparaisons, aux critères de jugement, et au design de l'étude (acronyme de Patients Interventions Comparisons Outcomes Study design PICOS).
METHODES	
Protocole et enregistrement de l'étude	Indiquer si un protocole existe et où il est disponible, fournir le numéro d'enregistrement de l'étude.
Critères d'éligibilité	Spécifier les caractéristiques des études (PICOS, durée du suivi), rapporter les critères utilisés pour sélectionner les études (année, langue et statut de la publication).
Sources d'information	Décrire toutes les sources d'information (les bases de données utilisées avec la date que couvre la recherche, contact aux auteurs pour identifier des études supplémentaires) avec la date de dernière consultation.
Recherche bibliographique	Présenter l'équation de recherche électronique d'au moins une base de données avec les limites utilisées.
Sélection des études	Décrire les méthodes utilisées pour la sélection des études.
Extraction des données	Décrire les méthodes utilisées pour l'extraction des données et le processus pour obtenir un consensus entre les investigateurs.
Les données	Lister et définir toutes les variables pour lesquelles des données ont été recherchées ainsi que les hypothèses et les simplifications apportées.
Risque de biais des essais	Décrire les méthodes utilisées pour évaluer le risque de biais ainsi que comment cette information va être utilisée dans l'analyse des données.
Mesures résumées	Mentionner les mesures résumées principales (risque relatif, différence de moyenne).
Analyse des résultats	Décrire les méthodes utilisées pour combiner les résultats en incluant une mesure pour évaluer l'hétérogénéité pour chaque méta-analyse (I^2).
Risque de biais dans l'étude	Décrire les méthodes utilisées pour évaluer le risque de biais qui peut affecter l'effet traitement combiné (biais de publication).
Analyses supplémentaires	Décrire les méthodes utilisées pour faire des analyses supplémentaires (analyse de sensibilité, analyse en sous-groupe, méta-régression).
RESULTATS	
Sélection des études	Donner le nombre d'études présélectionnées, évaluées et incluses dans la revue avec les raisons d'exclusions à chaque étape, idéalement sous la forme d'un « flow chart ».
Caractéristiques des études	Pour chaque étude présenter les caractéristiques pour lesquelles des données ont été extraites.
Risque de biais des essais	Présenter l'évaluation du risque de biais de chaque étude.
Résultats des essais	Pour tous les critères de jugement considérés présenter pour

	chaque étude : les données résumées pour chaque groupe d'intervention, l'estimation de l'effet traitement avec un intervalle de confiance idéalement sous la forme d'un Forest plot.
Analyse des résultats	Présenter les résultats de chaque méta-analyse avec un intervalle de confiance et une mesure pour évaluer l'hétérogénéité pour chaque méta-analyse (I^2).
Risque de biais dans l'étude	Présenter l'évaluation des risques de biais de l'étude.
Analyses supplémentaires	Présenter les résultats des analyses supplémentaires.
DISCUSSION	
Principaux résultats	Résumer les principaux résultats avec les forces de l'étude.
Limites	Discuter des limites de l'étude et des critères de jugement (risque de biais, biais de publication).
Conclusions	Fournir une interprétation générale des résultats en comparant avec les autres études, et les implications pour des futures recherches
FINANCEMENT	
Financement	Décrire les sources de financement pour la revue systématique et le rôle du promoteur dans l'étude.

Tableau 2: Les éléments devant être dans un rapport de revue systématique selon MECIR

No item	Status	Item name
Title and authors		
1.	Highly desirable	Format of title Format of title
2.	Mandatory	Authors: List names and affiliations of all authors
Abstract		
3.	Mandatory	Writing the abstract: Prepare a structured abstract to provide a succinct summary of the review.
4.	Mandatory	Abstract, Background : Summarize the rationale and context of the review
5.	Mandatory	Abstract, objective(s): State the main objective(s)
6.	Mandatory	Abstract, Search methods: the date of the last search, databases and other sources.
7.	Mandatory	Abstract, Selection criteria
8.	Mandatory	Abstract, Data collection and analysis: Summarize any noteworthy methods for selecting studies, collecting data, evaluating risk of bias and synthesizing findings
9.	Mandatory	Abstract, Main results: number of studies and participants
10.	Highly desirable	Abstract, Main results: study characteristics
11.	Mandatory	Abstract, Main results: bias assessment
12.	Mandatory	Abstract, Main results: findings
13.	Mandatory	Abstract, Main results: adverse effects
14.	Mandatory	Abstract, Main results: format of numerical results
15.	Highly desirable	Abstract, Main results: interpretability of findings
16.	Mandatory	Abstract, Authors' conclusions
17.	Mandatory	Completeness of main review text
18.	Mandatory	Consistency of summary versions of the review
Background		
19.	Mandatory	Background: A concise description of the condition or problem, definition of the intervention and how it might work, and why it is important to do the review
20.	Highly desirable	Background headings: Include the four standard headings when writing the Background
21.	Mandatory	Background references
22.	Mandatory	Background text: Avoid the use of plagiarized text.
23.	Mandatory	Main objective
24.	Highly desirable	Secondary objectives
25.	Mandatory	Economic evidence
26.	Mandatory	Qualitative research evidence
Methods		
27.	Highly desirable	Reference protocol
Criteria for considering studies for this review		
28.	Mandatory	Eligibility criteria for types of study: study designs
29.	Mandatory	Eligibility criteria for types of study: study reports
30.	Mandatory	Eligibility criteria for types of participants
31.	Mandatory	Eligibility criteria for types of interventions
32.	Mandatory	Role of outcomes: If measurement of particular outcomes is used as an eligibility criterion
33.	Mandatory	Outcomes of interest: Primary and secondary outcomes, and define acceptable ways of measuring them.
Search methods for identification of studies		
34.	Mandatory	Search sources : List all sources searched, including: databases, trials registers, web sites and grey literature
35.	Mandatory	Latest searches : Provide the date of the last search and the issue / version number (where relevant) for each database
36.	Mandatory	Search timeframe: Any restrictions placed on the time period covered by the search.
37.	Mandatory	Searches for different types of evidence
38.	Mandatory	Search strategies for bibliographic databases
39.	Highly desirable	Search strategies for other sources
Data collection and analysis		
40.	Mandatory	Inclusion decisions
41.	Mandatory	Data collection process

42.	Highly desirable	Requests for data
43.	Mandatory	Data items : List the types of information that were sought from reports of included studies
44.	Mandatory	Transformations of data
45.	Highly desirable	Missing outcome data
46.	Mandatory	Tools to assess risk of bias in individual studies
47.	Mandatory	Effect measures: the effect measures used by the review authors to describe effect sizes
48.	Mandatory	Quantitative synthesis: Describe any methods for combining results across studies
49.	Mandatory	Addressing risk of bias
50.	Mandatory	Non-standard designs: If designs other than individually randomized, parallel-group randomized trials are included, describe any methods used to address clustering, matching or other design features of the included studies.
51.	Mandatory	Studies with more than two groups: Explain how they are addressed and incorporated into syntheses
52.	Highly desirable	Risk of reporting bias across studies
53.	Mandatory	Subgroup analyses: If subgroup analysis (or meta-regression) was performed, state the potential effect modifiers with rationale for each, stating whether each was defined a priori or post hoc.
54.	Highly desirable	Summary of findings
Results		
Description of studies		
55.	Mandatory	Flow of studies: Studies from the number(s) of references identified in the search to the number of studies included in the review, ideally using a flow chart.
56.	Highly desirable	Lack of included studies: If a review identifies no eligible studies, restrict the Results section to a description of the flow of studies and any brief comments about reasons for exclusion of studies.
57.	Mandatory	Excluded studies: List key excluded studies and provide justification for each exclusion.
58.	Highly desirable	Studies awaiting classification : List the characteristics of any studies that have been identified as potentially eligible but have not been incorporated into the review
59.	Mandatory	Ongoing studies
60.	Mandatory	Table of 'Characteristics of included studies
61.	Mandatory	Included studies : Provide a brief narrative summary of any included studies
62.	Mandatory	Table of "Characteristics of included studies": sample sizes
63.	Mandatory	Table of "Characteristics of included studies": methods
64.	Mandatory	Table of "Characteristics of included studies": participants
65.	Mandatory	Table of "Characteristics of included studies": interventions
66.	Mandatory	Table of "Characteristics of included studies": outcomes
67.	Highly desirable	Table of "Characteristics of included studies": dates
68.	Mandatory	Table of "Characteristics of included studies": funding source
69.	Mandatory	Table of "Characteristics of included studies": declarations of interest
70.	Highly desirable	Choice of intervention groups in multi-arm studies.
71.	Mandatory	References to included studies
Risk of bias in included studies		
72.	Mandatory	Risk of bias' table : Present a 'Risk of bias' table for each included study, with judgements about risks of bias, and explicit supports for these judgements
73.	Highly desirable	Summary assessments of risk of bias: Summarize the risk of bias across domains for each key outcome for each included study
74.	Mandatory	Risk of bias in included studies: Provide a brief narrative summary of the risks of bias among the included studies
Effects of interventions		
75.	Mandatory	Use of "Data and analysis" headings
76.	Highly desirable	Presenting data
77.	Mandatory	Number of studies and participants
78.	Highly desirable	Source of data
79.	Mandatory	Multiple outcome data : Describe any <i>post hoc</i> decisions that might give rise to accusations of selective outcome reporting
80.	Highly desirable	Ordering of results and "Data and analysis" section
81.	Mandatory	Pre-specified outcomes: To avoid selective outcome reporting (in truth or in perception), the review should address all outcomes specified in the protocol.

82.	Mandatory	Statistical uncertainty
83.	Highly desirable	P values
84.	Mandatory	Tables and Figures
85.	Highly desirable	Number of Tables and Figures
86.	Mandatory	Consistency of results : Ensure that all statistical results presented in the main review text are consistent between the text and the “Data and analysis” tables
87.	Mandatory	Different scales : Explain how studies measuring an outcome of interest using different scales
88.	Mandatory	Interpretability of results: Ensure that key findings are interpretable, or are re-expressed in an interpretable way.
89.	Mandatory	Studies without usable data: Comment on the potential impact of studies that apparently measured outcomes but did not contribute data that allowed the study to be included in syntheses.
90.	Highly desirable	Missing outcome data: Discuss the implications of missing outcome data from individual participants.
91.	Highly desirable	Skewed data: Discuss the possibility and implications of skewed data when analysing continuous outcomes.
92.	Highly desirable	Forest plots
93.	Highly desirable	Multiple subgroup analyses and sensitivity analyses
94.	Mandatory	Labels on plots : Label the directions of effect and the intervention groups in forest plots with the interventions being compared
95.	Highly desirable	Risk of bias across studies
96.	Highly desirable	Reporting biases
97.	Highly desirable	« Summary of findings » table
98.	Mandatory	Assessments of the quality of the body of evidence: Provide justification or rationale for any measures of the quality of the body of evidence for each key outcome. If a ‘Summary of findings’ table is used, use footnotes to explain any downgrading or upgrading according to the GRADE system.
Discussion		
99.	Highly desirable	Discussion headings
100.	Mandatory	Limitations
Authors’ conclusions		
101.	Mandatory	Conclusions: implications for practice
102.	Mandatory	Conclusions: implications for research
Acknowledgements		
103.	Mandatory	Acknowledgements : Acknowledge the contribution of people not listed as authors of the review, including any assistance from the Cochrane review Group, non-author contributions to searching, data collection, study appraisal or statistical analysis, and the role of any funders
Contributions of authors		
104.	Mandatory	Contributions of authors
Declarations of interest		
105.	Mandatory	Declarations of interest
Differences between protocol and review		
106.	Mandatory	Changes from the protocol
107.	Highly desirable	Methods not implemented
Sources of support		
108.	Mandatory	Funding

Tableau 3: Les éléments devant être dans un rapport de revue systématique selon AMSTAR

Les items d'AMSTAR : qui permettent d'évaluer la qualité méthodologique des revues systématiques

1. Y avait-il un protocole de disponible ?

La question clinique et les critères d'éligibilité devraient être établis avant la conduite de la revue systématique.

2. Y avait-il une double sélection des études et une double extraction des données ?

Il devrait avoir au moins deux lecteurs indépendants et une procédure de consensus entre les lecteurs pour régler les désaccords.

3. Y avait-il une recherche exhaustive de la littérature ?

Au moins deux bases de données électroniques doivent être recherchées. La publication doit rapporter la date de la recherche, le nom des bases de données utilisées (par exemple CENTRALE, EMBASE, et MEDLINE). Les mots clés et / ou les termes MeSH doivent être indiqués et si possible, la stratégie de recherche doit être fournie. Toutes les recherches doivent être complétées par la consultation de manuels, de registres spécialisés, et des références des études trouvées.

4. Y avait-il une sélection des études selon le statut de la publication (publié ou non publié)?

Les auteurs doivent indiquer qu'ils ont consulté ou exclu les rapports indépendamment ou en fonction du statut de la publication et de la langue.

5. Y avait-il une liste des études de disponible ?

Une liste d'études incluses et exclues doit être fournie.

6. Est-ce que les caractéristiques des études incluses étaient rapportées ?

Sous une forme agrégée (un tableau), les données des études sur les participants, les interventions et les critères de jugement doivent être fournies. Les caractéristiques de toutes les études analysées doivent être fournies par exemple l'âge, le sexe, les données socio-économiques, la maladie (statut, durée, gravité).

7. Y avait-il une évaluation de la qualité des études incluses de fournir ?

A priori les méthodes d'évaluation doivent être fournies.

8. Y avait-il une formulation appropriée des résultats de l'évaluation de la qualité dans les conclusions ?

Les résultats de l'évaluation méthodologique et la qualité des études devraient être considérés dans l'analyse et les conclusions de la revue systématique, et explicitement formulés dans les recommandations.

9. Y avait-il des méthodes appropriées pour combiner les résultats des études ?

Évaluation de l'homogénéité des études (tester l'homogénéité, I^2). Si une hétérogénéité, existe un modèle à effets aléatoires doit être utilisé et/ou la pertinence clinique doit être prise en considération (par exemple est-il judicieux de combiner ces études?).

10. Y avait-il une évaluation du biais de publication ?

L'évaluation du biais de publication devrait inclure une combinaison d'aides graphiques (par exemple, graphique en entonnoir, d'autres tests disponibles) et/ou de test statistique (par exemple, tests de régression Egger).

11. Y avait-il une mention des conflits d'intérêts ?

Les sources de financement doivent être clairement citées dans la revue systématique et les études incluses.

1.1.7. Limites des méta-analyses classiques dans l'évaluation thérapeutique

Bien que les revues systématiques qui incluent une méta-analyse soient reconnues comme la meilleure source d'information pour évaluer une intervention thérapeutique⁷⁷, cette méthode comporte certaines limites qu'il faut souligner. Les limites peuvent être divisées en deux catégories : celles liées à l'hétérogénéité et celles liées aux choix méthodologiques pour réaliser une méta-analyse comme l'évaluation du biais de publication ou du risque de biais, ou alors comment combiner l'ensemble des études disponibles²⁵.

L'hétérogénéité est définie comme la variation de l'estimation de l'effet traitement entre les études, elle est inhérente à la méta-analyse. En effet, les essais au sein d'une méta-analyse sont nécessairement au minimum légèrement différents. L'hétérogénéité peut être d'ordre clinique avec des variations dans les caractéristiques de la population (par exemple : âge, sexe, maladie) ou dans les interventions dans le bras expérimental et de référence (dose, durée...) ou bien d'ordre méthodologique (variation dans les méthodes utilisées entre les études) ou les caractéristiques cliniques (par exemple, la diversité des populations de patients ou d'interventions). Elle peut être évaluée sur le plan statistique par le Q de Cochran, la statistique I^2 et la variance inter-essais τ^2 . Explorer les sources d'hétérogénéité est nécessaire pour pouvoir combiner les résultats de plusieurs études. Cependant, l'évaluation des relations entre une caractéristique moyenne (âge moyen des patients) et l'effet traitement doit être interprétée avec précaution. Des relations importantes peuvent être oubliées et des relations factices peuvent être trouvées. En effet, une relation à l'échelle de l'étude ne signifie pas forcément qu'il existe une relation à l'échelle du patient et inversement. Par exemple l'effet traitement peut être lié à l'âge des patients au sein de chaque essai mais peut ne pas être associé à l'âge moyen des patients entre les essais d'une même méta-analyse, et inversement. Par ailleurs, des facteurs de confusion peuvent affecter les résultats de cette évaluation^{92,93}.

En outre, l'hétérogénéité représente un défi pour l'interprétation en pratique clinique, même si elle est évaluée et ses causes sont explorées. En effet l'incertitude générée par la combinaison de plusieurs ECRs provenant de sources disparates (avec des différences dans les patients, les interventions et les résultats évalués), et l'absence d'informations nécessaires à un clinicien pour appliquer les résultats à un patient, peuvent limiter la capacité de faire des inférences causales. Il a été recommandé d'interpréter les résultats comme une association plutôt qu'un effet de causalité.

Les différents choix méthodologiques pour réaliser une méta-analyse sont nombreux. Par exemple une des questions cruciales de la méta-analyse est: « Peut-on combiner toutes les études disponibles ? » Plusieurs recommandations sont possibles, souvent contradictoires car il y a un manque de clarté ou de consensus sur la meilleure façon de traiter ces questions. Par exemple la Collaboration Cochrane recommande soit de restreindre la méta-analyse aux essais à faible risque de biais, soit de présenter une analyse stratifiée selon le risque de biais^{19,29}. Il s'ensuit des publications avec diverses approches et analyses, ce qui peut rendre obscure l'interprétation des conclusions. Cette question a été illustrée dans une étude réalisée par Dechartres et ses collègues⁹⁴. Les auteurs ont évalué l'impact de différents choix stratégiques pour l'analyse sur l'effet traitement ainsi que sur les conclusions de la méta-analyse dans 163 méta-analyses (1240 ECRs), dont 92 (705 ECRs) évaluées un critère de jugement subjectif et 71 (535 ECR) évaluées un critère de jugement objectif. Pour chaque méta-analyse, ils ont comparé l'effet traitement obtenu selon 5 stratégies d'analyse : une méta-analyse de tous les essais, une méta-analyse restreinte aux essais les plus grands (dernier quartile de la distribution de l'effectif), effet traitement de l'essai le plus précis (défini comme l'essai pour lequel la variance de l'effet traitement était la plus faible), une méta-analyse de tous les essais avec un ajustement sur le « small study effect »⁹⁵ (La tendance qu'ont les petits essais à montrer des effets traitements plus

importants que les essais ayant une taille plus importante au sein d'une méta-analyse), et une méta-analyse restreinte aux essais à faible risque de biais. L'effet traitement était plus important dans la méta-analyse de tous les essais que dans l'essai le plus précis. Il y avait une petite différence de l'effet traitement entre les méta-analyses de tous les essais et celle restreinte aux essais les plus grands, et une plus grande différence entre les méta-analyses de tous les essais et les méta-analyses limitées. Pour 3 domaines du *Rob tool* (la séquence de randomisation, l'assignation secrète et l'aveugle), les effets traitements sont plus grands pour les essais à haut risque de biais ou à risque de biais non défini par rapport aux essais à faible risque de biais (à la fois pour les critères de jugement subjectifs et objectifs), cependant, il n'y avait pas de différence significative pour le domaine des données manquantes. En conclusion les différentes stratégies aboutissaient à des résultats différents qui peuvent entraîner une altération des conclusions.

Des solutions peuvent être apportées pour améliorer la validité des méta-analyses. L'hétérogénéité doit être reconnue et explorée de manière approfondie et ainsi nuancer les conclusions de la méta-analyse. La transparence de la recherche, la qualité de présentation des publications sont essentielles pour améliorer l'interprétation des résultats. De nombreuses recommandations (tel PRISMA⁸²) sont disponibles pour améliorer la présentation des publications de méta-analyses.

Une autre limite des méta-analyses est comment évaluer l'ensemble des traitements disponibles pour une condition clinique. Une méta-analyse classique permet de fournir une estimation de l'efficacité de chaque traitement par rapport à un traitement de référence, mais pas de la différence d'efficacité de l'ensemble de ces nouveaux traitements deux à deux. Dans de très nombreuses situations cliniques plusieurs traitements co-existent. Par exemple plus d'une soixantaine d'antidépresseurs sont disponibles. Bien qu'il existe un traitement actif, les ECRs évaluant de nouvelles interventions sont souvent conduits contre

placebo à cause des coûts, des intérêts commerciaux ou pour obtenir leur autorisation de mise sur le marché et de nombreuses comparaisons dites de « face-face » ne sont pas faites, ou ne sont pas publiées⁹⁶⁻⁹⁸. La plupart des ECRs, des revues systématiques et des méta-analyses classiques actuellement publiés ne permettent donc pas de répondre à la question pourtant cruciale du choix de l'intervention la plus efficace parmi l'arsenal thérapeutique disponible dans une situation donnée. Ainsi les professionnels de santé ont souvent du mal à faire les bons choix thérapeutiques. Il devient alors nécessaire de développer des techniques statistiques afin d'évaluer l'ensemble des traitements actuellement disponibles dans une situation clinique donnée et de les classer selon leur efficacité ou et leur sécurité relative.

1.2. Méthodologie des méta-analyses en réseau

1.2.1. L'introduction des méta-analyses en réseau

On observe une augmentation constante du nombre d'interventions pour une même indication clinique rendant le nombre de comparaisons à entreprendre important. De plus de très nombreuses comparaisons n'ont pas été étudiées. En effet dans de nombreux domaines thérapeutiques, les traitements concurrents sont évalués préférentiellement *versus* placebo et dans d'autres domaines thérapeutiques⁹⁶⁻⁹⁸, les comparaisons directes sont impossibles à effectuer, par exemple pour des raisons de faisabilité (dans le cas des maladies rares).

Pour pallier au manque d'essais comparatifs des traitements actifs entre eux, le recours aux méta-analyses en réseau (« network meta-analysis, multiple-treatment meta-analysis, mixed treatment comparison meta-analysis ») est envisageable. Cette méthode permet d'estimer toutes les comparaisons deux à deux entre les traitements, même en absence de certains essais et d'établir un classement des traitements⁹⁹. Contrairement à la méta-analyse classique qui permet d'évaluer une intervention relative à une intervention donnée.

L'utilisation des méta-analyses en réseau a augmenté de manière considérable ces dernières années. Une recherche électronique du nombre de référence via Pubmed à la date du 25 juillet 2014, montre une augmentation constante du nombre de références ayant pour thématique la méta-analyse en réseau. Les méta-analyses en réseau répondent à la nécessité de déterminer le meilleur traitement disponible. C'est pourquoi les cliniciens et les chercheurs, ainsi que les agences nationales d'évaluation des technologies de santé et les organismes de réglementation des médicaments utilisent de plus en plus ces méthodes¹⁰⁰⁻

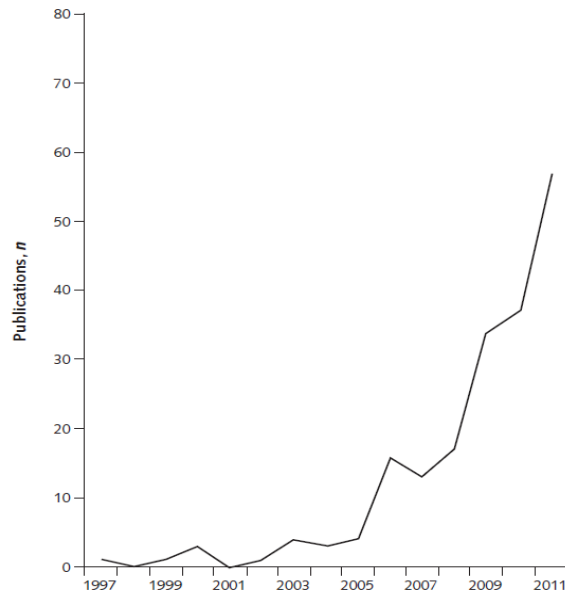


Figure 3 : Evolution du nombre de publications ayant pour thématique la méta-analyse en réseau. Adaptée de la source¹⁰³

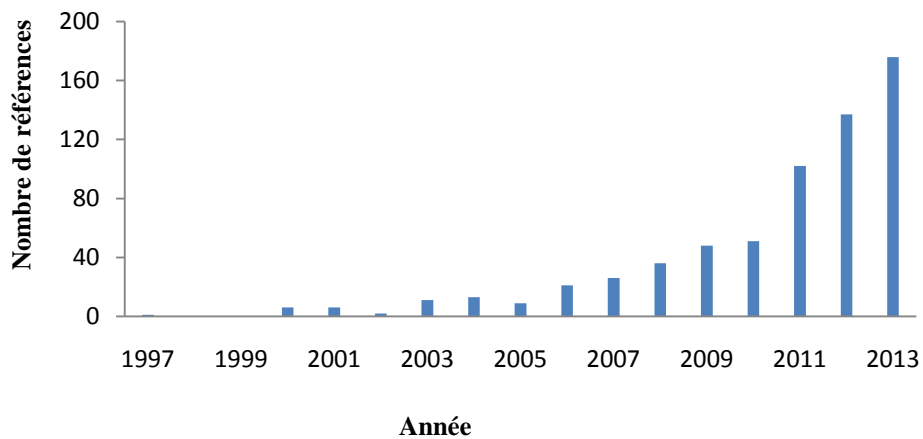


Figure 4 : Evolution du nombre de références indexées dans Pubmed ayant pour thématique la méta-analyse en réseau

1.2.2. Définition d'un réseau d'essais

Un réseau d'essais (« network of evidence, network trials ») se définit par l'ensemble de toutes les comparaisons entre toutes les interventions pour une même indication clinique. Il permet de synthétiser l'ensemble de l'information apporté par divers essais d'un domaine thérapeutique. Un réseau d'essais peut être représenté par un schéma géométrique dans laquelle un nœud représente une intervention disponible et les liens entre les nœuds représentent les comparaisons entre les différentes interventions respectives. Le réseau doit être connecté c'est-à-dire toutes les interventions incluses dans le réseau sont liées entre elles par au moins une chaîne de comparaison directe. La taille des nœuds est proportionnelle au nombre de patients qui ont reçu l'intervention et l'épaisseur des lignes est proportionnelle au nombre d'essais.

Plusieurs géométries sont possibles^{104,105}:

- Réseau en étoile (« Star network »), toutes les interventions ont un seul comparateur commun.
- Réseau en échelle (« Ladder network ») il existe au moins deux comparateurs communs et aucune boucle fermée.
- Réseau avec une boucle fermée (« closed network ») impliquant au moins trois interventions connectées entre elles.
- Réseau entièrement connecté (« Fully connected network »), toutes les interventions ont été comparées les unes par rapport autres dans plusieurs essais.
- Réseau complexe (« Complex network ») plusieurs boucles fermées coexistent avec des interventions qui peuvent avoir des connexions rares.

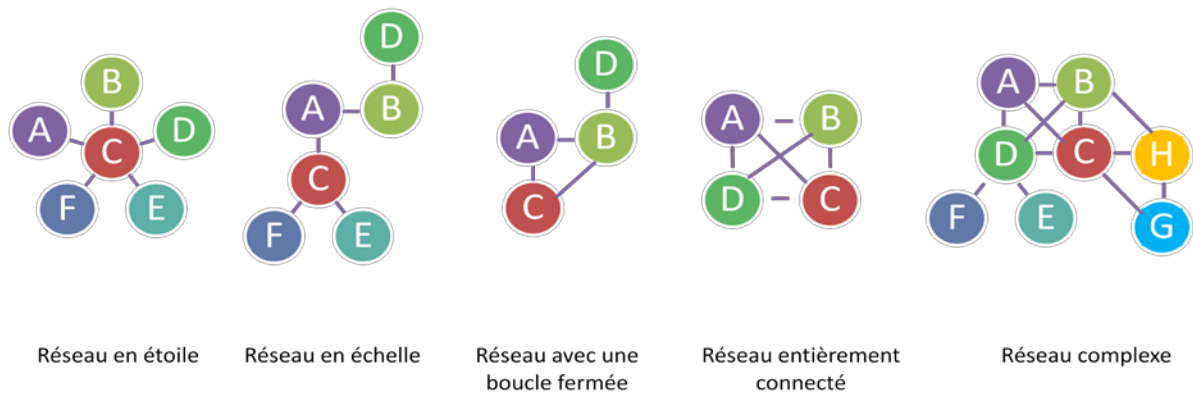
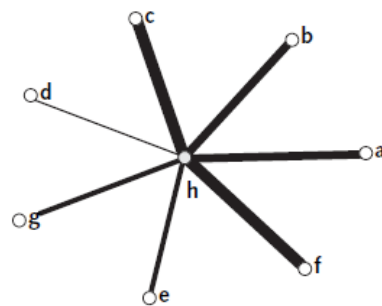


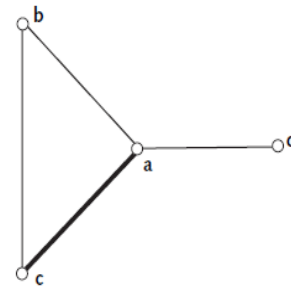
Figure 5 : Exemple des différentes géométries de réseau possible

Second-generation antiepileptic drugs in partial epilepsy



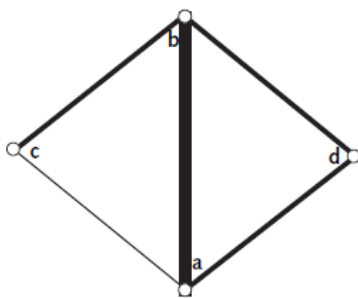
a: levetiracetam, b: gabapentin, c: lamotrigine, d: oxcarbazepine, e: tiagabine, f: topiramate, g: zonisamide, h: placebo

Self-monitoring of glucose in type 2 diabetes



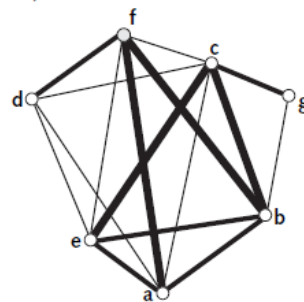
a: self-monitoring of blood glucose, b: self-monitoring of urine glucose, c: no self-monitoring, d: self-monitoring of blood glucose with regular feedback

Prophylaxis for *Pneumocystis carinii* in HIV-infected patients



a: trimethoprim-sulfamethoxazole, b: pyrimethamine, c: dapsons, d: dapsons-pyrimethamine

Antihypertensive treatment (incidence of diabetes)



a: diuretic, b: ACE-i, c: CCB, d: ARB, e: β -blocker, f: placebo, g: β -blocker or diuretic

Figure 6 : Exemples de réseau existants dans la littérature. Adaptée de la source¹⁰⁵

Le premier réseau sur la gauche est un réseau dit en étoile, le deuxième et le troisième réseau sont des réseaux avec au moins une boucle fermée ; enfin le quatrième réseau est un réseau complexe.

Le réseau peut être décrit par trois mesures la diversité, la co-occurrence et l'homophilie¹⁰⁵⁻¹⁰⁷. La diversité est faible quand un réseau a peu d'interventions et que celles-ci ne sont pas également représentées (certaines interventions sont comparées plus souvent que les autres). La co-occurrence évalue si des comparaisons sont surreprésentées ou sous-représentées. Enfin l'homophilie mesure si les interventions dans un même groupe (par exemple pour les médicaments de la même classe thérapeutique) sont plus susceptibles d'être comparées les unes par rapport aux autres qu'avec les interventions d'autres groupes.

Un élément clé pour comprendre un réseau est l'évaluation de sa géométrie¹⁰⁵. Quels traitements ont été comparés dans des essais dits de face-face, quels traitements sont reliés indirectement à travers un ou plusieurs comparateurs communs, et quel est le niveau de preuve pour chaque comparaison. En examinant les liens entre les interventions d'une manière graphique, un lecteur peut déterminer la solidité des preuves du réseau dans son ensemble, à savoir quelles comparaisons sont surreprésentées ou sous-représentées, et si le réseau est bien connecté. Mieux est connecté le réseau, plus fiable seront les estimations. La diversité et la force d'un réseau sont déterminées par le nombre d'interventions et les comparaisons disponibles, la façon dont elles sont représentées dans le réseau, et la quantité de preuves qu'elles portent. Un déséquilibre important en termes de quantité de preuves pour chaque intervention peut avoir une incidence sur la puissance et la fiabilité de l'estimation globale^{108,109}.

La géométrie du réseau reflète le contexte clinique et les stratégies de recherche entreprises. La structure du réseau a pu être façonnée par des choix rationnels pour les comparateurs ou

par des biais spécifiques par exemple lorsqu'on observe une prédilection pour les essais contre placebo. L'exploration de la géométrie du réseau peut aider à mettre en évidence des « gaps » ou un agenda de recherche biaisé (des comparaisons importantes sont évitées) même si les essais dans un domaine sont parfaitement conçus et réalisés. Dans plusieurs domaines, ont été mis en évidence un agenda de recherche biaisé par l'examen de la géométrie: la polyarthrite rhumatoïde, les antifongiques, les stratégies thérapeutiques du cancer avancé du sein, du myélome multiple ou les traitements du carcinome baso cellulaire^{98,107,110-112}.

1.2.3. Les comparaisons indirectes ajustées

Les premières méthodes proposées pour effectuer des comparaisons indirectes consistaient à regrouper les patients selon l'intervention qu'ils avaient reçue et à comparer les bras traitements actifs des essais ou comparer l'estimation ponctuelle l'efficacité A versus C à celle de B versus C^{113,114}. De telles comparaisons (dites naïves) sont biaisées car elles ne tiennent pas compte de la méthodologie des essais, des populations, ni de la nature randomisée des essais¹¹⁵. Par exemple on pourrait imaginer que tous les essais A versus C ont eu lieu dans les années 1980 alors que les essais B versus C ont eu lieu récemment. De nombreux paramètres comme la définition du stade de la maladie, le contexte de soin, la population (mode de vie, facteurs de risques...) ont changé pendant ce temps. Ils peuvent modifier le risque de base de la population et interagir avec l'efficacité des traitements. Les populations ainsi obtenues ne sont plus comparables. Les estimations issues de ce type de comparaisons sont biaisées et ne doivent pas être réalisées.

Une technique plus élaborée est la comparaison indirecte «ajustée», intégrant la nature randomisée de chaque essai. Cette approche consiste à collecter tous les ECRs comparant par exemple, une intervention A et une intervention C et tous les ECRs comparant une

intervention B et C, où C est un comparateur commun pour estimer l'effet traitement de A versus B. L'effet de A par rapport à B est estimé par $\delta_{AB} = \delta_{AC} + \delta_{CB}$ et la variance associée par $var(\delta_{AB}) = var(\delta_{AC}) + var(\delta_{CB})$ ¹¹⁶.

L'hypothèse principale de ce modèle est que l'efficacité relative d'un traitement est la même dans tous les essais inclus dans la comparaison indirecte.

Lorsqu'il existe un ou plusieurs essais comparant l'intervention A et l'intervention B, on peut privilégier la source directe et ne pas considérer la source indirecte, en estimant que l'essai dit de face-face est la meilleure source d'information. Cependant, ce type d'essais n'est pas indemne de biais par exemple ceux liés aux défauts méthodologiques, ou le type de financement^{117,118}. Il existe aucune preuve théorique que les comparaisons indirectes soient biaisées systématiquement dans le sens d'une surestimation ou d'une sous-estimation de l'effet traitement. Par ailleurs, une revue de 44 méta-analyses avec des comparaisons directes et indirectes ne retrouvait que trois discordances significatives entre les deux estimations¹¹⁹.

C'est pourquoi il est parfois souhaitable de présenter les deux sources d'information directe et indirecte. Une des limites de cette approche est qu'elle ne peut pas être appliquée à des grands réseaux d'essai ou ceux avec une boucle fermée.

1.2.4. Les comparaisons mixtes

Dans une méta-analyse en réseau, l'information concernant l'effet traitement d'une intervention peut provenir de deux sources directes ou indirectes. Assez fréquemment, l'information directe est très peu représentée ou souvent unique. Il est possible de combiner comparaisons directes et indirectes pour obtenir une comparaison mixte (« mixed treatment

comparaison ») en tenant compte de la totalité de l'information disponible. L'information directe et indirecte, par exemple par une moyenne pondérée par l'inverse de la variance :

$$Var \delta_{AB}^{combiné} = \left(\frac{\delta_{AB}^{ind}}{var(\delta_{AB}^{ind})} + \frac{\delta_{AB}^{dir}}{var(\delta_{AB}^{dir})} \right) / \left(\frac{1}{var(\delta_{AB}^{ind})} + \frac{1}{var(\delta_{AB}^{dir})} \right)$$

Cette méthode requière une cohérence (« consistency ») des données, c'est-à-dire les résultats des comparaisons directes et ceux des comparaisons indirectes doivent être concordants. Il s'ensuit une augmentation de la puissance et de la précision des estimations.

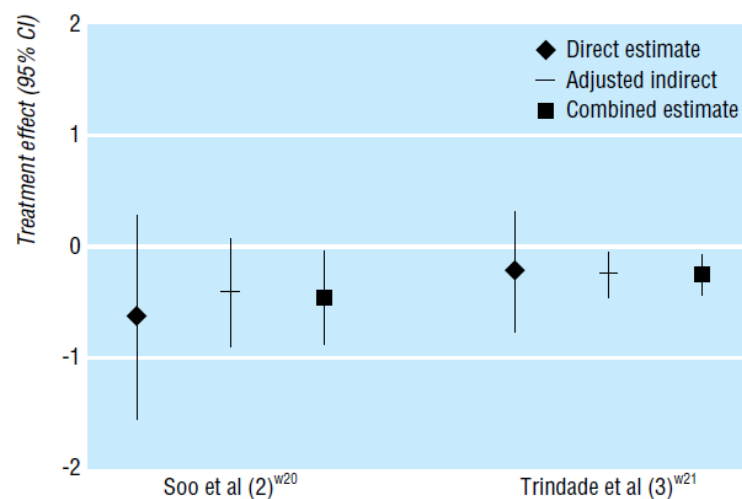


Figure 7 : Exemple de comparaison mixte. Adaptées de la source¹²⁰

Dans cet exemple, la meilleure estimation est celle qui combinait les résultats des comparaisons directes et indirectes.

Pour combiner l'information directe et indirecte, il est important de définir alors les paramètres basiques du réseau qui sont les mesures d'efficacité relative des traitements par rapport au traitement de référence. Dans un réseau de 4 traitements A, B, C et D avec le

traitement C comme traitement de référence, seule A et D n'ont pas été comparés directement (figure 8).

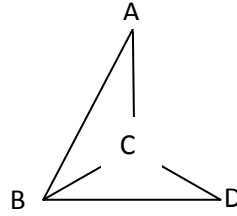


Figure 8 : Représentation graphique d'un réseau de 4 interventions

Les paramètres basiques sont les suivants : d_{AC} , d_{BC} et d_{DC} . Dans un réseau composé d'essais à deux bras, ou K traitements sont évalués, il faut K-1 paramètres basiques pour décrire complètement le réseau. Sous l'hypothèse de cohérence les autres paramètres dits fonctionnels se déduisent des précédents par les relations de cohérence:

$$\begin{aligned}d_{AB}^{dir} &= d_{AB}^{ind} \\d_{BD}^{dir} &= d_{BD}^{ind}\end{aligned}$$

Il n'y a pas d'équation de cohérence concernant d_{AD} car il y a aucune information directe pour cette comparaison.

et donc :

$$\begin{aligned}d_{AB}^{dir} &= d_{CB}^{dir} - d_{CA}^{dir} \\d_{BD}^{dir} &= d_{CD}^{dir} - d_{CB}^{dir}\end{aligned}$$

Le choix d'une intervention de référence et l'hypothèse de cohérence permet de réduire la dimensionnalité du réseau. Autrement dit, elles permettent de réduire le nombre d'estimation de $K \cdot (K-1)/2$ contrastes deux à deux entre interventions à l'estimation de (K-1) paramètres basiques, en synthétisant l'information directe et indirecte. Dans notre exemple,

il passe de 5 à 3 paramètres (initialement d_{AC} , d_{BC} , d_{AB} , d_{BD} et d_{DC} à par la suite d_{AC} , d_{BC} , d_{DC}). Il est possible de choisir d'autres paramètres basiques que d_{AC} , d_{BC} et d_{DC} . En effet, n'importe quel graphe reliant tous les points du réseau pourrait convenir, à condition que chacune de ses branches soit constituée d'information directe et qu'elle atteigne tous les points du réseau.

1.2.5. Puissance dans une méta-analyse en réseau

L'information indirecte permet d'améliorer la précision d'une méta-analyse en réseau par rapport à une méta-analyse standard. Prenons par exemple un réseau de 4 traitements A, B, C et D où seule l'information directe existe pour les relations d_{AB} , d_{AC} et d_{AD} , pour chaque comparaison nous ne disposons que d'un seul essai de variance 1. La précision de cet essai est également 1 (inverse de la variance). Une comparaison indirecte issue de deux essais de variance 1 aura pour variance 2 (Si l'on suppose que tous les essais ont la même variance, on obtient : somme des variances = 1+1) et pour précision 0,5. Chaque nouvelle comparaison indirecte informant notre comparaison étudiée apportera un gain de précision de 0,5 (leurs précisions se somment car les essais sont indépendants entre eux).

Cependant, le calcul de la puissance théorique d'une méta-analyse en réseau peut être relativement complexe pour plusieurs raisons :

- la puissance est difficile à déterminer pour le réseau de manière globale, mais peut être calculée pour chaque « branche » du réseau.
- Ce calcul dépend de nombreux facteurs : nombre d'essais et d'événements dans chaque branche du réseau et de la variabilité au sein de chaque branche (hétérogénéité) et l'existence ou non d'une incohérence.

Il est possible d'estimer la puissance à l'aide du modèle proposé par Thorlund et Mills¹⁰⁹, mais cette estimation nécessite des hypothèses fortes et multiples, concernant la variance des essais, leur taille et la présence ou non d'incohérence.

1.2.6. Les différentes approches statistiques

Plusieurs approches sont possibles pour estimer l'effet traitement dans une méta-analyse en réseau. L'approche bayésienne est la plus fréquemment utilisée^{99,121} elle repose sur la proposition de méta-analyses Bayésiennes faite par Higgins et Whitehead¹²². Cette approche fait appel à des techniques d'intégration par Monte Carlo des chaînes de Markov (MCMC) –notamment un type d'échantillonnage de Gibbs, une classe de méthodes MCMC. Dans ce modèle, il est important de définir une loi *a priori* informative ou non informative pour les paramètres basiques et pour la variance inter-essais. Dans le cadre des méta-analyses en réseau, il est très fréquent que le choix de la loi *a priori* soit non informative, cela implique que le choix du traitement de référence aura un impact négligeable sur les lois *a posteriori*; de plus les résultats seront similaires à ceux de l'approche fréquentiste. Il est techniquement possible de réaliser une méta-analyse en réseau avec une approche fréquentiste, mais elle est rarement utilisée. D'autres techniques ont été décrites, notamment l'ajustement de modèles linéaires généralisés ou des modèles de méta-régression¹²³⁻¹²⁷.

1.2.7. L'approche Bayésienne décrite par Lu et Ades

En 2004, Lu et Ades⁹⁹ propose un modèle pour combiner des comparaisons directes et indirectes. Ce modèle découle du précédent modèle proposé par Smith et ses collègues¹²⁸. Nous décrivons ici le modèle dans le cas d'un critère de jugement binaire. Cette approche se

construit de la façon suivante : si nous considérons l'intervention t comme le traitement étudié et l'intervention c comme le traitement contrôle dans l'essai i . r_i^t, r_i^c sont respectivement le nombre d'événements dans les groupes traitement et contrôle ; n_i^t, n_i^c sont respectivement le nombre de patient dans les groupes traitement et contrôle ; et p_i^t, p_i^c sont respectivement les probabilités de survenue de l'événement dans les groupes traitement et contrôle. δ_i décrit la différence observée entre le traitement et le contrôle, et μ_i la moyenne du taux d'événement de l'essai i . Le paramètre δ_i suit une loi normale centrée sur la différence vrai d d'effet des traitements t et c à savoir $\delta_i \sim N(d, \sigma^2)$. Il s'agit donc d'un modèle de méta-analyse aléatoire.

Avec ces conventions, le modèle est :

$$r_i^t \sim \text{Bin}(n_i^t, p_i^t)$$

$$r_i^c \sim \text{Bin}(n_i^c, p_i^c)$$

$$\text{logit}(p_i^t) = \mu_i + \delta_i / 2$$

$$\text{logit}(p_i^c) = \mu_i - \delta_i / 2$$

$$\delta_i \sim N(d, \sigma^2)$$

$$\delta_i = \text{logit}(p_i^t) - \text{logit}(p_i^c)$$

L'extension du modèle à K-comparaison et en choisissant l'intervention 1 comme l'intervention de référence dans l'ensemble du réseau, nous avons :

$$r_{ik} \sim \text{Bin}(n_{ik}, p_{ik})$$

$$\text{logit}(p_{i1}) = \mu_i - \delta_{i2}/K - \delta_{i3}/K - \dots - \delta_{iK}/K$$

$$\text{logit}(p_{i2}) = \mu_i + (K-1)\delta_{i2}/K - \delta_{i3}/K - \dots - \delta_{iK}/K$$

$$\text{logit}(p_{iK}) = \mu_i - \delta_{i2}/K - \delta_{i3}/K - \dots + (K-1)\delta_{iK}/K$$

1.2.8. Problèmes spécifiques aux essais multi-bras

Il est possible qu'un essai comporte plus de deux groupes expérimentaux (utilisant chacun un traitement différent) qui sont tous comparés au même groupe de référence. On parle alors d'essai multi-bras. Il est souvent fréquent qu'une méta-analyse en réseau inclut au moins un essai à plus de deux bras. Dans une revue publiée dans le Lancet en 2005, a montré que sur un échantillon de 519 ECRs publiés en Décembre 2000 et indexés dans Medline, 27% avaient plus de deux groupes expérimentaux¹²⁹. La prise en considération de ce type d'essai dans une méta-analyse pose le problème de la non indépendance statistique des résultats car ces essais utilisent le même groupe de référence. Comme pour la méta-analyse conventionnelle, on peut dédoubler les essais multi-bras en plusieurs essais à deux bras. Les calculs sont alors simplifiés, mais cela fait compter deux fois le bras contrôle des essais à trois bras (et trois fois pour ceux à quatre bras), ce qui donne donc un poids excessif à certains bras (contrôle) et introduit une corrélation dans les estimations des effets des deux traitements expérimentaux. Le risque est de biaiser l'estimation finale, surtout si le nombre de patients provenant d'essais multi-bras est important. L'approche de Lumley ne prend pas en compte les réseaux incluant des essais multi-bras. L'approche de Lu et Ades permet de tenir compte de ces essais. L'inconvénient est qu'il n'existe pas d'algorithme général pour paramétrer le modèle et que l'écriture du modèle n'est pas unique mais dépend du choix des paramètres basiques et de l'ordre des traitements. Higgins et al ont récemment décrit un modèle d'interaction traitement-plan d'expérience qui permet de modéliser les équations d'incohérence dans un cadre plus général¹³⁰.

1.2.9. Les hypothèses sous-jacentes de la méta-analyse en réseau

Dans une méta-analyse conventionnelle à effets aléatoires, nous faisons l'hypothèse d'interchangeabilité (« exchangeability »), c'est-à-dire que l'intensité des effets du traitement d'un essai à l'autre ne peut être différenciée¹³¹. La seule hypothèse supplémentaire dans une méta-analyse en réseau par rapport à la méta-analyse conventionnelle est que ces hypothèses d'interchangeabilité s'étendent à l'ensemble de tous les essais. Autrement dit, l'interchangeabilité suppose que si tous les essais avaient inclus toutes les interventions du réseau, chaque essai conduirait à estimer les mêmes quantités d'effet, échangeables, entre toutes les interventions du réseau^{132,133}. L'hypothèse de cohérence (« consistency ») découle de l'hypothèse d'interchangeabilité, elle désigne l'absence de discordance entre les résultats issus de comparaisons directes et ceux issus de comparaisons indirectes¹³⁴⁻¹³⁶. Comme pour la méta-analyse conventionnelle l'hypothèse d'homogénéité (« homogeneity ») est nécessaire, elle désigne l'absence de variabilité excessive dans les effets traitements entre les différents essais du réseau¹³⁴. Enfin, l'hypothèse de similarité (« similarity/transitivity ») est l'absence de différence excessive entre les comparateurs communs dans l'évaluation de effet traitement entre les deux séries d'essais A comparant C et ceux comparant B et C dans la comparaison indirecte^{137,138}.

1.2.9.1. Evaluation des hypothèses

1.2.9.1.1. Cohérence

La cohérence est l'absence de discordance entre les résultats issus de comparaisons directes et ceux issus de comparaisons indirectes, les causes d'incohérence doivent être analysées: limites méthodologiques dans les comparaisons directes ou indirectes (biais), diversité clinique importante entre patients inclus. Elle est partiellement évaluable, en effet la vérification de la cohérence entre les résultats des comparaisons directes et indirectes est

possible lorsqu'il existe une boucle fermée. Il n'y a pas de consensus universel sur la méthodologie à utiliser. Dans un réseau d'essais, plusieurs chemins sont possibles pour comparer deux traitements ; il est important qu'il n'y a pas d'incohérence entre ces différents chemins (par exemple entre plusieurs comparaisons indirectes). En cas d'incohérence, les résultats ne peuvent pas être combinés.

Plusieurs outils statistiques sont possibles pour évaluer ou ajuster l'incohérence^{123,139}:

- Le modèle de Bucher : Pour toute boucle fermée abc du réseau, on peut estimer d_{AB}^{dir} par comparaison directe et $d_{AB}^{ind} = d_{AC}^{dir} - d_{BC}^{dir}$ par comparaison indirecte ajustée. On peut ensuite estimer l'incohérence sur la boucle par $\varphi_{ABC} = d_{AB}^{dir} - d_{AB}^{ind}$ avec la variance associée $var(\varphi_{ABC}) = var(d_{AB}^{dir}) + var(d_{AB}^{ind})$. On peut alors tester l'hypothèse nulle de cohérence sur la boucle en utilisant la statistique $Z_{ABC} = \frac{\varphi_{ABC}}{var(\varphi_{ABC})}$. Dans des réseaux avec plusieurs boucles, cette approche peut être appliquée de manière répétée à toutes les boucles fermées du réseau. Néanmoins, le test est peu puissant et l'approche soulève un problème de multiplicité des tests^{140,141}.
- Les modèles d'incohérence : Plusieurs modèles ont été proposés pour ajuster une incohérence. Les approches les plus utilisées sont ceux de Lumley, Lu et Ades. Pour prendre en compte l'incohérence dans un réseau, on ajoute dans les équations de cohérence des paramètres supplémentaires, nommés paramètres d'incohérence (φ). Si nous reprenons le réseau de la figure 8, on obtient des équations du type :

$$d_{AB}^{dir} = d_{AB}^{ind} + \varphi_{CAB}$$

$$d_{BD}^{dir} = d_{BD}^{ind} + \varphi_{CAD}$$

Et donc :

$$d_{AB}^{dir} = d_{CB}^{dir} - d_{CA}^{dir} + \varphi_{CAB}$$

$$d_{BD}^{dir} = d_{CD}^{dir} - d_{CB}^{dir} + \varphi_{CAD}$$

Les paramètres d'incohérence permettent d'assouplir l'hypothèse de cohérence, et autorisent l'existence d'une discordance entre information directe et indirecte. Dans une méta-analyse en réseau Bayésienne, il faut préciser la distribution *a priori* des paramètres d'incohérence, généralement ils suivent une loi normale centrée et de variance $\sigma^2\varphi$.

Le nombre de paramètres d'incohérence dépend du nombre de boucles fermées indépendantes dans le réseau, c'est-à-dire des boucles ne pouvant pas être obtenues par la combinaison d'autres boucles. Dans notre exemple, les boucles «ABC», «CBDC» et «CABDC» ne sont pas indépendantes, la dernière étant la somme des deux premières. On appelle ICDF (pour « *inconsistency degrees of freedom* ») le nombre de degrés de liberté d'incohérence. Il correspond au nombre de paramètres d'incohérence à introduire dans un réseau pour prendre en compte l'incohérence. Pour une méta-analyse en réseau d'essais à deux bras, l'ICDF = C – K + 1 où C est le nombre de comparaisons pour lesquelles on dispose d'information directe, et K le nombre de traitements dans le réseau. En présence d'essai multi-bras, il n'existe pas de formule simple pour calculer l'ICDF. Dans notre exemple, l'ICDF valait donc deux. Plusieurs propositions peuvent être faites pour évaluer l'incohérence d'un réseau. On peut comparer l'ajustement des modèles avec et sans paramètres d'incohérence; déterminer la probabilité d'incohérence, à savoir la probabilité que la variance de la distribution *a posteriori* des paramètres d'incohérence soit supérieure à celle de l'hétérogénéité inter-essais: $p(\sigma^2\varphi > \sigma^2)$; calculer un rapport de variances (représentant la part de la variance d'incohérence sur la somme des variances d'incohérence et d'hétérogénéité); faire une *validation croisée* qui consiste à réaliser dans chaque boucle fermée une estimation directe et à la

comparer à chaque itération à celle obtenue sur le reste des données (incluant donc uniquement les comparaisons indirectes). On obtient alors une probabilité d'incohérence au sein de cette boucle, ensuite on répète l'opération pour chaque boucle. Cette dernière méthode est plus lourde, mais elle a l'avantage de permettre de localiser les comparaisons à l'origine de l'incohérence. La localisation de l'incohérence peut être réalisée par l'analyse des résidus des modèles avec et sans termes d'incohérence. On pourra revalider les résultats du modèle sans les essais à l'origine de l'incohérence.

1.2.9.1.2. Similarité

L'hypothèse de similarité évalue si il y a des différences dans la distribution des modificateurs de l'effet traitement entre les deux séries d'essais A comparant C et ceux comparant B et C dans la comparaison indirecte¹³⁴. Les modificateurs de l'effet traitement sont des variables qui influent sur l'effet traitement ou sont liées à la validité externe (caractéristiques des patients, des pathologies, des interventions, nombre des co-interventions, durée de suivi, des critères de jugement, risque de biais dans les études). Lorsque la distribution des modificateurs de l'effet traitement est la même entre les deux séries d'essais A comparant C et ceux comparant B et C dans la comparaison indirecte, on peut supposer que les effets traitements sont transitifs. La transitivité peut être considérée comme l'extension de la l'homogénéité clinique et méthodologique des comparaisons dans tous les essais qui comparent les interventions. Dans les réseaux complexes, l'hypothèse de transitivité doit tenir dans l'ensemble des estimations indirectes ou mixtes.

Celle-ci peut être approchée par des analyses en sous-groupes, des analyses de sensibilité ou une méta-régression^{126,142-144}. Il est intéressant dans une méta-analyse en réseau de présenter les caractéristiques cliniques et méthodologiques des essais pour chaque comparaison du réseau pour que le lecteur puisse juger de la similarité^{115,145,146}. Par

exemple dans une méta-analyse qui comparait l'efficacité du paracétamol combiné à la codéine au paracétamol seul dans la douleur postopératoire réalisée par Zhang et ses collègues^{120,147}. Basé sur les résultats de 13 essais, l'estimation directe montrait une différence significative (différence moyenne de 6,97, IC à 95% 3,56 à 10,37). La comparaison indirecte basée sur 43 essais contre placebo, montrait qu'il n'y avait pas de différence significative entre les interventions (- 1,16, IC à 95% - 6,95 à 4,64). La discordance entre l'estimation directe et indirecte était significative (P = 0,02). Cependant, lorsqu'on examinait les caractéristiques des essais, la plupart des essais de la comparaison directe utilisés 600 à 650 mg de paracétamol et 60 mg de codéine par jour, alors que la plupart des essais de la comparaison indirecte utilisés 300 mg de paracétamol et 30 mg de codéine par jour. L'analyse a été refaite en incluant uniquement les essais qui utilisaient 600 à 650 mg de paracétamol et 60 mg de codéine, l'estimation indirecte (5,72, IC à 95% - 5,37 à 16,81) n'était plus significativement différente de l'estimation directe (7,28, IC à 95% 3,69 à 10,87). Ainsi, les discordances entre directe et indirecte pourraient être expliquées par le fait que de nombreux essais utilisés de faibles doses de paracétamol (300 mg) et de codéine (30 mg). Cet exemple montre que l'évaluation des caractéristiques des essais impliqués dans des comparaisons indirectes est essentielle pour permettre de juger de la similarité.

1.2.9.1.3. Homogénéité

La méta-analyse en réseau offre une opportunité d'évaluer l'homogénéité ou hétérogénéité de tous les essais dans chacune des comparaisons. L'hétérogénéité statistique peut être évaluée par des tests tels que le Q de Cochran ou quantifier par des indicateurs tels que I^2 . L'évaluation de l'hétérogénéité peut se faire avec les mêmes méthodes utilisées dans les méta-analyses conventionnelles¹³⁴. Dans la méta-analyse en réseau les comparaisons impliquées sont multiples. Pour cette raison, l'hétérogénéité doit être évaluée à la fois au sein

de chaque essais et toutes les comparaisons du réseau. Par ailleurs, il est important de rapporter une mesure d'hétérogénéité, en particulier la variance inter-essais τ^2 . Cependant, le nombre d'essais dans chacune des comparaisons peut être insuffisant pour détecter statistiquement une éventuelle hétérogénéité. L'hétérogénéité peut être vérifiée sur le plan conceptuel (les différences dans les méthodes des études, les populations étudiées, paramètres, le suivi, et les co-interventions).

1.2.10. Présentation et évaluation des hypothèses dans les méta-analyses en réseau

Plusieurs études méthodologiques ont évalué comment les hypothèses étaient rapportées et examinées dans les comparaisons indirectes et les méta-analyses en réseau.

Song et ses collègues¹³⁴ ont évalué les méthodes et la présentation des hypothèses dans 88 comparaisons indirectes et méta-analyses en réseau publiées entre 2000 et 2007. Ils ont montré que l'évaluation des hypothèses dans les comparaisons indirectes et les méta-analyses en réseau était inappropriée ou négligée. L'hypothèse de similarité était mentionnée dans seulement 40 des 88 revues. L'hypothèse de cohérence n'était pas évaluée dans 18 des 30 revues qui combinaient directes et indirectes comparaisons.

Donegan et ses collègues¹⁴⁸ ont également évalué si les auteurs de comparaisons indirectes et de méta-analyses en réseau utilisaient des méthodes pour évaluer les hypothèses. Cette revue incluant 43 études publiées entre 1992 et 2007 a montré que les hypothèses essentielles à la méta-analyse en réseau n'étaient évaluées et discutées dans la majorité des revues. L'hypothèse de similarité était évaluée (en utilisant des analyses de sensibilité, ou des analyses en sous-groupe, ou des méta-régressions) dans 19 des 41 revues. Parmi les 17 revues qui combinaient des comparaisons directes et indirectes; 6 revues évaluaient l'incohérence.

Sobieraj et ses collègues¹⁴⁹ ont montré que les méthodes pour évaluer l'homogénéité et la cohérence n'étaient pas décrites dans l'ensemble des revues. Cette étude incluait 34 méta-analyses en réseau Bayésienne publiées entre 2006 et 2011.

Nikolakopoulou et ses collègues¹⁵⁰ ont montré également que les hypothèses ne sont pas adressées par les auteurs de méta-analyses en réseau. Cette étude portait sur 186 méta-analyses en réseau publiées entre 1997 et 2012. Les auteurs ont employé des méthodes inadéquates pour évaluer la cohérence dans 36 des 151 revues qui combinaient directes et indirectes comparaisons. Dans 67 des 151 revues, ne rapportaient pas de méthode pour évaluer la cohérence.

1.2.11. Limites de l'évaluation des hypothèses dans les méta-analyses en réseau

Bien que la validité de la méta-analyse en réseau dépende de la vérification des hypothèses d'homogénéité, de similarité et de cohérence, cette évaluation comporte certaines limites. Premièrement, le nombre d'essais dans chacune des comparaisons peut être insuffisant pour détecter statistiquement une éventuelle hétérogénéité ou une incohérence, ces tests statistiques sont soumis aux erreurs de type II (manque de puissance, lorsque l'information est insuffisante). Deuxièmement, lorsque de nombre de tests à effectuer est grand, comme dans un réseau très complexe, cela augmente la probabilité de détecter de faux positifs en particulier (erreurs de type I). Cela signifie que, dans la plupart des méta-analyses en réseau, l'absence de signaux –notamment d'incohérence- n'exclut pas totalement sa présence, ou à contrario trouver un signal peut parfois être un faux positif. Comme on l'a vu précédemment, la validité de la comparaison indirecte dépend également de la similarité des essais en termes de modificateurs de l'effet traitement, un problème souvent lié à la validité externe ou généralisation des résultats des essais. Par exemple dans la comparaison

indirecte ajustée de B versus C, les résultats des essais de A versus B doivent être généralisables aux patients dans les essais A versus C, de même pour les résultats des essais A versus C doit être généralisables aux patients essais A versus B. Si les modificateurs de l'effet traitement sont inégalement répartis, la généralisation des résultats entre les essais sera problématique. Il n'existe pas de méthode robuste permettant de certifier une similarité¹³⁸. L'évaluation de l'hypothèse de similarité est basée essentiellement sur un jugement subjectif pour décider si les distributions des modificateurs de l'effet traitement entre les études sont similaires. Toutefois, il est parfois difficile en pratique d'identifier les modificateurs de l'effet traitement. En effet, les modificateurs de l'effet traitement ne sont souvent pas rapportés dans les études et le nombre d'études disponibles par comparaison est souvent insuffisant¹⁵¹. Récemment des études ont essayé d'améliorer la validité des comparaisons indirectes et mixtes en ajustant sur les covariables des études¹⁵²⁻¹⁵⁴. Cependant, ces modèles de méta-régression sont souvent limités à cause du nombre d'études disponibles. Par ailleurs, des différences entre les études par exemple les caractéristiques des patients, des interventions, n'impliquent pas nécessairement une modification de l'effet traitement¹³⁶.

Une récente étude proposait un outil pour évaluer la similarité et la cohérence pour prédire les discordances entre les estimations provenant des comparaisons directes et indirectes¹³⁸. Cet outil comporte 3 éléments 1) l'évaluation de la similarité clinique (basé sur les caractéristiques principales des études) ; 2) l'évaluation de la similarité de la qualité des études ; 3) l'évaluation de la cohérence entre les estimations directes et indirectes. Les auteurs ont appliqué cet outil sur 94 revues systématiques Cochrane qui fournissaient les données pour comparer 2 interventions en utilisant les comparaisons directes et indirectes. Ils ont montré que les scores obtenus pour évaluer la similarité et la cohérence n'étaient pas associés à une discordance significative entre comparaison directe et indirecte.

1.2.12. Présentation des résultats des méta-analyses en réseau

Pour présenter les résultats d'une méta-analyse en réseau trois points sont essentielles :

- 1) Les résultats de la méta-analyse en réseau doivent être présentés sous la forme de quantités d'effet accompagné d'un intervalle de confiance (ou d'un intervalle de crédibilité) autour de la quantité d'effet vraie. Pour K interventions, on estime les $K \times (K-1)/2$ quantités d'effet (rapports des cotes dans le cas d'un critère de jugement binaire) pour toutes les comparaisons possibles deux à deux.
- 2) Il est important de rapporter une mesure d'hétérogénéité, en particulier la variance inter-essais τ^2 .
- 3) Il est utile de considérer les comparaisons deux à deux pour lesquelles il existe une source directe et de comparer les quantités d'effet estimées via la méta-analyse en réseau à celles obtenues via la méta-analyse conventionnelle. En examinant chaque estimé et l'intervalle de confiance associé, on peut ainsi évaluer la manière dont l'information indirecte influence ou non l'information directe.

1.2.12.1. Les formats pour présenter les résultats d'une méta-analyse en réseau

Plusieurs formats existent pour rapporter les résultats d'une méta-analyse en réseau: une matrice, graphiquement sur un Forest plot ou sur le schéma de la géométrie du réseau, un tableau.

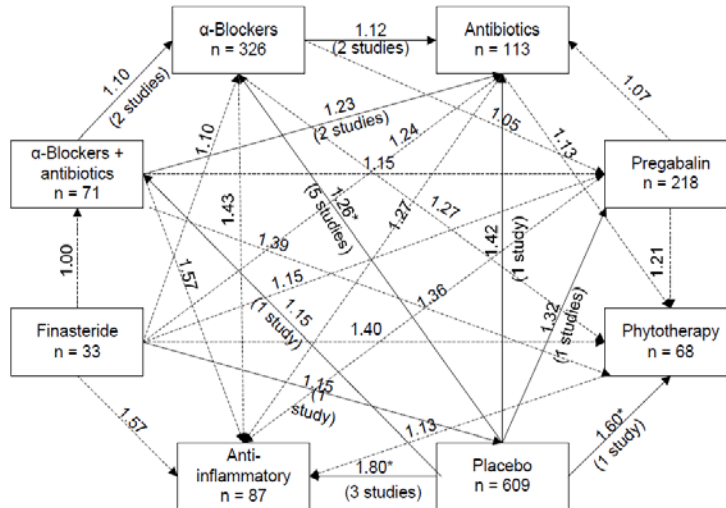


Figure 9 : Représentation des résultats sur le schéma de la géométrie du réseau Adaptée de la source¹⁵⁵

Les lignes en plein représentent les comparaisons directes entre les traitements, celles en pointillés les comparaisons indirectes. Le nombre au dessus de la ligne indique l'effet traitement; le nombre entre parenthèse se réfère au nombre d'études.

Mixed treatment comparison				
Standard Meta-Analysis	Intervention A	OR _{A-B_MTC} (95% CrI)	OR _{A-C_MTC} (95% CrI)	OR _{A-D_MTC} (95% CrI)
	OR _{A-B_MA} (95% CrI)	Intervention B	OR _{B-C_MTC} (95% CrI)	OR _{B-D_MTC} (95% CrI)
	OR _{A-C_MA} (95% CrI)	OR _{B-C_MA} (95% CrI)	Intervention C	OR _{C-D_MTC} (95% CrI)
	Not calculable	OR _{B-D_MA} (95% CrI)	OR _{C-D_MA} (95% CrI)	Intervention D

Tableau 5 : Matrice des résultats dans une méta-analyse en réseau. Adaptée de la source¹⁵⁶

Toutes les estimations deux à deux sont représentées. Le triangle supérieur représente les résultats issus de la méta-analyse en réseau, le triangle inférieur représente les résultats issus des comparaisons directes.

1.2.12.2. Classement des interventions

Il est également possible de présenter un classement des interventions de la plus efficace à la moins efficace en utilisant des probabilités. La probabilité que l'intervention j soit la k -ième meilleure intervention parmi les K interventions évaluées est définie comme $P(j = k)$, $k=1, 2, \dots, K$. Nous appelons ces probabilités les probabilités de rang. Dans l'approche bayésienne, à chaque itération dans la méthode de MCMC, chaque traitement peut être classé en fonction de la quantité d'effet estimée relativement à l'intervention de référence c du réseau

$$\text{Rang}(j) = \begin{cases} \sum_{k=1}^K I(d_{kc} \leq d_{jc}) & \text{si le critère est négatif} \\ K + 1 - \sum_{k=1}^K I(d_{kc} \leq d_{jc}) & \text{si le critère est positif} \end{cases}$$

Ainsi, on peut soit utiliser les probabilités que chaque traitement soit classé le premier, le deuxième, puis le troisième ; ou les probabilités cumulées qui représentent les probabilités qu'une intervention soit parmi les meilleurs interventions. Dans l'approche fréquentiste, ces probabilités peuvent être estimées via la méthode de ré-échantillonnage par exemple par bootstrap. Enfin, il est possible d'estimer le rang médian de chaque intervention et l'intervalle de confiance (crédibilité). Pour chaque intervention et chaque rang, la somme de ces probabilités vaut 1. Ces probabilités peuvent être représentées graphiquement en utilisant une courbe des probabilités de rang («rankograms») ou une courbe des probabilités de rang cumulées pour chaque traitement. Enfin, on peut calculer la surface sous la courbe des probabilités cumulées («surface under the cumulative ranking» SUCRA) pour chaque traitement: cette aire vaut 1 lorsqu'il est certain que le traitement est la meilleure et 0 lorsqu'il est certain qu'il est le pire¹⁵⁷.

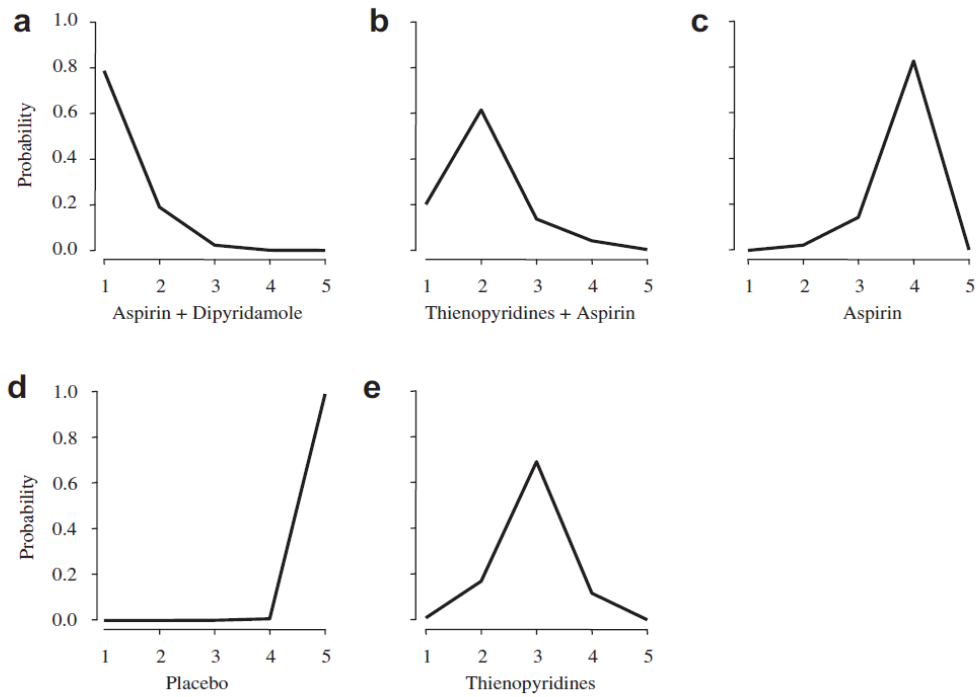


Figure 10 : Courbe des probabilités des rangs. Adaptée de la source¹⁵⁷

En abscisse est représentée les cinq rangs possibles et en ordonné la probabilité d'un traitement d'atteindre chaque rang. Par exemple, La combinaison Aspirine plus Dipyridamole a une probabilité forte d'être classée à la première place, et le placebo a une probabilité forte d'être à la dernière place.

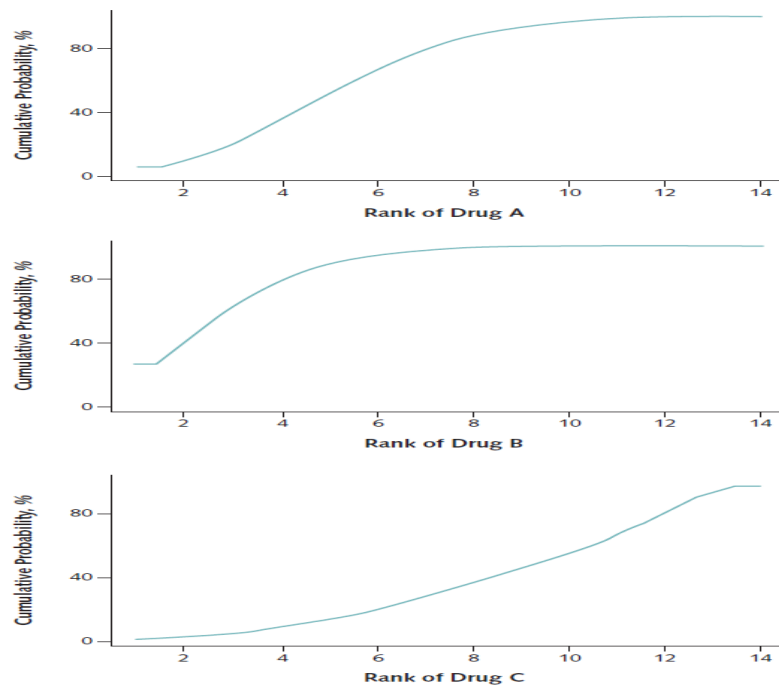


Figure 11 : Courbe des probabilités de rang cumulées. Adaptée de la source¹⁰³

Ces « rankograms » montrent la distribution des probabilités cumulées de chaque traitement d'être classé à chacune des 14 rangs possibles. Par exemple, les traitements A et B ont une probabilité plus forte d'être parmi les meilleurs traitements, et le traitement C a une probabilité plus forte d'être parmi les pires.

Treatment	SUCRA (%)	Median rank (95% credible interval)	Probability to be no worse than 1.1-fold
Aspirin + dipyridamole	94	1 (1, 2)	Reference
Thienopyridines + aspirin	74	2 (1, 3)	51
Thienopyridines	52	3 (2, 4)	18
Aspirin	30	4 (3, 4)	2
Placebo	0	5 (5, 5)	0

Tableau 4 : Rangs des différents traitements antiplaquettaire. Adaptée de la source¹⁵⁷

Ce tableau indique les rangs médians (avec les intervalles crédibles à 95%), les probabilités de ne pas être le pire. Les valeurs « Sucra » sont représentées en pourcentage.

1.2.13. Les différents biais dans la méta-analyse en réseau

La méta-analyse en réseau est une extension de la méta-analyse classique¹⁵⁸, comme pour celle-ci, elle est probablement soumise au même type de biais. Elle doit se réaliser selon les méthodes rigoureuses des revues systématiques pour limiter les risques de biais.

Le biais de publication est une des principales menaces à la validité des résultats^{50,159-161}.

De plus, il n'existe pas de test pour le détecter ou le corriger dans la méta-analyse en réseau.

Récemment une étude publiée dans *Plos One* évaluait l'impact du biais de publication.

Cette étude utilisait les données provenant de la FDA de 74 ECRs évaluant 12 antidépresseurs contre placebo et leurs 51 publications correspondantes. Pour chaque ensemble de données (données de la FDA et données publiées), les auteurs ont estimé à l'aide d'une méta-analyse en réseau tous les effets traitements deux à deux (66 comparaisons possibles), les probabilités d'être le meilleur médicament et le classement des 12 antidépresseurs. Ils ont comparé les résultats de la méta-analyse en réseau avec les 51 essais publiés et ceux avec les 74 essais enregistrés dans la FDA. Cette étude a mise en évidence des différences dans les estimations de l'effet traitement et dans le classement des traitements selon les sources des données (FDA ou données publiées). Ils ont aussi montré que les biais de publication affectant un seul médicament peuvent influencer le classement de toutes les autres molécules.

Un autre problème de la méta-analyse en réseau est le choix des interventions à inclure dans le réseau. Par exemple : Faut-il inclure tous les essais contre placebo ? Cela peut aboutir à des conclusions différentes. Il a été montré que des méta-analyses en réseau sur le même sujet peuvent avoir des conclusions différentes essentiellement du aux critères d'éligibilités¹⁶².

Cette question du choix des interventions à inclure dans le réseau a été illustrée dans une étude publiée dans *BMJ* portant sur 18 réseaux d'essais incluant 757 ECRs et 1036 701

patients¹⁶³. Les auteurs ont évalué l'impact de l'exclusion d'un ou plusieurs traitements sur l'estimation des effets traitements et sur le classement des interventions dans le réseau. Pour chaque méta-analyse en réseau, ils ont comparé les effets traitements et les classements de la méta-analyse en réseau de toutes les interventions avec une méta-analyse excluant l'intervention ayant le plus grand impact dans le réseau, une méta-analyse excluant des traitements actifs, une méta-analyse excluant le placebo ou pas de traitement comme comparateur. Parmi les 18 réseaux éligibles, 11 n'incluaient pas tous les comparateurs et seulement 10 incluaient un comparateur placebo ou pas de traitement. Dans 7 réseaux, il y avait au moins une intervention dont le retrait a provoqué une légère modification des effets traitements, et modifié le top trois des interventions dans 9 réseaux. L'exclusion du placebo ou pas de traitement comme comparateur a provoqué des changements importants sur les effets traitements (un changement en moyenne 1,16 à 3,10 fois) pour 4 des 10 réseaux qui avaient initialement un placebo ou pas de traitement comme comparateur. En conclusion l'exclusion de traitements du réseau peut provoquer des modifications importantes des résultats de la méta-analyse en réseau.

1.2.14. Les recommandations pour rapporter et évaluer les méta-analyses en réseau

Une présentation adéquate et précise des rapports de méta-analyse en réseau est essentielle pour que les professionnels des soins de santé puissent comprendre ce qui a été réalisé afin de traduire correctement les résultats de la recherche en pratique clinique. Actuellement des recommandations comme le PRISMA ne sont pas encore disponibles. Une extension du PRISMA à la méta-analyse en réseau est en cours de préparation.

De nombreuses revues ont essayé de décrire les éléments clés devant figurer dans un rapport de méta-analyse en réseau^{101,135,136,164}. Cependant ces recommandations sont basées

essentiellement sur les aspects statistiques de la méta-analyse en réseau. Quelques uns de ces documents adressent la question de la recherche bibliographique, des critères d'éligibilités et de l'évaluation du risque de biais dans les études incluses. Par exemple le groupe ISPOR (acronyme de the International Society for Pharmacoeconomics and Outcomes Research) a publié trois documents pour la conduite et l'interprétation des méta-analyses en réseau^{134,135,156}. Le groupe ISPOR propose une liste d'éléments importants de chaque section de la revue, notamment le titre, l'introduction, les méthodes, les résultats, la discussion devant figurer dans un rapport de méta-analyse en réseau. Récemment ils ont développé un questionnaire portant sur 22 questions pour évaluer si les résultats de la méta-analyse en réseau fournissent une réponse valide. Ces questions s'intéressent à la quantité de preuve disponible, les méthodes d'analyse, la qualité de la présentation et de la transparence, l'interprétation des données, et les conflits d'intérêts.

Par ailleurs, un groupe de travail du GRADE a développé une approche en quatre étapes pour évaluer la qualité des preuves d'une méta-analyse en réseau en se basant sur : 1) identifier les estimations des effets traitements provenant des sources directes et indirectes pour chaque comparaison ; 2) puis évaluation de la qualité des estimations des effets traitements provenant des sources directes et indirectes pour chaque comparaison ; 3) identifier les estimations des effets traitements provenant de la méta-analyse en réseau ; 4) évaluation de la qualité des estimations des effets traitements provenant de la méta-analyse en réseau pour chaque comparaison¹⁶⁵.

Tableau 6: Les éléments devant figurer dans un rapport de méta-analyse en réseau selon

ISPOR

Report section	Checklist Item	What to look for in the paper
Introduction	Are the rationale for the study and the study objectives stated clearly?	A clear rationale for the review A clear objective or research question that pertains to the network meta-analysis
Methods	Does the methods section include the following? Description of eligibility criteria Information sources Search strategy Study selection process Data extraction (validity/quality assessment of individual studies) Are the outcome measures described? Is there a description of methods for analysis/synthesis of evidence? Do the methods described include the following? Description of analyses methods/models Handling of potential bias/inconsistency Analysis framework	A systematic review of the literature in accordance with Centre for Reviews and Dissemination (CRD) guidelines and PRISMA [52,54,55] Justification of outcome measures selected for analysis Description and justification of statistical model(s) used: multiple meta-analysis of pairwise comparisons vs. network meta-analysis models; fixed- vs. random-effects models; models without or with covariate (interactions) Description of whether analyses were performed with a frequentist or Bayesian approach Description of how possible bias/inconsistency was evaluated (either qualitative or quantitative, e.g., comparison of direct evidence with the indirect evidence). If meta-regression models are used, rationale for selection of covariates in models Description of relative-effect estimates used for presentation of findings (e.g., odds ratio, relative risk, hazard ratio, difference in change from baseline) Description of whether relative-effect measures were transformed into expected (absolute) outcomes (e.g., proportion of responders)
	Are sensitivity analyses presented?	Rationale for and description of sensitivity analyses Studies included Prior distributions for model parameters in Bayesian framework
Results	Do the results include a summary of the studies included in the network of evidence? Individual study data? Network of studies? Does the study describe an assessment of model fit? Are competing models being compared? Are the results of the evidence synthesis (ITC/MTC) presented clearly? Sensitivity/scenario analyses	Description of results of study identification and selection process Table/list of studies with information regarding study design and patient characteristics (that might act as effect modifiers); these are important to judge potential similarity/consistency issues Figure of network of studies Table with raw data by study and treatment as used for the analysis/model. (Optionally present relative effects of available direct comparisons of each study) Justification of model results Table/ figure with results for the pairwise comparisons as obtained with analyses; Point estimates and measure of uncertainty (95% CIs) In Bayesian framework, probability to reflect decision uncertainty (i.e., probability of which treatment is best if multiple treatments are being compared and probability that one treatment is better than the comparator) Description of (different) findings with sensitivity/scenario analysis
Discussion	Does the discussion include the following? Description/summary of main findings Internal validity of analysis External validity Implications of results for target audience	Summary of findings Internal validity (individual trials, publication bias, differences across trials that might violate similarity and consistency assumptions) Discussion regarding generalizability of findings (given patient population within and across trials in network.) Interpretation of results from a biological and clinical perspective

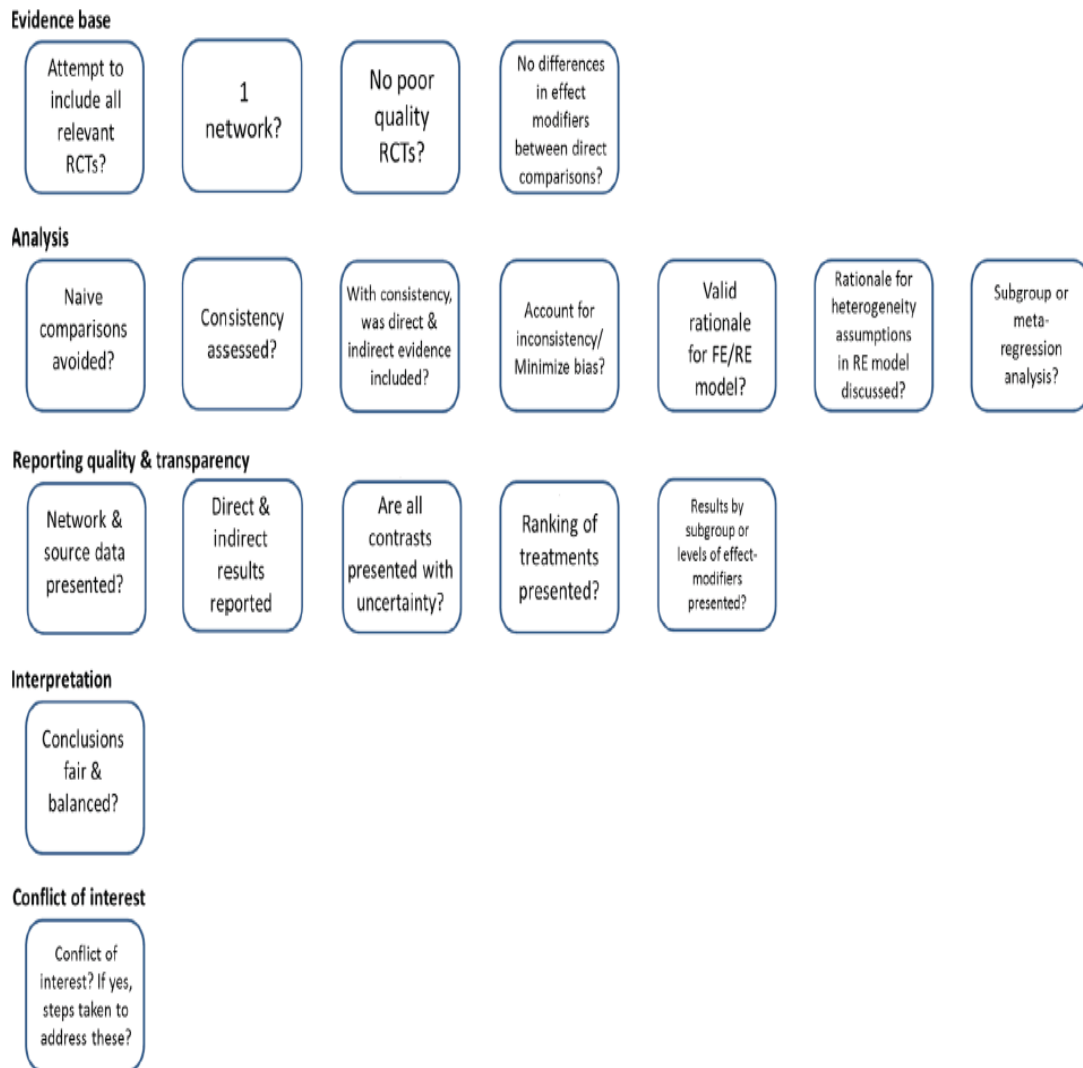


Figure 12 : Les différents domaines de l'évaluation de la validité d'une méta-analyse en réseau selon ISPOR. Adaptée de la source¹⁶⁴

FE : fixed effect, RCTs : randomised controlled trials; RE: random effects.

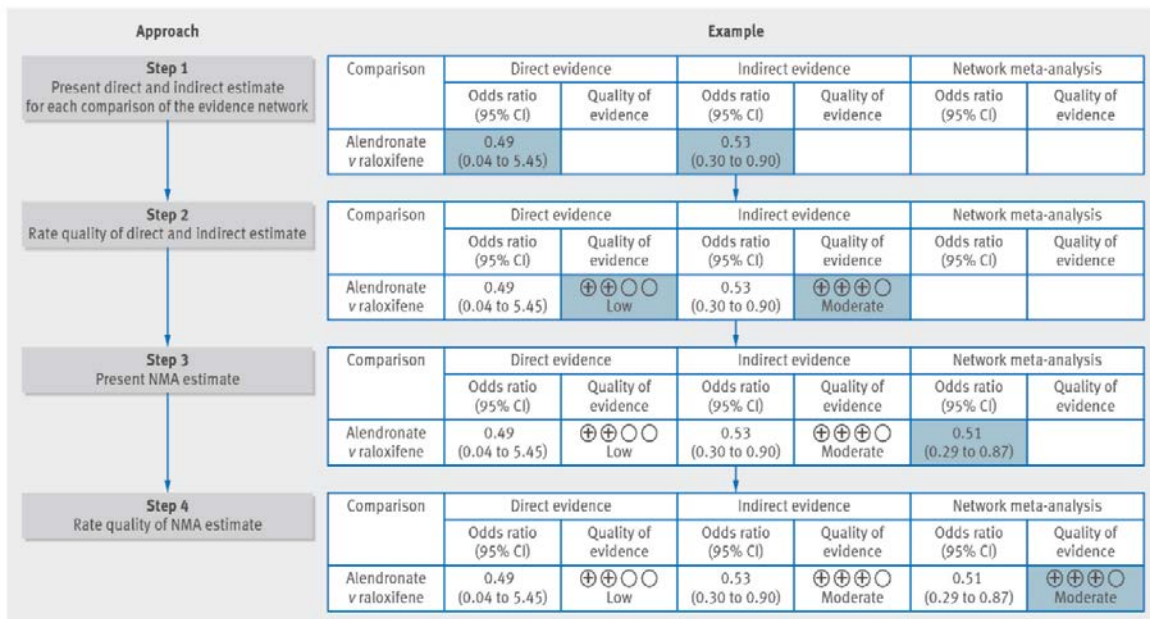


Figure 13 : L'approche GRADE pour évaluer la qualité des méta-analyses en réseau.

Adapté de la source¹⁶⁵

1.3. Justification de la démarche de la thèse

Comme nous l'avons vu précédemment, la fréquence d'utilisation des méta-analyses en réseau a augmenté notamment depuis le milieu des années 2000^{103,134,166}. Ainsi que le nombre de publications méthodologiques pour conduire des comparaisons indirectes et des méta-analyses en réseau s'est multiplié^{100,132,134,148,167}. Cette rapide évolution des méthodes a soulevé des inquiétudes sur la conduite et la présentation des méta-analyses en réseau. De nombreuses revues ont tenté d'exposer les éléments clés pour améliorer la mise en œuvre des méta-analyses en réseau^{102,132,135,136,168}. Elles se sont principalement intéressées aux méthodes statistiques et à la vérification des hypothèses.

Les méta-analyses en réseau restent des méta-analyses, elles sont soumises aux mêmes risques de biais¹⁵⁸ et doivent se réaliser selon les méthodes rigoureuses des revues systématiques. Cependant, aucune étude ne s'était intéressée aux éléments essentiels de la réalisation d'une revue systématique dans les méta-analyses en réseau comme les objectifs, les critères d'éligibilités, la recherche de la littérature, l'évaluation du risque de biais et le biais de publication ou les éléments à produire dans un rapport de méta-analyse pour présenter les résultats. Une présentation inadéquate des rapports de méta-analyses en réseau peut avoir un impact sur l'interprétation des résultats.

Pour répondre à ces questions, notre travail sera organisé en deux temps. Nous avons réalisé tout d'abord une revue exhaustive de littérature afin d'identifier toutes les méta-analyses en réseau publiées. Dans le premier travail nous avons déterminé si les méta-analyses en réseau suivaient les principales recommandations méthodologiques pour la conduite et la présentation des revues systématiques. Enfin dans le deuxième travail nous sommes intéressés à la présentation des résultats des méta-analyses en réseau. Nous avons évalué quelles caractéristiques étaient rapportées dans les publications de méta-analyse en

réseau comme la géométrie du réseau, les estimations de l'effet traitement issues des comparaisons directes, indirectes et mixtes.

Pour chaque travail, nous résumerons les méthodes ainsi que les principaux résultats avant de présenter le tiré-à-part de la publication. Finalement le chapitre 4 de cette thèse sera consacré à la discussion de nos résultats et les perspectives de nos travaux de recherche.

2. Chapitre : Evaluation des composantes méthodologiques des revues systématiques dans les publications de méta-analyse en réseau

2.1. Justification de l'étude

Précédemment plusieurs études s'étaient intéressées à la manière dont les comparaisons indirectes étaient conduites et rapportées^{134,148,161,167}. Ces études étaient axées sur les méthodes statistiques et la vérification des hypothèses (cohérence, homogénéité, similarité). Certains éléments importants liés à la conduite et la présentation des revues systématiques n'ont pas été étudiés. Les méta-analyses en réseau sont principalement des méta-analyses, et donc sont soumises aux mêmes risques de biais¹⁵⁸. Ainsi elles doivent être effectuées en respectant les règles méthodologiques des revues systématiques pour limiter les risques de biais.

2.2. Objectifs

L'objectif de ce premier travail était de déterminer si les méta-analyses en réseau suivent les principales recommandations pour la conduite et la présentation des revues systématiques.

2.3. Méthodes

Nous avons réalisé une revue systématique. La première étape de cette étude a été d'identifier toutes les méta-analyses en réseau comparant l'efficacité clinique d'au moins trois interventions (en excluant les réseaux de trois interventions sans boucle fermée) indexées dans Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, Medline, et Embase publiées entre 1960 et le 12 juillet 2012. Nous avons également examiné toutes les références d'articles méthodologiques sur les méta-analyses en réseau et recherché des articles citant les papiers statistiques sur la méta-

analyse en réseau via le Web of Science ou SCOPUS. Deux lecteurs indépendants ont sélectionné sur le titre et le résumé les publications, ensuite tous les textes intégraux des articles ont été retrouvés

Pour chaque méta-analyse en réseau sélectionnée, nous avons extrait les données suivantes par deux lecteurs indépendants :

1) les caractéristiques générales comme le nom du journal, l'année de publication, le pays de l'auteur correspondant, le domaine médical, le type de financement et le type d'interventions évalué.

2) Nous avons aussi évalué des informations concernant la structure du réseau comme le nombre d'interventions, le nombre de comparaisons effectuées par au moins un ECR (comparaison directe), et le nombre d'ECRs inclus dans le réseau.

En se basant sur les recommandations de PRISMA⁸² et d'AMSTAR^{88,89}, nous avons évalué si les éléments liés à la conduite et la présentation des revues systématiques ont été rapportées ou non. Les éléments étaient les suivantes: les objectifs, les critères d'éligibilités, le critère de jugement principal, la recherche bibliographique, les méthodes pour la sélection et l'extraction des données, les méthodes pour évaluer et incorporer les risques de biais, les méthodes pour évaluer le biais de publication, les caractéristiques des essais inclus dans le réseau (descriptions des interventions, population, design de l'étude), les hypothèses nécessaires à la méta-analyse en réseau, la déclaration des conflits d'intérêts.

Lorsque certains éléments étaient bien rapportés, nous avons évalué si les procédés utilisés étaient adéquats. Nous avons considéré les procédés suivants comme inadéquats: une recherche électronique sur une seule base de données bibliographiques, restreindre la sélection des études à ceux publiés, une sélection et une extraction des données par un seul lecteur.

Finalement, pour résumer les données, nous avons construit deux critères composites basés sur les recommandations de Li et ses collègues⁹⁰ et MECIR¹⁹ pour évaluer la présentation ou la conduite inadéquate. Le premier critère composite basé sur Li et ses collègues était considéré comme inadéquate s'il avait un des items suivants : une recherche bibliographique non rapportée ou inadéquate ; l'évaluation du risque de biais non présentée. Le deuxième critère composite composé de 7 items obligatoires recommandés par MECIR considéré comme inadéquate si il avait un des items suivants : les objectifs non rapportés ; les critères d'éligibilités non rapportés ; une recherche bibliographique non rapportée ou inadéquate ; la sélection et l'extraction des données non rapportée ou inadéquate ; les caractéristiques des études inclus dans le réseau non présentées ; l'évaluation du risque de biais non présentée.

2.4. Résultats

Nous avons identifié 121 méta-analyses en réseaux représentant de nombreux domaines médicaux. Les 121 méta-analyses de réseau étaient publiées dans 75 différents journaux, avec 55 (45%) des méta-analyses de réseau publiées dans des journaux généralistes et 66 (55%) dans des journaux spécialisées.

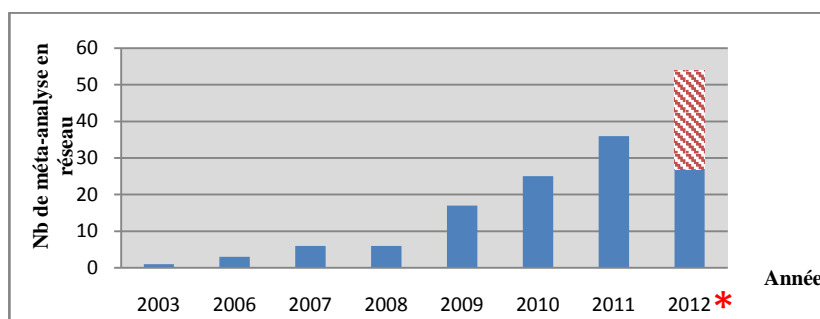


Figure 14 : Nombre de publications de méta-analyse en réseau par an.

*Les données pour le 2ème semestre 2012 ont été imputées.

Parmi les 121 méta-analyses en réseau, 100 (83%) portaient sur l'évaluation d'interventions pharmacologiques et 11 (9%) sur des interventions non pharmacologiques. Le nombre médian d'interventions évaluées dans une méta-analyse en réseau était de 7 (intervalle interquartile 5-9), et le nombre médian d'ECRs inclus dans une méta-analyse en réseau était de 22 (15-40).

La stratégie de recherche électronique n'était pas décrit dans 88 (73%) méta-analyses réseau. Dans 36 (30%) méta-analyses en réseau, le critère de jugement principale n'étaient pas clairement identifiable ; 61 (50%) méta-analyses en réseau ne décrivaient pas de méthode pour évaluer le risque de biais des études incluses dans le réseau et 103 (85%) ne décrivaient pas de méthode pour évaluer le biais de publication.

Enfin pour le critère composite basé sur les recommandations de Li et ses collègues, 87 (72%) méta-analyses réseaux avaient une présentation inadéquate ou conduite inadéquate de la recherche bibliographique ou n'évaluaient pas le risque de biais des essais. Cette mesure ne différait pas selon le type de journal (journaux généralistes, 69% (intervalle de confiance à 95% [57% ; 81%]) ; journaux spécialisée, (IC à 95% 74% [63% à 85%]) ; $P = 0,5$) ou la source de financement (financement public, 67% (74% à 79%); financement privé, 79% (67% à 90%); $P = 0,2$). Pour le second critère composite basé sur 7 items obligatoires recommandés par la collaboration Cochrane, 120 (99%) méta-analyses en réseaux avaient soit une présentation inadéquate ou conduite inadéquate d'au moins un des 7 éléments méthodologiques des revues systématiques.

2.5. Discussion

Nos résultats montrent que les éléments méthodologiques essentiels aux revues systématiques sont manquants dans la plupart des publications de méta-analyse en réseau. Ces résultats ne diffèrent pas selon le type de publication (revue généraliste ou spécialisée) ou selon le type de financement (public ou privé).

Nous avons mis en évidence que plus de la moitié des méta-analyses en réseau ne décrivaient pas de recherche exhaustive de la littérature. Hors un des éléments clés des revues systématiques est d'identifier toutes les études pertinentes. La recherche documentaire doit être la plus exhaustive possible pour maximiser la probabilité de retrouver toutes les études pertinentes et minimiser le biais de publication^{24,26,27}.

Le biais de publication pourrait affecter les résultats des méta-analyses en réseau^{169,170}. Ce facteur est d'autant plus important que nous n'avons pas de méthode validée pour détecter ou exclure le biais de publication dans le réseau. Bien que le biais de publication puisse avoir un effet important sur les conclusions de méta-analyse en réseau, il est négligé dans la majorité des publications de méta-analyse en réseau, 85% méta-analyses en réseau ne présentaient pas de méthode pour évaluer le biais de publication.

La qualité des études inclus pourrait influencer sur les résultats des méta-analyses en réseau¹⁷¹. Cependant, 58% de nos méta-analyses en réseau ne présentaient pas l'évaluation du risque de biais.

Notre étude comporte certaines limites. Premièrement, évaluer la qualité de réalisation des méta-analyses en réseau uniquement dans les rapports publiés pourrait ne pas être fiable, comme cela a été démontré pour les essais¹⁷². Les auteurs auraient pu utiliser des méthodes adéquates, mais des détails importants sont omis dans les publications ou des informations

clés peuvent avoir été supprimées au cours du processus de publication. Cependant, nous étions dans la position du lecteur, qui ne peut évaluer uniquement ce qui a été rapporté. De plus nous avons exclu les rapports issus des agences d'évaluation, en effet il se pourrait que la qualité de ces rapports soit meilleure que ceux analysés.

Un des mécanismes qui peut expliquer ces résultats est l'absence de recommandations claires comme PRISMA pour rapporter les méta-analyses en réseau. L'absence de certains éléments importants liés aux revues systématiques dans les méta-analyses en réseau peut soulever des doutes sur la validité de leurs conclusions. La réalisation d'une méta-analyse en réseau en dehors d'une revue systématique, sans une recherche bibliographique approfondie et sans une évaluation des risques de biais des essais, augmente le risque de produire des résultats biaisés. Ainsi les défauts potentiels dans la conduite des méta-analyses en réseau pourraient affecter leurs résultats. Des lignes directrices claires sont nécessaires pour améliorer la qualité de présentation et de conduite des méta-analyses en réseau.

Article

Aïda Bafeta, Ludovic Trinquart, Raphaèle Seror, Philippe Ravaud.
Analysis of the systematic reviews process in reports of network meta-analyses: methodological systematic review. *British Medical Journal* 2013;347:f3675

RESEARCH

Analysis of the systematic reviews process in reports of network meta-analyses: methodological systematic review

 OPEN ACCESS

Aïda Bafeta *PhD student*^{1,2}, Ludovic Trinquart *senior statistician*^{1,2,3,4}, Raphaèle Seror *associate professor of rheumatology*^{1,2,4}, Philippe Ravaut *professor of epidemiology and director*^{1,2,3,4,5}

¹INSERM U738 Research Unit, Paris, France; ²Assistance Publique-Hôpitaux de Paris, Centre d'Epidémiologie Clinique, Hôpital Hôtel-Dieu, 75004 Paris, France; ³Université Paris Descartes-Sorbonne Paris Cité, Paris, France; ⁴French Cochrane Centre, Paris, France; ⁵Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, USA

Abstract

Objective To examine whether network meta-analyses, increasingly used to assess comparative effectiveness of healthcare interventions, follow the key methodological recommendations for reporting and conduct of systematic reviews.

Design Methodological systematic review of reports of network meta-analyses.

Data sources Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, Medline, and Embase, searched from inception to 12 July 2012.

Review methods All network meta-analyses comparing clinical efficacy of three or more interventions based on randomised controlled trials, excluding meta-analyses with an open loop network of three interventions. We assessed the reporting of general characteristics and key methodological components of the systematic review process using two composite outcomes. For some components, if reporting was adequate, we assessed their conduct quality.

Results Of 121 network meta-analyses covering a wide range of medical areas, 100 (83%) assessed pharmacological interventions and 11 (9%) non-pharmacological interventions; 56 (46%) were published in journals with a high impact factor. The electronic search strategy for each database was not reported in 88 (73%) network meta-analyses; for 36 (30%), the primary outcome was not clearly identified. Overall, 61 (50%) network meta-analyses did not report any information regarding the assessment of risk of bias of individual studies, and 103 (85%) did not report any methods to assess the likelihood of publication bias. Overall, 87 (72%) network meta-analyses did not report the literature search, searched only one database, did not search other sources, or did not report an assessment of risk of bias of individual studies. These

methodological components did not differ by publication in a general or specialty journal or by public or private funding.

Conclusions Essential methodological components of the systematic review process—conducting a literature search and assessing risk of bias of individual studies—are frequently lacking in reports of network meta-analyses, even when published in journals with high impact factors.

Introduction

Assessing the comparative effectiveness of many or all available interventions for a clinical indication is challenging.¹ Direct evidence from head to head trials is often lacking. Indirect comparison between two interventions can still be estimated from the results of randomised controlled trials, without jeopardising the randomised comparisons within each trial, if each intervention is compared with an identical common comparator.^{6,7} When both direct and indirect comparisons are available, the two sources of information can be combined by using network meta-analyses—also known as multiple treatments meta-analyses or mixed treatment comparison meta-analyses.^{8,9} These methods allow for estimating all possible pairwise comparisons between interventions and placing them in rank order.

In the past few years, network meta-analyses have been increasingly adopted for comparing healthcare interventions.¹⁰⁻¹² Network meta-analyses are attractive for clinical researchers because they seem to respond to their main concern: determining the best available treatment. Moreover, national agencies for health technology assessment and drug regulators increasingly use such methods.¹³⁻¹⁶

Extra material supplied by the author (see <http://www.bmj.com/content/347/bmj.f3675?tab=related#webextra>)

Web appendix: Supplementary material

Correspondence to: A Bafeta aida.bafeta@htd.aphp.fr

Several reviews previously evaluated how indirect comparisons have been conducted and reported.^{7 10-12 17 18} These reviews focused on checking validity assumptions (based on homogeneity, similarity, and consistency) and did not assess the key components of the systematic review process. Network meta-analyses are primarily meta-analyses, and should therefore be performed according to the explicit and rigorous methods used in systematic reviews and meta-analyses to minimise bias.¹⁹ We performed a methodological systematic review of published reports of network meta-analyses to examine how they were reported and conducted. In particular, we assessed whether the reports adequately featured the key methodological components of the systematic review process.

Methods

Search strategy

We systematically searched for all published network meta-analyses in the following databases from their dates of inception: Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, Medline, and Embase. Search equations were developed for each database and were based partly on search terms used by Song and colleagues¹² and Salanti and colleagues²⁰ (web table 1). The equations were based on specific free text words pertaining to network meta-analyses or overviews of reviews. The date of the last search was 12 July 2012. We also screened the references of methodology papers and reviews^{11-17 21} and searched for papers citing landmark statistical articles through the Web of Science or SCOPUS database.^{6 8 22-24}

Eligibility criteria

All reports of network meta-analyses comparing the clinical efficacy of three or more interventions based on randomised controlled trials were eligible. We excluded reports of adjusted meta-analyses involving indirect comparisons (that is, an open loop network of three interventions)^{10 13 17} as well as publications of methodology, editorial style reviews or reports, cost effectiveness reviews, reviews based on individual patient data, and reviews not involving human participants. We also excluded reviews not published in English, French, German, Spanish, and Italian.

Selection of relevant reports of network meta-analyses

Two reviewers (AB, LT) independently selected potentially relevant reports of network meta-analyses on the basis of the title and abstract and, if needed, the full text according to prespecified eligibility criteria. Disagreements were discussed to reach consensus.

Data extraction

A standardised form was used to collect all data from the original reports of network meta-analyses and supplementary appendices when available. We collected all data for epidemiological and descriptive characteristics and those pertaining to the key methodological components of the systematic review process. Two reviewers (AB, RS) independently extracted all data from a random sample of 20% of reports. From all reports, two reviewers (AB, RS) independently extracted data for items that involved subjective interpretation. The corresponding items were questions referring to participants, interventions, comparisons, and outcomes; the type of comparator; how reporting bias was assessed; the

publication status of selected randomised controlled trials; and whether statistical assumptions (homogeneity, similarity, and consistency) were mentioned. Disagreements were resolved by discussion.

General characteristics

The following general characteristics were collected: journal name, year of publication, country of corresponding author, medical area, funding source, and type of interventions included in the network. We categorised journal types into general or specialty; we also identified journals with high impact factors (that is, the 10 journals with the highest impact factor for each medical subject category of the *Journal Citation Reports* 2011). Regarding network typology, we collected the number of interventions, the number of comparisons conducted by at least one randomised controlled trial (that is, direct comparisons), and the number of randomised controlled trials included in the network.

Reporting of key methodological components of the systematic review process

According to PRISMA²⁵ and AMSTAR²⁶ guidelines, we assessed whether key methodological components were reported or not. In the introduction and methods sections of each report, these factors included:

- Questions referring to participants, interventions, comparisons, and outcomes
- Existence of a review protocol
- Primary outcome explicitly specified in the article or abstract, in the primary study objectives, or as the only outcome reported in the article²⁷
- Information sources including databases searched, electronic search strategy, date of last search and period covered by the search of each database, any other sources (conference abstracts, unpublished studies, textbooks, specialty registers (for example, the US Food and Drug Administration), contact with study authors, reviewing the references in the studies found or any relevant systematic reviews), ongoing studies searched, and any or no restrictions related to language or publication status (that is, published studies or any other sources or ongoing studies)
- Methods for study selection and data extraction
- Methods used for assessing risk of bias of individual studies, such as use of any scales, checklists, or domain based evaluation recommended by the Cochrane Collaboration, with separate critical assessments for different domains (that is, allocation concealment, generation of sequence allocation, blinding, incomplete outcome data, or selective reporting)
- Methods to incorporate assessment of risk of bias of individual studies in the analysis or conclusion of review (that is, subgroup analysis, inclusion criteria, sensitivity analysis, or the grading recommendations²⁸ assessment development and evaluation (GRADE))
- Methods to assess publication or reporting bias (statistical or graphical evaluation).

Regarding the search strategy, we distinguished between a “de novo” search (a new comprehensive literature search) and a mixed search strategy (an initial search for systematic reviews complemented by an updated search of trials not covered by the search period of these systematic reviews).

In the results section, we assessed whether reports included the following components:

- Study selection, including the number of studies screened, assessed for eligibility, and included in the review; with reasons for exclusions at each stage (for example, a PRISMA flowchart) and the list of studies included and excluded
- Study characteristics, including a description of the trial network with the number of interventions and number of comparisons conducted by at least one randomised controlled trial (that is, direct comparisons), characteristics of patients (for example, age, sex, disease status, severity), duration of follow-up of patients and duration of interventions, description of interventions, number of study groups and patients, and funding sources
- Risk of bias within studies
- Risk of bias across studies (publication bias).

Finally, we assessed whether the discussion section included:

- A limitation regarding reporting bias
- Whether the authors mentioned or discussed the assumptions required in network meta-analyses (based on homogeneity, similarity, consistency, and exchangeability)
- Whether authors mentioned a conflict of interest.

Conduct quality when the reporting was adequate

When items were adequately reported, we subsequently assessed whether the conduct quality was adequate. We considered the following items as inadequate conduct: electronic search of only one bibliographic database, restricting study selection to published reports, lack of independent duplicate study selection, and lack of independent duplicate data extraction.

Composite outcome assessing inadequate reporting or inadequate conduct quality

To summarise data, we built two composite outcomes assessing inadequate reporting or inadequate conduct quality (when the reporting was adequate). The first composite outcome was based on guidelines from Li and colleagues.²⁹ Inadequate reporting or conduct quality was considered as one of the following:

- The authors did not report a literature search, they reported an electronic search of only one bibliographic database, or they did not supplement their investigation by searching for any other sources
- The authors did not report an assessment of risk of bias of individual studies.

The second composite outcome was defined as an all or none approach, based on seven methodological items identified as mandatory according to the methodological expectations of Cochrane intervention reviews.³⁰ Conduct or reporting quality was considered inadequate if:

- The authors did not report questions referring to participants, interventions, comparisons, and outcomes
- The authors did not report eligibility criteria
- The authors did not report a literature search; they reported an electronic search of only one bibliographic database; or they did not supplement their investigation by searching for references in studies found, any relevant systematic reviews, or ongoing studies (for example, clinicaltrials.gov or the World Health Organization International Clinical Trials Registry Platform)

- The authors did not report the methods of study selection and the data extraction process, or they reported inadequate methods
- The authors did not report the results of the selection of primary studies
- The authors did not provide a description of the primary studies
- The authors did not report an assessment of the risk of bias of individual studies.

Statistical analysis

Quantitative data were summarised by medians and interquartile ranges, and categorical data summarised by numbers and percentages. We compared general characteristics and reporting of key items for network meta-analyses with results published in general and specialty journals and with public and private funding sources. All or none measurements were analysed by χ^2 tests for categorical data. All tests were two sided, and $P < 0.05$ was considered significant. Statistical analyses involved use of SAS version 9.3 (SAS Institute).

Results

General characteristics

Our search identified 1226 citations, including 333 duplicate citations. Among the 893 potentially relevant publications, 118 were eligible reports of network meta-analyses. Three additional network meta-analyses were identified from methodological articles, resulting in 121 eligible reports of network meta-analyses (figs 1 and 2; web table 2 shows the main characteristics of the network meta-analyses^{w1-w12}).

Reports of network meta-analyses were published in 75 different journals, with 55 reports (45%) published in general journals, 66 (55%) in specialty journals, and 56 (46%) in journals with a high impact factor (table 1; web table 3). The network meta-analyses covered a wide range of medical areas, and 100 (83%) described the assessment of pharmacological interventions. The median number of interventions assessed per network meta-analysis was seven (interquartile range 5-9), and the median number of randomised controlled trials included per network meta-analysis was 22 (15-40).

Reporting of key methodological components of the systematic review process

Only three (2%) network meta-analyses did not report the databases searched in the methods section; of the remaining 118, 99 (82%) used a de novo search strategy and 19 (16%) a mixed search strategy. However, 42 (35%) did not report searching other sources, and 39 (32%) did not report any information regarding restriction or no restriction of study selection related to publication status (table 2). In addition, 42 (35%) and 32 (26%) network meta-analyses did not report the methods used for study selection and data extraction, respectively. A total of 61 (50%) network meta-analyses did not report any information regarding the assessment of risk of bias of individual studies. Of 121 network meta-analyses, 103 (85%) did not report any methods to assess the likelihood of publication bias.

In the results section, 95 (79%) network meta-analyses did not describe the characteristics of primary studies (that is, characteristics of the network, patient characteristics, and interventions); 70 (58%) did not report the risk of bias assessment within studies.

The similarity and consistency assumptions were not frequently mentioned in network meta-analyses (66% (n=80) and 44% (n=63) of reports, respectively; table 2). These findings did not differ by journal type or funding source (web table 4).

Conduct quality when the reporting was adequate

For network meta-analyses with adequate reporting, the conduct quality was frequently inadequate: 11% of these network meta-analyses involved a search of only one bibliographic database, 20% restricted the study selection to published articles, 43% lacked an independent duplicate study selection, and 21% lacked an independent extraction of duplicate data. These findings did not differ by journal type or funding source (table 3; web table 5).

Composite outcome assessing inadequate reporting or inadequate conduct quality

Overall, 87 network meta-analyses (72%) showed inadequate reporting of key methodological components or inadequate conduct quality according to guidelines from Li and colleagues.²⁹ This measurement did not differ by journal type (general journal, 69% (95% confidence interval 57% to 81%); specialty journal, 74% (63% to 85%); P=0.5) or funding source (public funding, 67% (74% to 79%); private funding, 79% (67% to 90%); P=0.2).

We used a second composite according to the seven mandatory items of methodological expectations of Cochrane intervention reviews.³⁰ Based on this composite, 120 network meta-analyses (99%) showed inadequate reporting of key methodological components or inadequate conduct. These findings did not differ by journal type or funding source.

Discussion

We identified 121 reports of network meta-analyses covering a wide range of medical areas. Key methodological components of the systematic review process were missing in most reports. The reporting did not differ by publication in a general or specialty journal or by public or private funding.

Several guidelines have been developed to assess the methodological quality and reporting of systematic reviews and meta-analyses,^{26-30,33} and according to all these guidelines, retrieving all relevant studies is crucial for the success of every systematic review. The literature search needs to be exhaustive and adequate to maximise the likelihood of capturing all relevant studies and minimise the effects of reporting biases that might influence the nature and direction of results.³⁴⁻³⁶ More than half of our network meta-analyses failed to report exhaustive searches; for those with adequate reporting, 11% searched only one electronic database (web table 6).

In addition, more than half of our network meta-analyses did not describe the characteristics of patients, the duration of follow-up or treatment, or the intervention. The lack of adequate reporting of the characteristics of primary studies raises several problems. Knowing precisely the type of intervention, characteristics, and duration of follow-up and treatment is a preliminary step for researchers and readers to assess the assumption of similarity required in network meta-analyses.^{12,37}

An important step of the systematic review process is to assess the risk of bias of individual studies.³⁸⁻⁴² The risk of bias in individual studies could affect the findings of network meta-analyses.⁴³ However, 58% of our network meta-analyses did not report the risk of bias of individual studies in the results sections, and 84% did not report a method to incorporate

assessment of risk of bias of individual studies in the analysis or conclusions of the review (web table 6).

Publication bias could affect the results of meta-analyses and network meta-analyses.⁴⁴⁻⁴⁷ This factor is all the more important, considering that we lacked the validated methods to detect or to appropriately adjust or exclude reporting bias in network meta-analyses. Moreover, reporting bias across trials could differentially affect the various comparisons in the network and modify the rank order efficacy of treatments. However, most of our network meta-analyses (85%) did not report a method to assess the publication bias. Although reporting bias could have a substantial effect on the conclusions of a network meta-analysis, it is probably neglected.

Strengths and limitations of study

To the best of our knowledge, no study investigated the systematic review process in network meta-analyses (mixed treatment, comparison meta-analyses).¹⁰⁻¹³ Recently, the Agency for Healthcare Research and Quality assessed the statistical methods used in a sample of published network meta-analyses, but did not analyse the conduct and reporting of the systematic review process.⁴⁸

Our study had some limitations. Assessing conduct quality from published reports alone could be unreliable, as has been shown for randomised trials.⁴⁹ The study authors may have used adequate methods but omitted important details from their reports, or key information may have been deleted during the publication process. However, we were in the position of the reader, who can only assess what was reported. In addition, we judged that the conduct quality was inadequate only if the reporting was adequate. Furthermore, we did not assess health technology assessment reports, which could have had higher quality or precision than the reports analysed.

Conclusions and implications

Our study identified some important methodological components in the reports of network meta-analyses that may raise doubts about the confidence we could have in their conclusions. Potential flaws in the conduct of network meta-analyses could affect their findings. Performing a network meta-analysis outside the realm of a systematic review, without extensive and thorough searches for eligible trials, and without risk of bias assessments of trials, would increase the risk of a network meta-analysis producing biased results. Guidelines are needed to improve the quality of reporting and conduct of network meta-analyses.

Contributors: AB was responsible for the study conception, search of trials, selection of trials, data extraction, data analysis, interpretation of results, and drafting the manuscript. LT was responsible for the study conception, search of trials, selection of trials, interpretation of results, and drafting the manuscript. RS was responsible for the data extraction, interpretation of results, and drafting the manuscript. PR was the guarantor and was responsible for the study conception, interpretation of results, and drafting the manuscript. All authors, external and internal, had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis.

Funding: This study was funded by an academic grant for the doctoral student from "Pierre et Marie Curie University." Our team is supported by an academic grant (DEQ20101221475) for the programme "Equipe espoir de la Recherche," from the Fondation pour la Recherche Médicale. The funding agencies have no role in the design or conduct of the study; collection, management, analysis, or interpretation of the data; or preparation and review of the manuscript.

What is already known on this topic

Network meta-analyses are primarily meta-analyses, and should be conducted by respecting the methodological rules of systematic reviews

Network meta-analyses are subject to the same methodological risks of standard pairwise systematic reviews, and because this method in network meta-analyses is complex, it is probably more vulnerable to these risks

What this study adds

Key methodological components of the systematic review process are frequently inadequately reported in publications of network meta-analyses

The level of inadequate reporting does not differ between reports published by a general or specialty journal or between reports with public or private funding

Inadequate reporting of results of network meta-analyses raises doubts about the ability of network meta-analyses to help clinical researchers determine the best available treatment

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Ethical approval not required.

Data sharing: No additional data available.

- 1 Dickersin K. Health-care policy. To reform US health care, start with systematic reviews. *Science* 2010;329:516-7.
- 2 Lathyris DN, Patsopoulos NA, Salanti G, Ioannidis JP. Industry sponsorship and selection of comparators in randomized clinical trials. *Eur J Clin Invest* 2010;40:172-82.
- 3 Hochman M, McCormick D. Characteristics of published comparative effectiveness studies of medications. *JAMA* 2010;303:951-8.
- 4 Volpp KG, Das A. Comparative effectiveness—thinking beyond medication A versus medication B. *N Engl J Med* 2009;361:331-3.
- 5 Estellat C, Ravaud P. Lack of head-to-head trials and fair control arms: randomized controlled trials of biologic treatment for rheumatoid arthritis. *Arch Intern Med* 2012;172:237-44.
- 6 Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997;50:683-91.
- 7 Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ* 2003;326:472.
- 8 Lumley T. Network meta-analysis for indirect treatment comparisons. *Stat Med* 2002;21:2313-24.
- 9 Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004;23:3105-24.
- 10 Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR task force on indirect treatment comparisons good research practices: part 1. *Value Health* 2011;14:417-28.
- 11 Donegan S, Williamson P, Gamble C, Tudur-Smith C. Indirect comparisons: a review of reporting and methodological quality. *PLoS One* 2010;5:e11054.
- 12 Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009;338:b1147.
- 13 Wells GA, Sultan SA, Chen L, Khan M, Coyle D. Indirect evidence: indirect treatment comparisons in meta-analysis. Canadian Agency for Drugs and Technologies in Health, 2009.
- 14 Glenny AM, Altman DG, Song F, Sakarovich C, Deeks JJ, D'Amico R, et al. Indirect comparisons of competing interventions. *Health Technol Assess* 2005;9:1-134, iii-iv.
- 15 Schöttker B, Luhmann D, Boulkheir D, Raspe H. Indirekte vergleichende von therapieverfahren. GMS Health Technology Assessment, 2009.
- 16 Cucherat M, Izard V. Les comparaisons indirectes Méthodes et validité. HAS-Service de l'Evaluation des Medicaments, 2009:66.
- 17 Hoaglin DC, Hawkins N, Jansen JP, Scott DA, Itzler R, Cappelleri JC, et al. Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ISPOR task force on indirect treatment comparisons good research practices: part 2. *Value Health* 2011;14:429-37.
- 18 Song F, Harvey I, Lilford R. Adjusted indirect comparison may be less biased than direct comparison for evaluating new pharmaceutical interventions. *J Clin Epidemiol* 2008;61:455-63.
- 19 Li T, Puhon MA, Vedula SS, Singh S, Dickersin K. Network meta-analysis-highly attractive but more methodological research is needed. *BMC Med* 2011;9:79.
- 20 Salanti G, Kavvoura FK, Ioannidis JP. Exploring the geometry of treatment networks. *Ann Intern Med* 2008;148:544-53.
- 21 Edwards SJ, Clarke MJ, Wordsworth S, Borrill J. Indirect comparisons of treatments based on systematic reviews of randomised controlled trials. *Int J Clin Pract* 2009;63:841-54.
- 22 Lu G, Ades AE, Sutton AJ, Cooper NJ, Briggs AH, Caldwell DM. Meta-analysis of mixed treatment comparisons at multiple follow-up times. *Stat Med* 2007;26:3681-99.
- 23 Salanti G, Higgins JP, Ades AE, Ioannidis JP. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008;17:279-301.
- 24 Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol* 2011;64:163-71.
- 25 Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Ann Intern Med* 2009;151:W65-94.
- 26 Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007;7:10.
- 27 Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004;291:2457-65.
- 28 Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924-6.
- 29 Li T, Vedula SS, Scherer R, Dickersin K. What comparative effectiveness research is needed? A framework for using guidelines and systematic reviews to identify evidence gaps and research priorities. *Ann Intern Med* 2012;156:367-77.
- 30 Higgins JPT, Green S, eds. Cochrane handbook for systematic reviews of interventions version 5.1.0. Cochrane Collaboration, 2011.
- 31 Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009;6:e1000100.
- 32 Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of reporting of meta-analyses. *Lancet* 1999;354:1896-900.
- 33 Mills EJ, Ioannidis JP, Thorlund K, Schunemann HJ, Puhon MA, Guyatt GH. How to use an article reporting a multiple treatment comparison meta-analysis. *JAMA* 2012;308:1246-53.
- 34 McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? *Lancet* 2000;356:1228-31.
- 35 Royle P, Milne R. Literature searching for randomized controlled trials used in Cochrane reviews: rapid versus exhaustive searches. *Int J Technol Assess Health Care* 2003;19:591-603.
- 36 Sampson M, Barrowman NJ, Moher D, Klassen TP, Pham B, Platt R, et al. Should meta-analysts search Embase in addition to Medline? *J Clin Epidemiol* 2003;56:943-55.
- 37 Xiong T, Parekh-Bhurke S, Loke YK, Abdelhamid A, Sutton AJ, Eastwood AJ, et al. Overall similarity and consistency assessment scores are not sufficiently accurate for predicting discrepancy between direct and indirect comparison estimates. *J Clin Epidemiol* 2013;66:184-91.
- 38 Gluud LL. Bias in clinical intervention research. *Am J Epidemiol* 2006;163:493-501.
- 39 Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. *BMJ* 2001;323:42-6.
- 40 Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008;336:601-5.
- 41 Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998;352:609-13.
- 42 Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12.
- 43 Dias S, Welton NJ, Marinho V, Salanti G, Higgins JP, Ades AE. Estimation and adjustment of bias in randomized evidence by using mixed treatment comparison meta-analysis. *J R Stat Soc* 2010;173:613-29.
- 44 Thornton A, Lee P. Publication bias in meta-analysis: its causes and consequences. *J Clin Epidemiol* 2000;53:207-16.
- 45 Trinquart L, Abbe A, Ravaud P. Impact of reporting bias in network meta-analysis of antidepressant placebo-controlled trials. *PLoS One* 2012;7:e35219.
- 46 Trinquart L, Chatellier G, Ravaud P. Adjustment for reporting bias in network meta-analysis of antidepressant trials. *BMC Med Res Methodol* 2012;12:150.
- 47 Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Res Synth Methods* 2012;3:80-97.
- 48 Coleman CI, Phung OJ, Cappelleri JC, Baker WL, Kluger J, White CM, et al. Use of mixed treatment comparisons in systematic reviews. *AHRQ Methods for Effective Health Care* 2012. www.ncbi.nlm.nih.gov/books/NBK107330/.
- 49 Vale CL, Tierney JF, Burdett S. Can trial quality be reliably assessed from published reports of cancer trials: evaluation of risk of bias assessments in systematic reviews. *BMJ* 2013;346:f1798.

Accepted: 31 May 2013

Cite this as: *BMJ* 2013;347:f3675

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works

on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>.

Tables

Table 1 | Epidemiological and descriptive characteristics of 121 reports of network meta-analyses

Item and subcategory	No (%) of reports
Journal type	
General journal	55 (45)
With high impact factor	29 (53)
Specialty journal	66 (55)
With high impact factor	27 (41)
Location of corresponding author	
Europe	69 (57)
North America	38 (31)
Asia	8 (7)
South America	3 (2)
Oceania	3 (2)
Funding source	
Private	41 (34)
Public	40 (33)
None	15 (12)
Both private and public	7 (6)
Unclear	18 (15)
Medical area	
Cardiology	27 (22)
Rheumatology	16 (13)
Endocrinology	12 (10)
Oncology	10 (8)
Infectious disease	9 (7)
Psychiatry/psychology	8 (7)
Neurology	8 (7)
Respiratology	6 (5)
Ophthalmology	5 (4)
Surgery	5 (4)
Other (≤ 3 reviews per medical area; 11 medical areas)	15 (12)
Type of intervention assessed	
Pharmacological intervention	100 (83)
Different class *	72 (60)
Same class *	35 (29)
Any dose of same drug *	12 (10)
Non-pharmacological intervention	11 (9)
Devices *	10 (8)
Surgery or procedure *	7 (6)
Therapeutic strategy or education	5 (4)
Both (pharmacological and non-pharmacological intervention)	10 (8)
Type of network	
No of interventions assessed per network†	7 (5-9)
No of comparisons conducted by at least one randomised controlled trial (per network meta-analysis)†	10 (6-15)
No of network meta-analyses with at least one closed loop	103 (85)
No of randomised controlled trials included in network meta-analysis (per network meta-analysis)†	22 (15-40)

Data are no (%) of reports unless stated otherwise.

Table 1 (continued)

Item and subcategory	No (%) of reports
*Multiple answers were possible, so the total does not equal 100%.	
†Data are median (interquartile range).	

Table 2 | Reporting of key methodological components of the systematic review process in network meta-analyses, by journal type

Items	Overall (n=121)	General journals (n=55)	Specialty journals (n=66)
Introduction			
Questions referring to participants, interventions, comparisons, outcomes, and study design	110 (91)	51 (93)	59 (89)
Methods			
Existence of systematic review protocol	15 (12)	9 (16)	6 (9)
Primary outcome(s)	85 (70)	44 (80)	41 (62)
Information sources searched			
Databases searched	118 (98)	53 (96)	65 (98)
Electronic search strategy for each database	33 (27)	19 (35)	14 (21)
Date of last search for each database	109 (90)	50 (91)	59 (89)
Period covered by search for each database	83 (69)	36 (65)	47 (71)
Search for any other sources (conference abstracts, unpublished studies, textbooks, specialty registers, contact with study authors, reviewing the references in the studies found or any relevant systematic reviews)	79 (65)	38 (69)	41 (62)
Reviewing the references in the studies found or any relevant systematic reviews	54 (45)	27 (49)	27 (41)
Search for ongoing studies	19 (16)	13 (24)	6 (9)
Restriction or no restriction related to language	92 (76)	42 (76)	50 (76)
Restriction or no restriction related to the publication status	82 (68)	38 (69)	44 (67)
Study selection and data collection process			
Process for selecting studies	79 (65)	35 (64)	44 (67)
Method of data extraction	89 (74)	42 (76)	47 (71)
Methods used for assessing risk of bias of individual studies	60 (50)	27 (49)	33 (50)
Methods to incorporate assessment of risk of bias of individual studies in the analysis or conclusions of review			
Subgroup analysis	9 (7)	4 (7)	5 (8)
Inclusion criteria	4 (3)	1 (2)	3 (5)
GRADE	4 (3)	4 (7)	0 (0)
Adjustment	2 (2)	2 (4)	0 (0)
Any of these methods	19 (16)	11 (20)	8 (13)
Assessment of risk of bias that may affect the cumulative evidence (publication bias)	18 (15)	11 (20)	7 (11)
Results			
Study selection			
No of studies screened, assessed for eligibility, and included in the review	90 (74)	43 (78)	47 (71)
List of studies included	117 (97)	53 (96)	64 (97)
List of studies excluded	9 (7)	6 (11)	3 (5)
Study characteristics			
Description of network	82 (68)	39 (71)	43 (65)
Characteristics of patients (for example, age, female:male ratio)	70 (58)	30 (55)	40 (61)
Duration of follow-up of patients	60 (50)	26 (47)	34 (52)
Duration of interventions	34 (28)	15 (27)	19 (29)
Description of interventions	64 (53)	30 (55)	34 (52)
No of study groups	95 (79)	42 (76)	53 (80)
No of patients	97 (80)	41 (75)	56 (85)
Funding source	13 (11)	8 (15)	5 (8)
Risk of bias within studies	51 (42)	23 (42)	28 (42)
Risk of bias across studies (publication bias)	18 (15)	11 (20)	7 (11)
Discussion			
Reporting of limitations at review level (reporting or publication bias)	58 (48)	32 (58)	26 (39)

Table 2 (continued)

Items	Overall (n=121)	General journals (n=55)	Specialty journals (n=66)
Other items			
Assumptions required in network meta-analysis			
Homogeneity assumption	107 (88)	50 (91)	57 (86)
Similarity assumption	41 (34)	18 (33)	23 (35)
Consistency or exchangeability assumption	58 (48)	32 (58)	26 (39)
Conflict of interest	95 (79)	47 (85)	48 (73)

Data are number (%) of reports featuring the corresponding item.

Table 3| Inadequate quality of conduct of the systematic review process with adequate reporting in network meta-analyses, by journal type

Item (reported and of inadequate quality of conduct)	Overall	General journal	Specialty journal
Electronic search of only one bibliographic database	13/118 (11)	8/53 (15)	5/65 (8)
Restriction of study selection based on publication status	16/82 (20)	6/38 (16)	10/44 (23)
Lack of independent duplicate study selection	34/79 (43)	13/37 (35)	21/42 (50)
Lack of independent duplicate data extraction	19/89 (21)	10/41 (24)	9/48 (19)

Data are number (%) / total number of reports. Denominators of fractions indicate the total number of reports in which the corresponding item was reported.

Figures

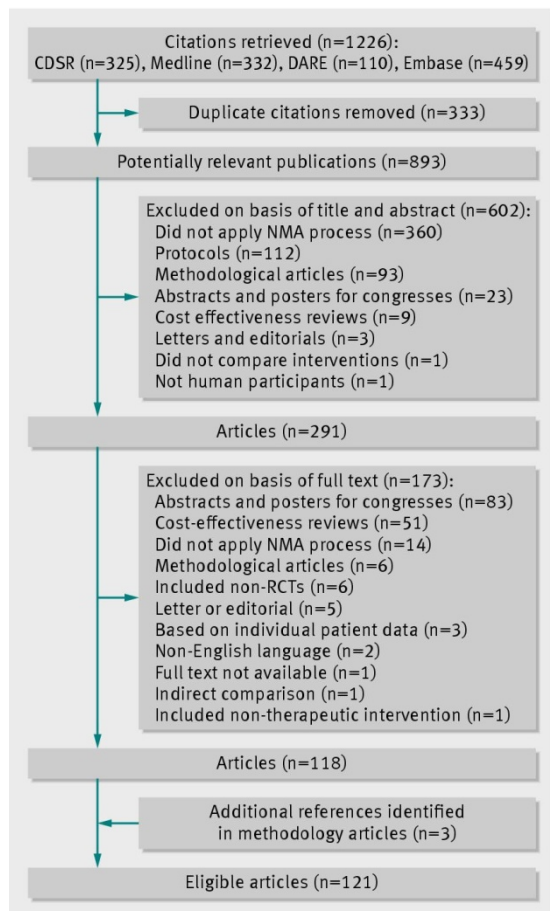


Fig 1 Flowchart of selection of network meta-analyses. CDSR=Cochrane Database of Systematic Reviews; DARE= Database of Abstracts of Reviews of Effects; NMA=network meta-analysis; RCT= randomised controlled trial

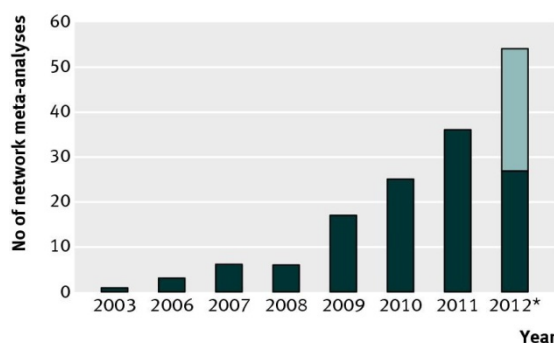


Fig 2 Number of reports of network meta-analyses published per year. *We estimated the number of reports published from July to December 2012 on the basis of the number of published network meta-analyses from the six previous months. Between January and June 2012, 27 reports were published (dark region); therefore, we estimated 27 reports from July to December 2012 (light region)

3. Chapitre : Présentation des résultats dans les méta-analyses en réseau

3.1. Justification de l'étude

Les méta-analyses en réseau sont de plus en plus utilisées dans l'évaluation thérapeutique¹⁶⁶. Cependant les résultats des méta-analyses en réseau peuvent être difficiles à interpréter pour les cliniciens et les décideurs de santé¹⁰⁴. Les publications de méta-analyse en réseau devraient permettre au lecteur d'évaluer l'ensemble des preuves disponibles dans le réseau et l'efficacité des interventions entre elles. Une présentation inadéquate des résultats peut affecter l'interprétation des méta-analyses en réseau et induire en erreur le lecteur¹⁷³.

3.2. Objectif

L'objectif est d'évaluer comment les résultats des méta-analyses en réseau étaient présentés.

3.3. Méthodes

Nous avons utilisé notre échantillon des 121 méta-analyses en réseau identifiées dans notre précédente étude. Nous avons extrait les données suivantes :

- 1) les caractéristiques générales : le type de critère de jugement, le nombre de critères de jugement évalués, le nombre de tables, de figures et de pages imprimées, l'existence d'une annexe.
- 2) Les aspects statistiques : nous avons évalué si les auteurs avaient utilisé une approche Bayésienne, puis pour ces publications utilisant une approche Bayésienne, nous avons évalué si les auteurs rapportaient une loi *a priori* non informative ou informative pour les paramètres basiques et la variance inter-essai. Basés sur les recommandations de Donegan et ses collègues¹⁷⁴, nous avons évalué la similarité. Nous avons considéré deux méthodes adéquates : évaluer les caractéristiques des patients et des essais par comparaisons ; ou examiner l'impact de covariables pouvant modifier l'effet traitement (analyse en sous-groupe

ou méta-régression).

3) La description du réseau et de sa géométrie : nous avons évalué si les auteurs rapportaient les interventions, les comparaisons directes, le nombre d'essais et de patients pour chaque branche du réseau, une représentation graphique du réseau) ;

4) Les estimations des effets traitements deux à deux entre les interventions: nous avons évalué si les résultats de tous les ECRs inclus dans le réseau (les données résumées pour chaque bras ou l'effet traitement), les estimations de l'effet traitement avec une mesure d'incertitude provenant des comparaisons directes, indirectes et mixtes. Nous avons examiné les différents formats possibles pour présenter les résultats: figure, table, matrice, graphique en forêt, graphique du réseau (chaque branche du réseau avec la taille de l'effet correspondant), ou dans le texte. Nous avons également évalué si les auteurs rapportaient un classement.

Enfin pour résumer les données, nous avons construit un critère composite basé sur les recommandations de Ioannidis¹⁰⁶, de Mills et ses collègues¹⁷⁵, de Cipriani et ses collègues¹⁰³, et d'Ades et ses collègues¹⁶⁸. Ce critère composite était considéré comme inadéquate en cas d'absence d'au moins un des items suivants : les interventions, les comparaisons directes entre les interventions, le nombre d'ECRs et de patients pour chaque branche du réseau; les estimations des effets traitements provenant des sources directes, indirectes, ou mixtes.

3.4. Résultats

Le nombre médian de critère de jugement évalué par méta-analyse en réseau était de 2 (intervalle interquartile 1-3). Dans l'ensemble, 70 (58%) publication de méta-analyses en réseau avaient un supplément ou une annexe. L'approche Bayésienne était appliquée dans 91 (75%) méta-analyses en réseau: parmi elles 51 décrivaient une loi a priori non informative pour les paramètres basiques, et 30 pour la variance inter-essai. Les auteurs présentaient une

méthode pour évaluer la similarité dans 43 (36%) des méta-analyses en réseau.

Les auteurs ne rapportaient pas tous les éléments (les interventions, les comparaisons directes, le nombre d'essais et de patients pour chaque branche du réseau) pour décrire le réseau et sa géométrie dans 100 (83%) des méta-analyses en réseau. Les estimations de l'effet traitement provenant de preuve directe, indirecte, et mixte n'ont pas été rapportées dans 48 (40%), 108 (89%), et 43 (36%) des articles, respectivement. Un classement des interventions était présenté dans 52 méta-analyses en réseau, parmi elles 43 ne rapportaient pas une mesure d'incertitude sur ce classement.

Le critère composite montre que 119 (98%) des méta-analyses en réseau ne rapportaient pas au moins un des éléments suivant : la description du réseau (les interventions, les comparaisons directes, le nombre d'essais ou de patients pour chaque comparaison directe) ou l'effet traitement provenant soit des sources directes, indirectes, ou mixtes. Ce résultat ne diffère pas selon le type de journal ou le type de financement.

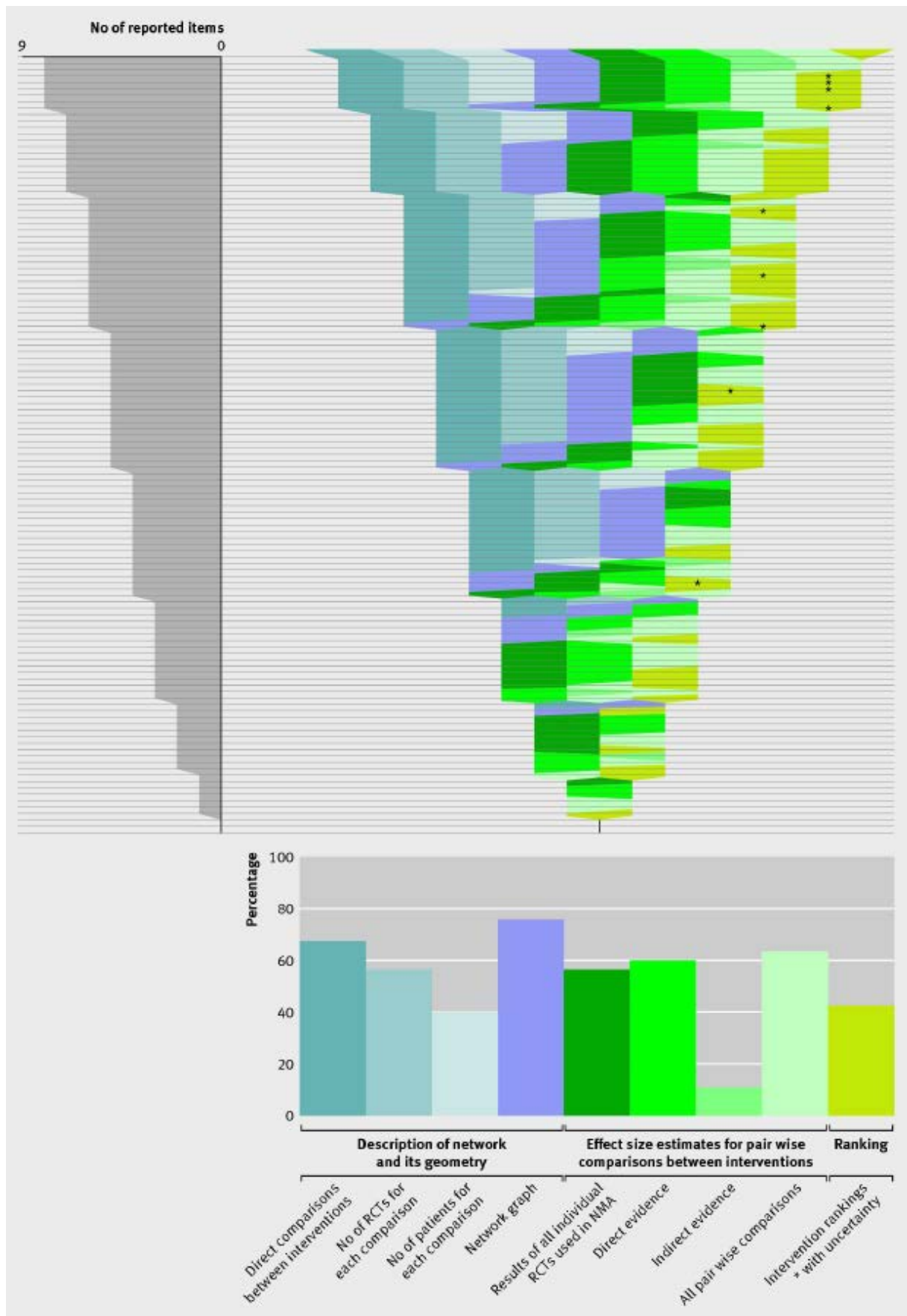


Figure 15 : Présentation des éléments relatifs aux résultats dans les méta-analyses en réseau.

L'idée était de présenter les éléments rapportés dans chacune des 121 méta-analyses en réseau. Chaque ligne horizontale correspond à une méta-analyse en réseau. Une couleur est attribuée à chacun des 10 items étudiés. Les bandes de couleur indiquent quels items sont rapportés pour chaque méta-analyse en réseau. Les items sont regroupés en trois catégories: 4 items en bleu relatifs à la description du réseau, un item en violet relatif à sa géométrie (graphique du réseau); 4 items en vert relatif à l'estimation de l'effet traitement pour toutes les comparaisons deux à deux entre les interventions; et un item relatif au classement des interventions. Les 121 méta-analyses en réseau sont classées selon le nombre total d'items rapportés, dans l'ordre décroissant. La figure à gauche montre la distribution du nombre total d'items rapportés dans les 121 méta-analyses en réseau. La figure en bas montre la proportion de publications rapportant chaque item.

3.5. Discussion

Nos résultats confirment la nécessité de directives claires pour rapporter les méta-analyses en réseau. Nous avons mis en évidence que la présentation des résultats est très hétérogène dans les publications de méta-analyses en réseau.

De nombreux éléments importants qui peuvent affecter l'interprétation et la compréhension des résultats des méta-analyses en réseau sont manquants ou mal rapportés. Les résultats des méta-analyses en réseau peuvent être difficiles pour les cliniciens à interpréter car le nombre d'interventions évaluées peut être important et les approches méthodologiques peuvent être complexes¹⁰⁴. Ainsi les lecteurs devraient pouvoir évaluer l'ensemble des preuves disponibles à l'aide d'éléments rapportés comme les interventions incluses, les comparaisons directes incluses, le nombre d'essais et de patients inclus dans chaque comparaison qui permettent de décrire le réseau. Cependant, seulement 17% des méta-analyses en réseau décrivent adéquatement le réseau et sa géométrie. Les publications devraient rapporter toutes les

estimations provenant des comparaisons directes et indirectes pour pouvoir juger si ces estimations sont suffisamment similaires pour les combiner¹³⁹. Seulement 10% des méta-analyses en réseau rapportaient à la fois les preuves directes et indirectes.

Dans notre étude, nous avons choisi d'utiliser un critère composite basé sur les recommandations disponibles pour évaluer la qualité de présentation des résultats. Cependant nous reconnaissons qu'il n'y a pas de consensus sur les éléments qui doivent être requis dans les publications de méta-analyse en réseau. Actuellement, cet axe est hautement débattu. Une autre limite est le nombre de mot réduit dans les journaux, ainsi des détails importants auraient pu être omis dans les publications ou des informations clés peuvent avoir été supprimées au cours du processus de publication. Néanmoins, la plupart des journaux ont un appendice en ligne et nous avons évalué toutes les données à la fois dans les articles et l'appendice lorsqu'il était disponible.

En conclusion, notre étude peut fournir de base pour le développement d'un outil d'évaluation de la qualité de présentation des méta-analyses en réseau et aider à l'interprétation.

Article

Aïda Bafeta, Ludovic Trinquart, Raphaèle Seror, Philippe Ravaud.
Reporting of results from network meta-analyses: methodological systematic
review. *British Medical Journal* 2014;348:g1741

RESEARCH

Reporting of results from network meta-analyses: methodological systematic review

 OPEN ACCESS

Aïda Bafeta *PhD student*¹, Ludovic Trinquart *postdoctoral research fellow*^{1 2 3 4}, Raphaèle Seror *associate professor of rheumatology*^{1 3}, Philippe Ravaud *professor of epidemiology and director*^{1 2 3 4}

¹Assistance Publique-Hôpitaux de Paris, Centre d'Epidémiologie Clinique, INSERM U1153, Hôpital Hôtel-Dieu, 75004 Paris, France; ²Université Paris Descartes-Sorbonne Paris Cité, Paris; ³French Cochrane Centre, Paris; ⁴Department of Epidemiology, Mailman School of Public Health, Columbia University, New York City, USA

Abstract

Objective To examine how the results of network meta-analyses are reported.

Design Methodological systematic review of published reports of network meta-analyses.

Data sources Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, Medline, and Embase, searched from inception to 12 July 2012.

Study selection All network meta-analyses comparing the clinical efficacy of three or more interventions in randomised controlled trials were included, excluding meta-analyses with an open loop network of three interventions.

Data extraction and synthesis The reporting of the network and results was assessed. A composite outcome included the description of the network (number of interventions, direct comparisons, and randomised controlled trials and patients for each comparison) and the reporting of effect sizes derived from direct evidence, indirect evidence, and the network meta-analysis.

Results 121 network meta-analyses (55 published in general journals; 48 funded by at least one private source) were included. The network and its geometry (network graph) were not reported in 100 (83%) articles. The effect sizes derived from direct evidence, indirect evidence, and the network meta-analysis were not reported in 48 (40%), 108 (89%), and 43 (36%) articles, respectively. In 52 reports that ranked interventions, 43 did not report the uncertainty in ranking. Overall, 119 (98%) reports of network meta-analyses did not give a description of the network or effect sizes from direct evidence, indirect evidence, and the network meta-analysis. This finding did not differ by journal type or funding source.

Conclusions The results of network meta-analyses are heterogeneously reported. Development of reporting guidelines to assist authors in writing and readers in critically appraising reports of network meta-analyses is timely.

Introduction

When several healthcare interventions are available for the same condition, the corresponding network of results from randomised controlled trials must be considered (that is, which of the considered interventions have been compared against each other or against a common comparator such as placebo).^{1,2} A network meta-analysis allows for a quantitative synthesis of the network by combining direct evidence from comparisons of interventions within randomised trials and indirect evidence across randomised trials on the basis of a common comparator.³⁻⁶

Networks of trials and network meta-analyses are increasingly used to evaluate healthcare interventions.⁷⁻¹¹ Compared with pairwise meta-analyses, network meta-analyses allow for visualisation of a larger amount of evidence, estimation of the relative effectiveness among all interventions (even if some head to head comparisons are lacking), and rank ordering of the interventions.

The conduct of network meta-analyses poses multiple challenges. Several reviews have explained how to approach these challenges in practice¹¹⁻¹³; others have evaluated how indirect comparison meta-analyses and network meta-analyses have been conducted and reported.^{8 9 14 15} Most reviews have focused on checking the validity of key assumptions required in network meta-analyses (homogeneity, similarity/transitivity, consistency, or an overarching assumption of exchangeability); others have assessed the presence of essential methodological components of the systematic review process (for example, conducting a literature search and assessing the risk of bias of individual studies). These reviews did not evaluate the presentation of findings in reports of network meta-analyses.

Although reporting guidelines have been developed for systematic reviews and pairwise meta-analyses, we lack reporting guidelines for network meta-analyses. An international

Correspondence to: A Bafeta aida.bafeta@htd.aphp.fr

Extra material supplied by the author (see <http://www.bmj.com/content/348/bmj.g1741?tab=related#webextra>)

group is preparing an extension of the PRISMA statement for network meta-analyses. Inadequate reporting of findings and inadequate evaluation of the required assumptions may impede confidence in the findings and conclusions of network meta-analyses.¹⁶ Reports of network meta-analyses should allow the reader to assess the amount of evidence in the network of randomised trials and the relative effect sizes between interventions, along with their uncertainty.^{9 17}

On the basis of a previous methodological systematic review of published reports of network meta-analyses,¹⁰ we aimed to examine how network meta-analysis results are reported.

Methods

Selection of articles

The search strategy and the selection criteria for reports of network meta-analyses have been described elsewhere.¹⁰ In brief, we considered networks of randomised trials that assessed three or more treatments but excluded meta-analyses with an open loop network of three interventions. The articles were screened from a sample of 893 potentially relevant publications indexed in the Cochrane Database of Systematic Reviews, the Database of Abstracts of Reviews of Effects, Medline, and Embase. The date of the last search was 12 July 2012. We identified three additional network meta-analyses from methodological articles. Two independent reviewers excluded non-network meta-analysis reports on the basis of the title and abstract of retrieved citations, and the full text articles were then evaluated according to pre-specified eligibility criteria. Disagreements were discussed to reach consensus.

Data extraction

We used a standardised data collection form to collect all data, for general characteristics and those pertaining to reporting of results, from 121 original reports of network meta-analyses and all supplementary appendices referenced in the articles, when available. Two reviewers independently extracted all data for a random sample of 20% of reports, except items that involved subjective interpretation, which were extracted independently and in duplicate for all reports. These items were the information on the structure of the network, whether the authors applied a Bayesian approach, which types of prior distributions were used, and whether the authors applied a method to investigate the impact of effect modifiers. Disagreements were resolved by discussion.

General characteristics

We collected data on the following general characteristics: type of journal (general or specialty), funding source (we classified network meta-analyses as having public funding if they were explicitly not funded and were conducted by academic authors or were funded by a public source, private funding if they were funded by at least one private source, and unclear funding if the authors did not report any information on the funding), type of interventions included in the network (pharmacological, non-pharmacological (defined as any intervention that did not include an active substance), or both), the type of outcome (binary, continuous, or time to event) and number of outcomes assessed per network meta-analysis, the number of printed pages (excluding the supplement or appendices), existence of a supplement or appendix, and the number of tables and figures (including the supplement or appendices).

Statistical aspects

We evaluated whether the authors applied a Bayesian approach, because this affects the reporting of results such as treatment rankings. In that case, we also assessed whether the authors used non-informative or informative priors for the basic parameters and the between trial variance.¹⁸ We also assessed whether the authors reported a method to investigate non-transitivity, a key to the validity of a network meta-analysis, which occurs when studies are not sufficiently similar in important clinical and methodological characteristics (effect modifiers).¹¹ In accordance with Donegan and colleagues,¹⁹ we considered two methods: comparing patients' or trials' characteristics across treatment comparisons (for example, using descriptive statistics describing the characteristics across the comparisons) and investigating potential treatment effect modifying covariates (for example, using subgroup analysis or network meta-regression analysis).

Presentation of results

In accordance with recommendations from Ioannidis,²⁰ Mills and colleagues,²¹ Cipriani and colleagues,¹¹ and Ades and colleagues,¹² we chose items to assess the reporting of the amount of evidence in the network of randomised trials, and of effect size estimates for pairwise comparisons between interventions. Furthermore, we assessed the reporting of rank ordering of interventions.

Amount of evidence in network of randomised trials

We assessed whether the reports included data to allow for gauging of the description of the network and its geometry. In particular, we assessed whether each report allowed for explicit identification of the interventions included in the network, the direct comparisons between the interventions, and the number of randomised trials and patients for each comparison. We also assessed whether a network graph or diagram was reported. In cases of star networks (that is, all interventions were compared with a common comparator but not against each other) and fully connected networks (all interventions were compared with each other), we considered that the network structure was clear provided that the authors explicitly mentioned the interventions and direct comparisons between the interventions included in the network.

Effect size estimates for pairwise comparisons between interventions

We evaluated whether the following components were reported: results of all individual randomised trials used in the network meta-analysis (summary data for each intervention group or estimates of treatment effect) and effect estimates with measures of uncertainty (such as confidence intervals, standard error, variance) from pairwise meta-analyses (direct evidence), indirect comparison meta-analyses (indirect evidence), or network meta-analyses (mixed evidence). In the case of the last of these, we assessed the reporting of all pairwise comparisons between interventions or only some selected pairwise comparisons between interventions (for example, all comparisons with placebo or reference intervention). We also assessed how these effect sizes were reported. Possible formats included a table, a matrix, a forest plot, a network graph (each edge labelled with the corresponding effect size), or narrative reporting in the results section.

Rank ordering of interventions

We assessed whether the reports gave some form of rank ordering of interventions. We assessed whether the authors presented the probabilities of each intervention being the best, the probabilities of each intervention taking each possible rank (in an absolute or cumulative ranking curve), or the mean ranks (or the surface under the cumulative ranking curve) and whether authors stated which intervention was the best. We also assessed whether the authors considered the uncertainty in rank ordering (by reporting confidence/credible intervals for the ranks or by reporting all probabilities for each intervention with each possible rank, in an absolute or cumulative ranking curve²²).

Composite outcome assessing inadequate reporting

We built a composite outcome describing inadequate reporting of results based on guidelines from Ioannidis,²⁰ Mills and colleagues,²¹ Ades and colleagues,¹² and Cipriani and colleagues.¹¹ We considered reporting to be inadequate in the case of absence of one of the following items: no reporting of interventions included in the network, direct comparisons between interventions, and number of randomised trials and patients for each comparison; or no reporting of results from direct, indirect, or mixed evidence.

Statistical analysis

We summarised quantitative data as medians and interquartile ranges and categorical data as numbers and percentages. We used the Fisher exact test for categorical data to analyse the composite outcome. All tests were two sided, and we considered $P < 0.05$ to be significant. We used SAS 9.3 for statistical analyses.

Results

General characteristics

We identified 121 eligible reports of network meta-analyses (table 1): 55 (45%) published in general journals and 66 (55%) in specialty journals; 55 (45%) network meta-analyses were funded by public sources and 48 (40%) by at least one private source. A total of 100 (83%) reports assessed pharmacological interventions. The median number of outcomes assessed per network meta-analysis was 2 (interquartile range 1-3). In all, 70 (58%) reports of network meta-analyses included a supplement or appendix.

Statistical aspects

Overall, 91 (75%) network meta-analyses applied a Bayesian approach: 51 used non-informative priors for the basic parameters, and 30 used them for the between trial variance. In all, 43 (36%) used a method to assess non-transitivity.

Amount of evidence in the network of randomised trials

Of 121 network meta-analyses, 52 (43%) did not report the number of randomised trials for each comparison and 96 (79%) did not report the number of patients (table 2). Overall, 100 (83%) network meta-analyses did not give a description of the network (that is, the interventions included in the network, direct comparisons between interventions, number of randomised trials and patients for each comparison) and its geometry (a network graph or information needed to know the network structure) (table 2, figure, appendix table A).

Effect size estimates for pair-wise comparisons between interventions

Of 121 reports of network meta-analyses, 53 (44%) did not give the results of all individual randomised trials used in the network meta-analysis. A total of 48 (40%) did not report the effect sizes for all pairwise comparisons between interventions with a measure of uncertainty derived from pairwise meta-analyses (that is, direct evidence), and 108 (89%) did not give indirect evidence (the effect sizes for all pairwise comparisons between interventions with a measure of uncertainty from indirect evidence). In all, 109 (90%) reports did not give both direct and indirect evidence and 71 (59%) did not give both direct and mixed evidence (that is, the effect sizes for all pairwise comparisons between interventions with a measure of uncertainty from network meta-analyses); 112 (93%) did not describe direct as well as indirect and mixed evidence (table 2, figure, appendix table A).

Rank order of interventions

Among 91 network meta-analyses with a Bayesian approach, 52 reports gave the rank order of interventions; 40 of these gave the probabilities of each intervention being the best, 7 used absolute or cumulative rankograms, and 5 stated which intervention was the best. However, 43 of the 52 reports did not report a measure of uncertainty in ranking (table 2). When reported, the uncertainty in ranking was assessed by absolute or cumulative rankograms (7/9) or confidence/credible intervals (2/9).

Composite outcome assessing inadequate reporting

Overall, 119 (98%) reports of network meta-analyses showed inadequate reporting of results (that is, the authors did not report the interventions included in the network or direct comparisons between the interventions; number of randomised trials and patients for each comparison; or results from direct, indirect, or mixed evidence). This finding did not differ by journal type (general journals: 100%, 95% confidence interval 94% to 100%; specialty journals: 97%, 93% to 100%; $P = 0.5$) or funding source (public: 98%, 94% to 100%; private: 100%, 93% to 100%; $P = 0.53$).

Discussion

Our sample of 121 network meta-analyses included a large number of recently published network meta-analyses covering a wide range of medical areas. Essential components of reporting results of network meta-analyses were missing in most of the articles.

An important element in understanding a network meta-analysis is assessing the network geometry.^{20,21} Firstly, knowing precisely the shape of the network (that is, which interventions are included in the network and which interventions have been compared against each other in randomised trials) is a preliminary step for researchers and readers to assess the available evidence base. The geometry of the network may show that some (head to head) comparisons have been ignored or that some comparators were preferred.²³⁻²⁶ Identifying trials that have not been conducted may help in designing the research agenda. Secondly, the reader should consider the number of randomised trials and patients for each comparison. In fact, differences in the numbers and sizes of trials across interventions and comparisons may affect the reliability of some network meta-analyses' estimates.²⁷ The amount of available evidence

may be imbalanced because of differential reporting bias across comparisons, which may substantially affect the results.²⁸

Thirdly, readers should be able to assess whether similar but not identical interventions were grouped, which may weaken the interpretability of a network meta-analysis. For transparency, authors of network meta-analyses should clearly report the interventions included in the network, the direct comparisons between the interventions, and the number of randomised trials and patients for each comparison and should give a graphical representation of the network.¹¹ However, only 17% of our reports of network meta-analyses adequately described the network and its geometry.

The most important element for readers may be the assessment of the relative effectiveness of the considered interventions. When direct and indirect evidence is available, a potential difficulty is choosing among direct, indirect, and mixed evidence, especially with discrepancies between direct and indirect evidence. Checking consistency is a preliminary step before proceeding to the network meta-analysis.²⁹ Empirical studies have detected statistically significant inconsistency between direct and indirect comparisons in 14% of 112 single loops of evidence and 2-9% of 303 loops from complete networks, depending on the measure of effect and the method for estimating heterogeneity.³⁰⁻³¹ However, excluding inconsistency on the basis of the results of statistical tests is difficult because the results may suggest no major inconsistency although it may be present. Moreover, statistically non-significant inconsistency does not necessarily imply clinical consistency.³⁰ Consequently, readers should be able to compare effect estimates from direct, indirect, and mixed evidence, even in the absence of statistical inconsistency. Indirect evidence may be considered of lower quality than evidence from head to head trials because it is prone to similarity/transitivity problems.³² However, in some cases, indirect evidence may be more reliable than direct and mixed evidence,³³⁻³⁵ and when trials are similarly biased an indirect comparison is not biased.³⁶ Transparency is important to allow readers to make their own judgment according to explicitly presented discrepancies. However, only 10% of our network meta-analysis reports gave both direct and indirect evidence and only 7% gave direct, indirect, and mixed evidence.

One appealing feature of a network meta-analysis is the rank ordering of interventions. This information is attractive for clinical researchers and physicians because it seems to respond to their main concern: determining the best available intervention. However, an intervention may be ranked highly even though it was assessed in only a few trials and with a few patients, which resulted in a misleadingly strong endorsement for the intervention despite large uncertainty. Discussion is ongoing as to whether rankings must be reported and are appropriate to derive inference. Clinical researchers and physicians should always be able to examine the uncertainty in the ranking of interventions, because the difference between interventions might be small and not clinically relevant. However, only 17% of our network meta-analysis reports that ranked interventions reported the uncertainty.

Strengths and limitations of study

To the best of our knowledge, no previous study has tackled the problem of presentation of results in a large sample, such as 121 network meta-analyses. Recently, the Agency for Healthcare Research and Quality assessed the statistical methods used in a sample of 34 closed loop Bayesian mixed treatment comparisons,^{15,37} and Tan et al explored the presentational approaches used in the United Kingdom for reporting evidence

synthesis of 19 reports of studies involving indirect and mixed treatment comparisons.³⁸ These authors focused mainly on the statistical methods or possible formats of results but did not assess reporting of the amount of evidence in the network of trials (that is, the interventions included in the network, direct comparisons between the interventions, number of trials and patients for each comparison, network graph/or diagram) or the reporting of effect estimates with uncertainty from direct, indirect, or mixed evidence.

Our study had some limitations. Several journals have space constraints, so the study authors may have omitted important details from their reports or key information may have been deleted during the publication process. Nevertheless, most journals now allow an online appendix for extended descriptions of methods and results, and we assessed all data from both the original reports of network meta-analyses and the supplementary appendices when available. Another limitation is that we searched only for reports of network meta-analyses published in journals and did not search for health technology assessment (HTA) reports. Network meta-analyses are increasingly being used to support HTAs. However, identifying network meta-analyses in HTA reports by usual search strategies is difficult. The reporting of results may differ in HTA reports and journal articles. Finally, we built a composite outcome based on the available recommendations from Ioannidis,²⁰ Mills and colleagues,²¹ Ades and colleagues,¹² and Cipriani and colleagues,¹¹ as a possible core set of information that we would like to see in any report of a network meta-analysis. However, we acknowledge that consensus is lacking on which items should be required in network meta-analysis reports. This is an area of ongoing debate.

Conclusions

Our study shows that the results of network meta-analyses are heterogeneously reported. These results might reflect the fact that we lack a general consensus on what should be reported in a network meta-analysis. This review reinforces the need to develop reporting guidelines for network meta-analyses to facilitate their reporting. It may provide a basis for the future development of a quality assessment tool for reporting network meta-analyses and aiding appropriate interpretation.

We thank Elise Diard (French Cochrane Center) and Monica Serrano (monicaserrano.com) for their help with the infographics and Laura Smales for language revision of manuscript. We are particularly indebted to Josefin Blomkvist for data extraction of items on the Bayesian approach and type of prior distribution for the basic parameters and the variance.

Contributors: AB was involved in the study conception, search for trials, selection of trials, data extraction, data analysis, interpretation of results, and drafting the manuscript. LT was involved in the study conception, search for trials, selection of trials, interpretation of results, and drafting the manuscript. RS was involved in data extraction, interpretation of results, and drafting the manuscript. PR was involved in study conception, interpretation of results, and drafting the manuscript. PR is the guarantor.

Funding: AB was funded by an academic grant for doctoral students from Pierre et Marie Curie University, France. Our team is supported by an academic grant for the programme Equipe Espoirs de la Recherche, from the Fondation pour la Recherche Médicale, France. The funding agencies had no role in the design or conduct of the study; collection, management, analysis, or interpretation of the data; or preparation and review of the manuscript.

What is already known on this topic

The findings from network meta-analyses may be difficult for clinicians and decision makers to interpret
 Inadequate reporting of results may affect the interpretation of network meta-analyses and mislead clinical researchers
 Reports of network meta-analysis should allow the reader to assess the amount of evidence in the network of trials and the relative effectiveness of the considered interventions

What this study adds

The findings from network meta-analyses are heterogeneously reported
 These results might reflect the lack of a general consensus on what should be reported in a network meta-analysis
 A clear need exists to extend reporting guidelines to network meta-analyses to improve the reporting of network meta-analyses

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Not needed.

Transparency: PR (the manuscript's guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Data sharing: No additional data available.

- Salanti G, Kavvoura FK, Ioannidis JP. Exploring the geometry of treatment networks. *Ann Intern Med* 2008;148:544-53.
- Ioannidis JP, Karassa FB. The need to consider the wider agenda in systematic reviews and meta-analyses: breadth, timing, and depth of the evidence. *BMJ* 2010;341:c4875.
- Higgins JP, Whitehead A. Borrowing strength from external trials in a meta-analysis. *Stat Med* 1996;15:2733-49.
- Lumley T. Network meta-analysis for indirect treatment comparisons. *Stat Med* 2002;21:2313-24.
- Salanti G, Higgins JP, Ades AE, Ioannidis JP. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008;17:279-301.
- Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004;23:3105-24.
- Jansen JP, Fleurence R, Devine B, Iztler R, Barrett A, Hawkins N, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. *Value Health* 2011;14:417-28.
- Donegan S, Williamson P, Gamble C, Tudur-Smith C. Indirect comparisons: a review of reporting and methodological quality. *PLoS One* 2010;5:e11054.
- Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009;338:b1147.
- Bafeta A, Trinquart L, Seror R, Ravaud P. Analysis of the systematic reviews process in reports of network meta-analyses: methodological systematic review. *BMJ* 2013;347:f3675.
- Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. *Ann Intern Med* 2013;159:130-7.
- Ades AE, Caldwell DM, Reken S, Welton NJ, Sutton AJ, Dias S. Evidence synthesis for decision making 7: a reviewer's checklist. *Med Decis Making* 2013;33:679-91.
- Hoaglin DC, Hawkins N, Jansen JP, Scott DA, Iztler R, Cappelleri JC, et al. Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 2. *Value Health* 2011;14:429-37.
- Edwards SJ, Clarke MJ, Wordsworth S, Borrill J. Indirect comparisons of treatments based on systematic reviews of randomised controlled trials. *Int J Clin Pract* 2009;63:841-54.
- Sobieraj DM, Cappelleri JC, Baker WL, Phung OJ, White CM, Coleman CI. Methods used to conduct and report Bayesian mixed treatment comparisons published in the medical literature: a systematic review. *BMJ Open* 2013;3:e003111.
- Abdelhamid AS, Loke YK, Parekh-Bhurke S, Chen YF, Sutton A, Eastwood A, et al. Use of indirect comparison methods in systematic reviews: a survey of Cochrane review authors. *Research Synthesis Methods* 2012;3:71-9.
- Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ* 2003;326:472.
- Thorlund K, Thabane L, Mills EJ. Modelling heterogeneity variances in multiple treatment comparison meta-analysis—are informative priors the better solution? *BMC Med Res Methodol* 2013;13:2.
- Donegan S, Williamson P, D'Alessandro U, Tudur Smith C. Assessing key assumptions of network meta-analysis: a review of methods. *Research Synthesis Methods* 2013;4:291-323.
- Ioannidis JP. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. *CMAJ* 2009;181:488-93.
- Mills EJ, Thorlund K, Ioannidis JP. Demystifying trial networks and network meta-analysis. *BMJ* 2013;346:f2914.
- Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol* 2011;64:163-71.
- Lathyrus DN, Patsopoulos NA, Salanti G, Ioannidis JP. Industry sponsorship and selection of comparators in randomized clinical trials. *Eur J Clin Invest* 2010;40:172-82.
- Estellat C, Ravaud P. Lack of head-to-head trials and fair control arms: randomized controlled trials of biologic treatment for rheumatoid arthritis. *Arch Intern Med* 2012;172:237-44.
- Rizos EC, Salanti G, Kontoyiannis DP, Ioannidis JP. Homophily and co-occurrence patterns shape randomized trials agendas: illustration in antifungal agents. *J Clin Epidemiol* 2011;64:830-42.
- Stamatakis E, Weiler R, Ioannidis JP. Undue industry influences that distort healthcare research, strategy, expenditure and practice: a review. *Eur J Clin Invest* 2013;43:469-75.
- Thorlund K, Mills EJ. Sample size and power considerations in network meta-analysis. *Syst Rev* 2012;1:41.
- Trinquart L, Abbe A, Ravaud P. Impact of reporting bias in network meta-analysis of antidepressant placebo-controlled trials. *PLoS One* 2012;7:e35219.
- Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE. Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. *Med Decis Making* 2013;33:641-56.
- Song F, Xiong T, Parekh-Bhurke S, Loke YK, Sutton AJ, Eastwood AJ, et al. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study. *BMJ* 2011;343:d4909.
- Veroniki AA, Vasiladi HS, Higgins JP, Salanti G. Evaluation of inconsistency in networks of interventions. *Int J Epidemiol* 2013;42:332-45.
- Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 8. Rating the quality of evidence—indirectness. *J Clin Epidemiol* 2011;64:1303-10.
- Song F, Harvey I, Lilford R. Adjusted indirect comparison may be less biased than direct comparison for evaluating new pharmaceutical interventions. *J Clin Epidemiol* 2008;61:455-63.
- Alkhafaji AA, Trinquart L, Baron G, Desvarieux M, Ravaud P. Impact of evergreening on patients and health insurance: a meta analysis and reimbursement cost analysis of citalopram/escitalopram antidepressants. *BMC Med* 2012;10:142.
- Madan J, Stevenson MD, Cooper KL, Ades AE, Whyte S, Akehurst R. Consistency between direct and indirect trial evidence: is direct evidence always more reliable? *Value Health* 2011;14:953-60.
- Song F, Clark A, Bachmann MO, Maas J. Simulation evaluation of statistical properties of methods for indirect and mixed treatment comparisons. *BMC Med Res Methodol* 2012;12:138.
- Coleman CI, Phung OJ, Cappelleri JC, Baker WL, Kluger J, White CM, et al. Use of mixed treatment comparisons in systematic reviews. AHRQ Publication No. 12-EHC119-EF. Agency for Healthcare Research and Quality, 2012 (available from www.effectivehealthcare.ahrq.gov/reports/final.cfm).
- Tan SH, Bujkiewicz S, Sutton A, Dequen P, Cooper N. Presentational approaches used in the UK for reporting evidence synthesis using indirect and mixed treatment comparisons. *J Health Serv Res Policy* 2013;18:224-32.

Accepted: 16 February 2014

Cite this as: *BMJ* 2014;348:g1741

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>.

Tables

Table 1 | General characteristics of 121 reports of network meta-analyses

Item and subcategory	No (%) of reports
Journal type:	
General journal	55 (45)
Specialty journal	66 (55)
Funding source:	
Private (at least one private source)	48 (40)
Public	55 (45)
Unclear	18 (15)
Type of intervention assessed:	
Pharmacological intervention	100 (83)
Non-pharmacological intervention	11 (9)
Both (pharmacological and non-pharmacological intervention)	10 (8)
Bayesian statistical approach:	
Basic parameter:	
Non-informative prior	51/91 (56)
Informative prior	0 (0)
Not reported	40/91 (44)
Between trial variance:	
Non-informative prior	30/91 (33)
Informative prior	4/91 (4)
Not reported	57/91 (63)
Methods to assess non-transitivity:	
Comparing patients' or trials' characteristics (across comparisons)	40/43 (93)
Investigating potential intervention effect modifying covariates	3/43 (7)
Type of outcomes:	
Binary	78 (64)
Continuous	28 (23)
Time to event:	15 (12)
No of outcomes assessed per network*	2 (1-3)
No of printed pages*	11 (9-15)
Supplement or appendix published	70 (58)
No of tables per network*	5 (3-6)
No of figures per network*	3 (2-5)

*Data are medians (interquartile range).

Table 2| Presentation of results in 121 reports of network meta-analysis. Values are numbers (percentages)

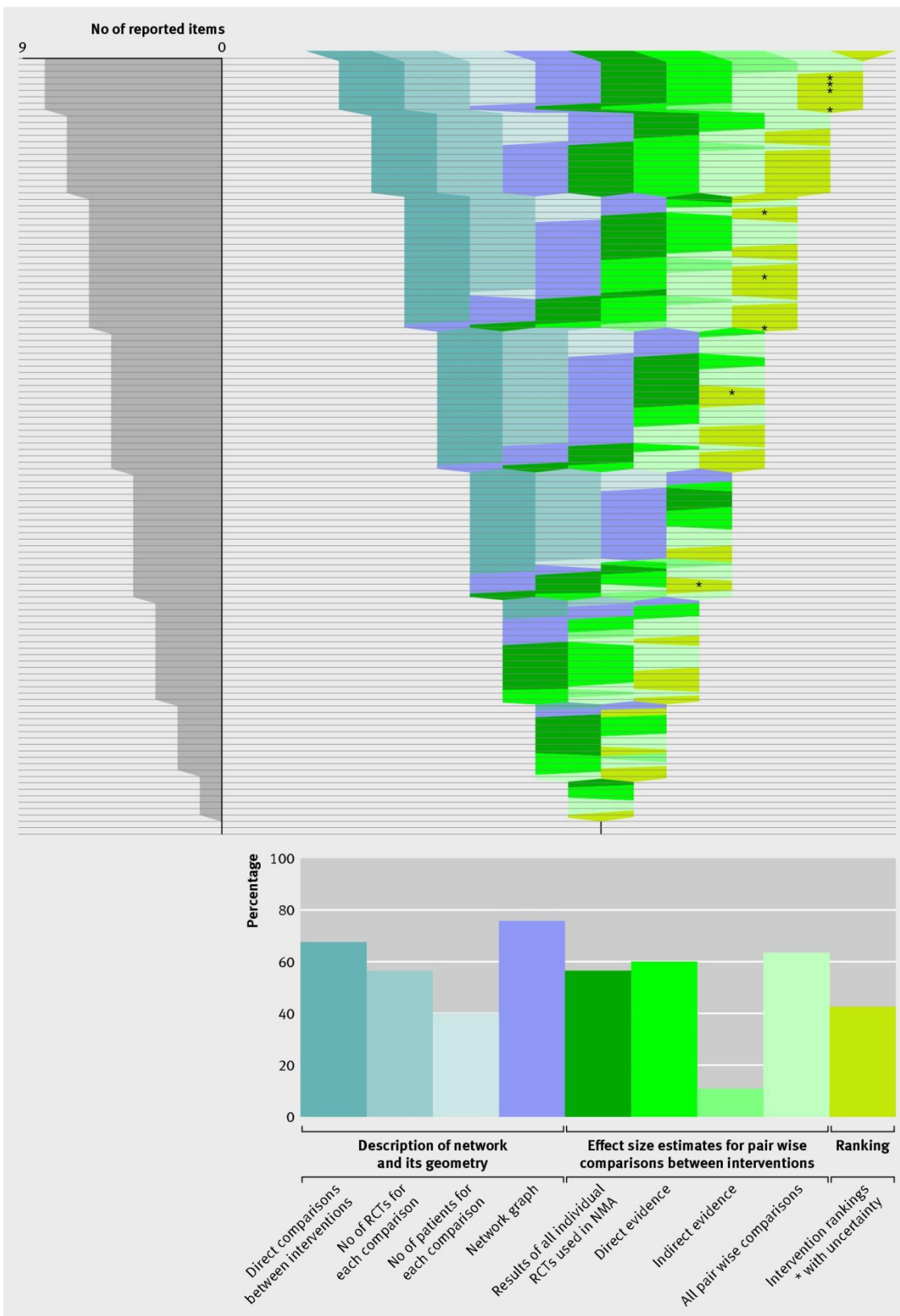
Items—yes if reported	Overall (n=121)	General journals (n=55)	Specialty journals (n=66)
Amount of evidence in network of randomised trials			
Interventions included in network	121 (100)	55 (100)	66 (100)
Direct comparisons between interventions	82 (68)	39 (71)	43 (65)
No of trials for each comparison	69 (57)	34 (62)	35 (53)
No of patients for each comparison	25 (21)	14 (25)	11 (17)
Network graph*	89 (74)	44 (80)	45 (68)
Effect size estimates for pairwise comparisons between interventions			
Results of all individual trials used in network meta-analysis	68 (56)	28 (51)	40 (61)
Direct evidence	73 (60)	33 (60)	40 (61)
Indirect evidence	13 (11)	6 (11)	7 (11)
Mixed evidence	78 (64)	34 (62)	44 (67)
Effect estimates for some selected pairwise comparisons between interventions from network meta-analysis	43 (36)	21 (38)	22 (33)
Format of results†:			
Table	41 (34)	24 (44)	17 (26)
Forest plot	7 (6)	3 (5)	4 (6)
Matrix	6 (5)	3 (5)	3 (5)
Figure	4 (3)	2 (4)	2 (3)
Text			
Rank order of interventions			
Rank order of interventions‡	52/91 (57)	22/38 (58)	30/53 (57)
Rank order of interventions with uncertainty	9/52 (17)	6/22 (27)	3/30 (10)

*7 star networks and 4 fully connected networks reported no network graph, but authors explicitly mentioned interventions and direct comparisons in network.

†Multiple answers were possible, so total does not equal 100%.

‡Network meta-analyses that did not use Bayesian approach cannot derive intervention rankings; 91 network meta-analyses used Bayesian approach.

Figure



Reporting items for results of network meta-analyses. The idea is to show which items were reported in each of the 121 network meta-analyses. Each horizontal line of the gap chart corresponds to one network meta-analysis report. A specific colour was attributed to each of the 10 items studied. The colour bands show which of these items (labelled at the top) were reported for each network meta-analysis. Items were grouped into three categories: 4 items in blue pertain to the description of network, and 1 item in purple pertains to its geometry (network graph); 4 items in green pertain to effect size estimates for pairwise comparisons between interventions; 1 item pertains to intervention ranking. The 121 network meta-analysis reports were sorted according to the total number of reported items, in decreasing order. The diagram on the left shows the distribution of the total number of reported items across the 121 network meta-analyses. The diagram at the bottom shows the proportion of network meta-analysis reports that reported each item

4. Chapitre: Discussion et Perspectives

4.1. Discussion

4.1.1. Principaux résultats

La fréquence d'utilisation des méta-analyses en réseau a considérablement augmenté ces dernières années, et cela dans de nombreux domaines de la médecine^{103,176}.

Actuellement, il existe une littérature abondante sur les outils statistiques utilisés dans les méta-analyses en réseau, par exemple la vérification des hypothèses (cohérence, similarité, hétérogénéité). Par exemple une revue de la littérature publiée en 2014 portant sur l'ensemble des études évaluant la qualité des méta-analyses en réseau et des comparaisons indirectes a montré que la majorité des études portaient essentiellement sur les concepts statistiques¹⁷⁷. Ainsi, il y a peu de recherche sur les éléments de revues systématiques comme les critères d'éligibilités, la recherche bibliographique, ou l'évaluation du risque de biais, dans les méta-analyses en réseau.

Par définition la méta-analyse en réseau est une méta-analyse, donc doit se réaliser dans le cadre rigoureux des revues systématiques pour limiter les risques de biais¹⁵⁸. Certains aspects dans la conduite et la présentation des méta-analyses ont une influence directe sur la validité des conclusions de celles-ci et leurs interprétations.

C'est pour cela que nous avons orienté ce travail de thèse sur :

- 1) l'évaluation des caractéristiques essentielles de la revue systématique dans les méta-analyses en réseau. Nous avons montré que bien la méta-analyse en réseau reste une méta-analyse, elle ne suit pas les recommandations essentielles à la conduite et la présentation des revues systématiques (Notamment, les étapes concernant la recherche bibliographique, l'évaluation de la qualité méthodologique des essais inclus, le biais de publication).
- 2) Dans un deuxième temps, nous avons évalué comment sont présentés les résultats dans les méta-analyses en réseau. Nous avons mis en évidence que la présentation des résultats dans

les méta-analyses en réseau était très hétérogène. Les auteurs des méta-analyses en réseau ne présentent pas ou rapportent de manière inadéquate les éléments essentiels permettant d'évaluer l'ensemble des preuves disponibles (les interventions incluses, les comparaisons directes incluses, le nombre d'essais et de patients inclus dans chaque branche du réseau qui permettent de décrire la forme du réseau, et les estimations deux à deux provenant des comparaisons directes, indirectes ou mixtes).

4.1.2. Implication

Il est important de pouvoir évaluer la qualité de réalisation d'une méta-analyse en réseau pour déterminer si les conclusions sont fiables. Pour d'améliorer la transparence et la reproductibilité, il est nécessaire de développer et promouvoir des recommandations pour la conduite et la présentation de ce type d'étude.

4.1.2.1. La recherche bibliographique dans les méta-analyses en réseau

Nos travaux ont montré que les recherches bibliographiques dans les méta-analyses en réseau étaient incomplètes. Les études non publiées -notamment les études en cours- ne sont pas recherchées. Identifier toutes les études pertinentes dans un sujet donné avec des interventions données, et des comparateurs donnés est complexe et nécessite du temps et des moyens plus importants que dans une méta-analyse classique. Au delà du biais de publication qui a tendance à surestimer les effets bénéfiques des traitements et à contrario de sous-estimer leurs toxicités^{52,75,178,179}, les résultats méta-analyses en réseau peuvent être biaisés par une mauvaise identification des essais à inclure. Les cliniciens et chercheurs sont donc potentiellement induits en erreur sur l'efficacité et l'innocuité des traitements. Des recherches méthodologiques sont nécessaires pour examiner l'impact de ces biais sur l'analyse et les

conclusions des méta-analyses en réseau.

4.1.2.2. L'évaluation du risque de biais dans les méta-analyses en réseau

Un des problèmes majeur de la méta-analyse en réseau est l'évaluation du risque de biais, la majorité des publications ne rapportent pas cette évaluation. Néanmoins, l'évaluation du risque de biais et sa prise en compte dans la méta-analyse en réseau est plus complexe que dans la méta-analyse conventionnelle.

Une différence fondamentale entre la méta-analyse conventionnelle et la méta-analyse en réseau est qu'elle estime tous les effets traitements deux à deux entre les interventions incluses dans le réseau. Ainsi, un biais peut affecter une seule estimation, et avoir des répercussions sur plusieurs estimations. De plus, le risque de biais peut varier selon les différentes comparaisons au sein du réseau. Une récente étude méta-épidémiologie en réseau¹⁸⁰ portant sur 32 réseaux d'essais représentant 613 ECRs, a évalué l'impact des caractéristiques des essais et la précision de la variance sur les résultats de la méta-analyse en réseau ; cette étude a mis en évidence que les essais avec une large variance ont tendance à surestimer l'effet traitement dans la majorité des méta-analyses en réseau. Néanmoins, les résultats de cette étude ne confirment pas que l'absence d'une assignation secrète, d'une séquence de randomisation, et d'aveugle sont en moyenne associés à une surestimation de l'effet traitement. Ces résultats diffèrent des autres études méta-épidémiologiques et notamment de l'étude méta-méta-épidémiologique BRANDO¹⁸¹ (« Bias in Randomized and Observational Studies ») qui combine les données de 7 études méta-épidémiologiques pour étudier l'influence de la méthode pour générer la séquence de randomisation, de l'assignation secrète, de l'aveugle sur l'effet traitement. Comme suggéré par les auteurs, cette différence peut être expliquée par des problèmes de puissance, en effet le nombre d'essais inclus avec des risques de biais est faible. Une autre explication pourrait être l'impact différentiel des

facteurs de risques dans les méta-analyses incluses dans les réseaux.

Comme pour la méta-analyse conventionnelle l'évaluation du risque de biais suscite de nombreuses questions et perspectives de recherche. De futures recherches s'ouvrent pour trouver des solutions pour répondre aux différentes questions : Comment prendre en compte les différentes variations du risque de biais entre les comparaisons directes et indirectes ? Comment évaluer l'impact du risque de biais dans la méta-analyse en réseau ? Comment identifier les biais propres à la méta-analyse réseau ? Comment développer des méthodes pour évaluer, résumer et présenter le risque de biais dans le réseau ?

4.1.2.3. La présentation des résultats dans les méta-analyses en réseau

Comment présenter les résultats de la méta-analyse en réseau en format simple et clair? Quels sont les éléments à produire dans un rapport de méta-analyse en réseau ? Nos travaux de recherche ont montré que les résultats des méta-analyses en réseau sont présentés de manière très hétérogène dans la littérature médicale.

Dans une méta-analyse en réseau, le nombre de comparaisons et d'interventions évaluées est plus élevé que dans une méta-analyse conventionnelle cela augmente le nombre de résultats, de tableaux, de figure. Par exemple un réseau de 10 interventions génère 45 comparaisons deux à deux, et ceci augmente de 105 pour un réseau incluant 15 interventions. De plus, une autre difficulté peut s'ajouté lorsque la méta-analyse en réseau utilise plusieurs critères de jugement. Ainsi les résultats des méta-analyses en réseau peuvent devenir obscurs pour les cliniciens.

Récemment, une étude publiée dans *the journal of clinical epidemiology* par Tan et ses collègues propose 3 formats pour rapporter les résultats d'une méta-analyse en réseau¹⁸². Le premier format est un tableau résumant tous les Forest plots sous forme d'une matrice. L'avantage de ce format est de présenter une vue d'ensemble de toutes les estimations deux à

deux sur le même graphique. Mais ce type de représentation devient vite illisible pour des réseaux de plus de 10 interventions. La figure 16 présente le premier format proposé.

Le deuxième format proposé est un tableau qui résume toutes les estimations deux à deux possibles. L'avantage de ce type de présentation est que la lecture de toutes les estimations deux à deux est facilitée, chaque intervention est associée à un comparateur. Un autre avantage de ce style de présentation est que les réseaux de grande taille peuvent être rapportés plus facilement en s'étendant sur plusieurs pages, le cas échéant. La figure 17 présente le deuxième format proposé.

Le troisième format est un graphique regroupant l'ensemble des rangs moyens des interventions avec un système de dégradé de couleurs du plus sombre (les moins efficaces) au plus clair (les plus efficaces). La figure 18 présente le troisième format proposé.

Aucun de ces 3 formats de présentation des résultats ne rapportent les estimations des comparaisons indirectes. Parmi les 3 formats, à mon sens le deuxième format offre la meilleure représentation car il est facile à lire et plus intuitif, cependant il manque une colonne qui indiquerait le nombre de patients pour chaque comparaison.

Des futures recherches sont nécessaires pour développer des outils graphiques standardisés, simples et accessibles au médecin pour améliorer leurs interprétations des résultats.

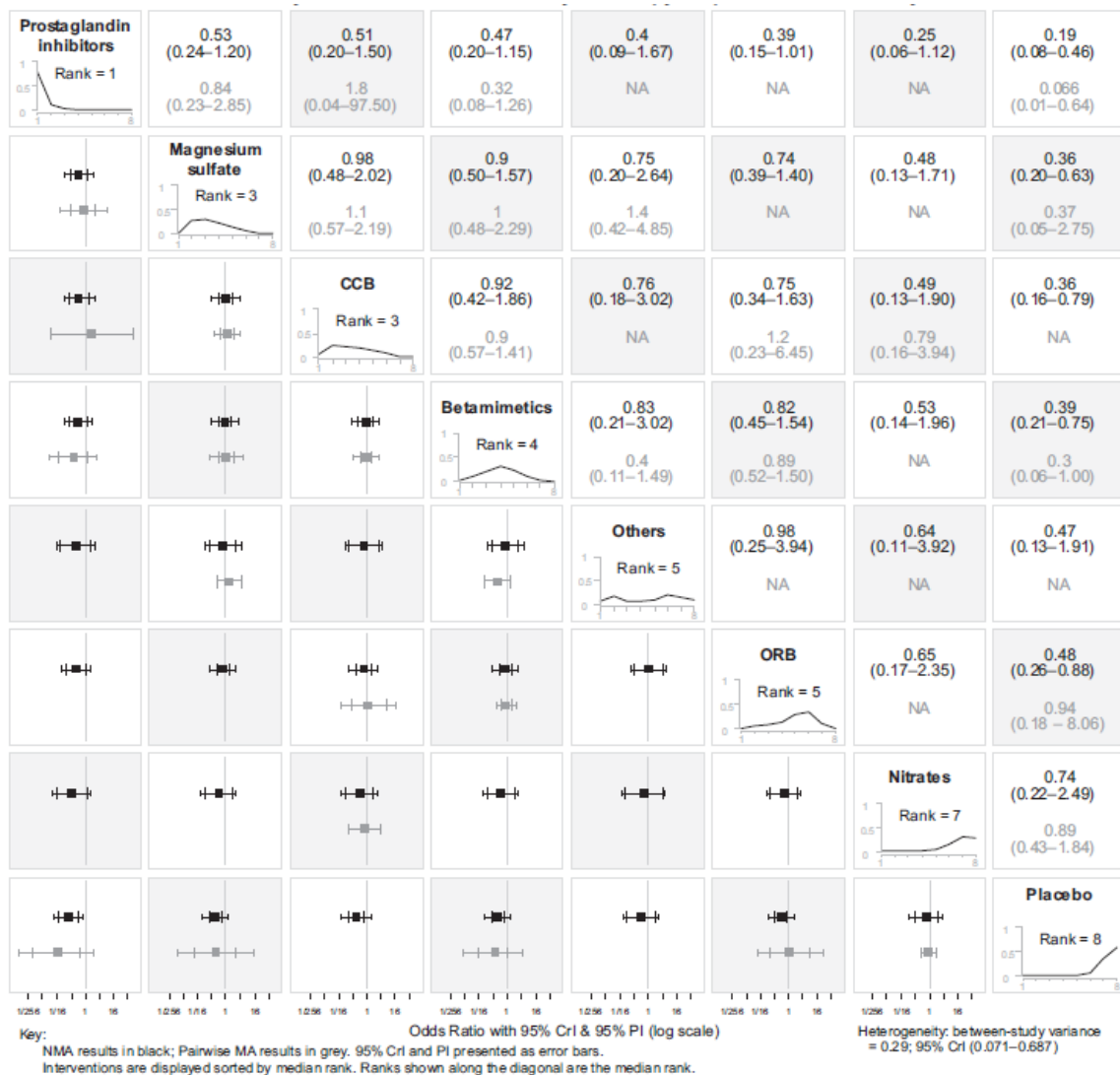


Fig. 1. Summary forest plot matrix. CCB, calcium channel blocker; ORB, oxytocin receptor blocker; NMA, network meta-analysis; CrI, credible interval; PI, prediction interval.

Figure 16 : Représentation de l'ensemble des Forest plots sous le format d'une matrice dans

l'exemple des traitements tocolytiques dans l'accouchement prématuré. Adaptée de la source ¹⁸²

Le long de la diagonale se trouve les interventions incluses dans le réseau avec leur rang médian. Les interventions sont ordonnées en fonction de leur rang médian. En dessous de la diagonale, dans le triangle inférieur toutes les comparaisons possibles deux à deux analysées sont représentées par les Forest plots. Les estimations issues de la méta-analyse en réseau sont représentées en noires et ceux de la comparaison directe en grises. Au dessus de la diagonale, dans le triangle supérieur les estimations numériques correspondantes sont présentées avec les mêmes codes couleurs pour les estimations provenant des méta-analyses en réseau et ceux

provenant des comparaisons directes.

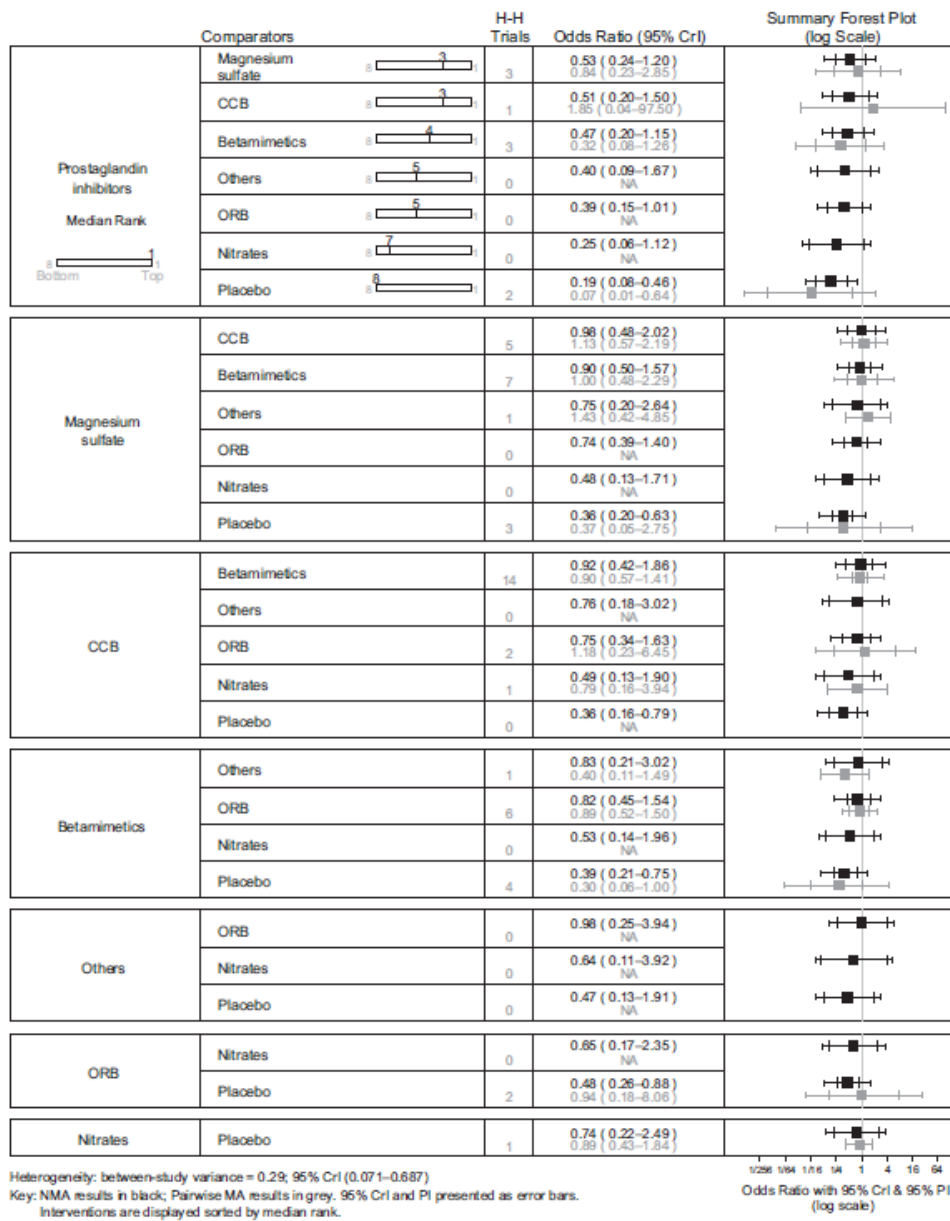


Fig. 2. Summary forest plot table. H-H trials, head-to-head trials; CCB, calcium channel blocker; ORB, oxytocin receptor blocker; NMA, network meta-analysis; CrI, credible interval; PI, prediction interval.

Figure 17 : Représentation de l'ensemble des Forest plots sous le format d'un tableau dans l'exemple des traitements tocolytiques dans l'accouchement prématuré. Adaptée de la source¹⁸²

Le tableau résume tous les résultats de toutes les combinaisons possibles. La première colonne présente une intervention, la seconde colonne présente les comparateurs par rapport à

l'intervention indiquée dans la première colonne. Les interventions ont été ordonnées par leur rang médian. La troisième colonne indique le nombre d'essais pour chaque comparaison. Dans la quatrième colonne, les estimations numériques correspondantes sont présentées. Les estimations issues de la méta-analyse en réseau sont représentées en noires et ceux de la comparaison directe en grises. Enfin, la colonne 5 présente les Forest plots avec les mêmes codes couleurs pour les estimations provenant des méta-analyses en réseau et ceux provenant des comparaisons directes.

Rank	Intervention
1	Prostaglandin inhibitors
2	
3	Calcium channel blockers Magnesium sulfate
4	Betamimetics
5	Others Oxytocin receptor blockers
6	
7	Nitrates
8	Placebo

Figure 18 : Représentation de l'ensemble des rangs moyens dans l'exemple des traitements

tocolytiques dans l'accouchement prématuré. Adaptée de la source ¹⁸²

Le dégradé de couleur du plus sombre (les moins efficaces des médicaments) au plus clair (les plus efficaces des médicaments) est utilisé dans ce graphique pour aider à attirer l'attention sur le(s) plus efficace(s) et le(s) moins efficace(s). Dans cette exemple, Les inhibiteurs de la prostaglandine est plus susceptible d'être le meilleure avec un rang médian de 1, tandis que les nitrates et placebo sont les pires, et les 5 autres interventions ont des classements similaires.

Un autre problème majeur de la présentation des résultats est quels éléments sont à rapporter dans un rapport de méta-analyse en réseau.

Il y a un véritable débat concernant les preuves qui doivent être rapportées dans les méta-analyses en réseau : Est ce que toutes les preuves issues des comparaisons directes, indirectes ou mixtes doivent être rapportées ? En effet lorsque les estimations issues des comparaisons directes et indirectes sont disponibles, une des difficultés majeures est de choisir parmi la source directe, indirecte ou mixte, notamment en présence d'incohérence. Vérifier la cohérence est une étape préliminaire avant de réaliser une méta-analyse en réseau. Cependant, il est difficile d'exclure une incohérence à l'aide des tests statistiques¹¹⁹. Par ailleurs, certaines études ont montré que parfois la preuve la moins biaisée peut provenir des comparaisons indirectes comme des comparaisons directes^{91,117,118,146}. De fait, les lecteurs doivent être en mesure de comparer les estimations des comparaisons directes, indirectes et mixtes, même en l'absence d'incohérence statistique.

Il est important de pouvoir évaluer la géométrie de réseau. En effet, elle permet d'évaluer l'ensemble des preuves disponibles en particulier, de repérer : 1) les comparaisons qui ont été ignorées, 2) les comparateurs préférés^{88,97-100}, 3) si des interventions ont été regroupées ou pas. Le lecteur doit être en mesure d'évaluer le nombre d'essais et de patients pour chaque comparaison, car des différences dans le nombre d'essais et de patients entre les différentes comparaisons peuvent affecter la fiabilité des résultats¹⁰⁹.

4.1.3. Conclusion

Les méta-analyses en réseau sont devenues très populaires, Les cliniciens, les agences d'évaluation de santé et les organismes de régulation sont de plus en plus confrontés à ces méthodes de synthèse^{103,134,148,150,176,177}. Une recherche électronique du nombre de citation via Web of Science à la date du 18 août 2014, montre une augmentation constante du nombre de

citations faisant référence à un article de méta-analyse en réseau dans la littérature médicale.

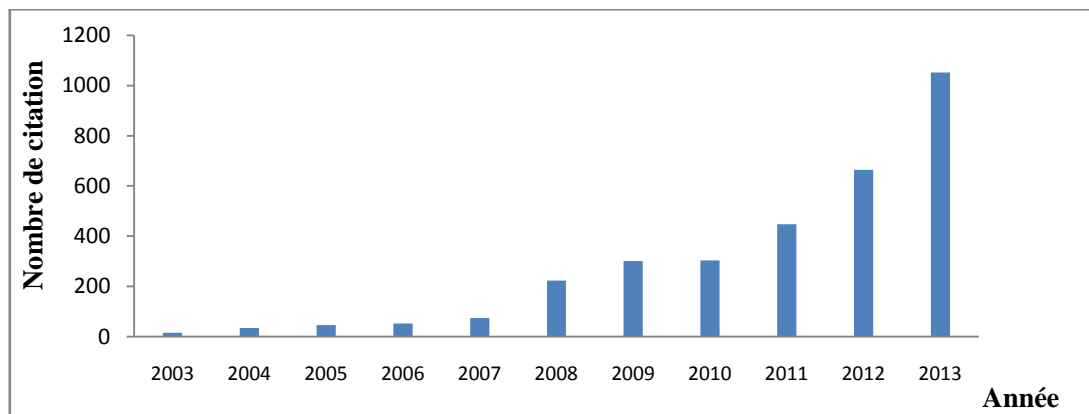


Figure 19 : Evolution du nombre de citations via Web Of Science faisant référence à un article de méta-analyse en réseau dans la littérature médicale.

Le nombre de citation total est de 3211 pour 260 publications.

La méta-analyse en réseau est une technique utile et puissante. Néanmoins, nous avons mis en évidence que dans la plupart des cas, certaines caractéristiques essentielles aux revues systématiques sont négligées. Il y a une incohérence entre le fait que des chercheurs réalisent des méta-analyses conventionnelles dans le cadre de revue systématique et que ces mêmes auteurs réalisent des méta-analyses en réseau en dehors de ce cadre. Cela soulève des doutes sur la fiabilité des résultats et des conclusions des méta-analyses en réseau. Or une recherche basée sur des données dont la fiabilité est remise en question pourrait avoir l'ensemble de sa validité menacée.

Nous espérons que nos travaux ont montré la nécessité de réaliser les méta-analyses en réseau dans le cadre des revues systématiques.

Nous avons décrit les questions conceptuelles et techniques qu'un clinicien chercheur doit se poser pour réaliser et interpréter une méta-analyse en réseau comme la qualité de la recherche bibliographique, comment identifier tous les essais dans une méta-analyse en réseau, comment

évaluer le risque de biais, comment présenter les résultats. La méta-analyse en réseau a le potentiel d'être la meilleure source pour éclairer les prises de décision en médecine. Cependant, un certain nombre de questions concernant sa réalisation ne sont pas résolus et de nombreux défis seront nécessaires pour fournir une synthèse fiable et transparente.

4.2. Perspectives

Nous allons poursuivre nos travaux de recherche sur la méta-analyse en réseau selon plusieurs axes principaux :

Premièrement, la présentation et la communication des résultats d'une méta-analyse en réseau sous un format accessible et compréhensible est un véritable challenge. Dans une méta-analyse en réseau, il est possible de calculer les probabilités que le traitement soit le meilleur, le deuxième ou le troisième parmi tous les traitements inclus dans le réseau. Le classement des interventions est le format le plus simple à comprendre. Cette approche est très intéressante pour les cliniciens, car elle répond à leur préoccupation principale: déterminer le meilleur traitement disponible. Cependant ce classement des interventions doit être interprété avec prudence car la différence entre les traitements pourrait être faible et cliniquement non significative. De plus plusieurs processus tels que le biais de publication, ou l'exclusion d'intervention du réseau peuvent l'affecter¹⁶⁹. Dans nos travaux, nous avons montré que le classement des interventions était utilisé dans 43% des publications avec une majorité qui ne rapportent pas de mesure d'incertitude sur le classement¹⁸³. Dans un premier temps nous allons ré-analyser les méta-analyses en réseau pour obtenir un classement des interventions et calculer un degré d'incertitude. Pour cela, nous avons identifié les essais pour lesquelles les données sont disponibles dans notre échantillon de 121 méta-analyses en réseau. Dans un deuxième temps, nous ont évalué les résumés de ces méta-analyses en réseau qui rapporte un classement des interventions. Dans un troisième temps, un groupe de lecteurs vont évaluer la proportion d'article ayant des différences avec les classements obtenus et ceux présents dans les résumés. Finalement nous souhaitons mener un ECR pour évaluer l'impact d'une présentation inadéquate des classements des interventions sur l'interprétation des cliniciens dans les résumés de méta-analyse en réseau.

Les caractéristiques des études incluses dans une méta-analyse en réseau ont probablement un

impact sur l'estimation des effets traitements. Cependant, peu d'études méta-épidémiologiques ont été réalisées pour évaluer les caractéristiques associées à l'effet traitement dans les méta-analyses en réseau^{119,163}. Nous proposons donc de transposer les méthodes utilisées dans les études méta-épidémiologiques d'essais randomisés pour réaliser des études méta-épidémiologiques en réseau. La méta-épidémiologie en réseau fait l'hypothèse d'un biais homogène en termes d'impact et de direction. L'effet traitement est comparé entre les essais avec et sans la caractéristique d'intérêt au sein de chaque réseau. L'impact moyen de la caractéristique est ensuite évalué sur l'ensemble des réseaux. Nous avons constitué un échantillon de méta-analyses en réseau pour lesquelles nous avons les données des essais inclus. Nous allons pouvoir explorer l'influence de certaines caractéristiques de ces études comme le regroupement d'intervention (par exemple le groupe traitement placebo est regroupé avec le groupe pas d'intervention ou traitement habituel) en utilisant les méthodes statistiques décrites précédemment.

Il est important de pouvoir avoir accès à la recherche, les chercheurs et cliniciens devraient être en mesure de retrouver de manière fiable et rapide les méta-analyses en réseau publiées dans la littérature médicale. La recherche des rapports de méta-analyse en réseau dans Medline pose deux défis. Premièrement, « the National Library of Medicine's Medlars », dans sa procédure d'indexation n'inclut pas "les méta-analyses en réseau" comme un "type de publication ». Deuxièmement, le nombre d'appellations de la méta-analyse en réseau est très varié contrairement à la revue systématique et la méta-analyse. Il est nécessaire de développer des stratégies de recherche adaptées aux différents lecteurs de la méta-analyse en réseau. Nous allons élaborer une base de données d'articles portant sur la méta-analyses en réseau à partir d'une recherche minutieuse à la main dans les principales revues médicales de haut impact factor clinique revues, puis évaluer la pertinence des articles. Nous allons identifier les mots clés, les termes et les stratégies de recherche proposées. Enfin nous allons comparer la

performance de différents termes individuellement, puis des différentes combinaisons de ces termes et enfin les différentes stratégies de recherche pour d'identifier les méta-analyses en réseau.

La médecine fondée sur les preuves nécessite l'utilisation de sources comme les ECRs, les revues systématiques les méta-analyses, et maintenant les méta-analyses en réseau. L'utilisation de ces sources suppose une certaine familiarité avec les méthodes, la présentation des résultats et leurs interprétations. Cependant les méthodes et l'interprétation des résultats de méta-analyses peuvent paraître obscures aux cliniciens. Ainsi, l'évaluation des connaissances, des attitudes et de leurs utilisations professionnelles chez les cliniciens est un élément important de santé publique. A notre connaissance aucunes études n'ont étudié l'utilisation des méta-analyses en réseau chez les médecins. Par conséquent, nous souhaitons décrire les connaissances, les attitudes et le comportement professionnel des médecins à l'égard des méta-analyses en réseau.

Nous réfléchissons à des approches permettant de planifier l'agenda de recherche à l'aide des méta-analyses en réseau. Comme nous l'avons souligné, la géométrie du réseau reflète en partie la façon dont le programme de recherche, dans son ensemble, a été construit. Sa structure a pu être influencée par des biais spécifiques, c'est-à-dire en évitant certaines comparaisons et en préférant d'autres, notamment les comparaisons contre placebo. Les essais cliniques sont généralement conçus pour tester l'efficacité relative de deux ou plusieurs interventions avec une puissance suffisante. Cependant ils sont conçus et réalisés de manière isolés et ne prennent pas en compte l'ensemble de l'information déjà disponible dans le domaine. Nous souhaitons utiliser les caractéristiques du réseau pour pouvoir fournir des informations non seulement sur la taille des essais, le nombre d'essais requis, mais aussi sur les interventions qui devraient être comparées dans un futur essai.

Références

1. Stolberg HO, Norman G, Trop I. Randomized controlled trials. *AJR Am J Roentgenol*. Dec 2004;183(6):1539-1544.
2. Meldrum ML. A brief history of the randomized controlled trial. From oranges and lemons to the gold standard. *Hematol Oncol Clin North Am*. Aug 2000;14(4):745-760, vii.
3. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c869.
4. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLoS Med*. Sep 2010;7(9):e1000326.
5. Cochrane AL. *Effectiveness and Efficiency: Random Reflections on Health Services*: Nuffield Provincial Hospitals Trust.; 1972.
6. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. *JAMA*. Jul 8 1992;268(2):240-248.
7. Fergusson DA, Hebert PC, Mazer CD, et al. A comparison of aprotinin and lysine analogues in high-risk cardiac surgery. *N Engl J Med*. May 29 2008;358(22):2319-2331.
8. Fergusson D, Glass KC, Hutton B, Shapiro S. Randomized controlled trials of aprotinin in cardiac surgery: could clinical equipoise have stopped the bleeding? *Clin Trials*. 2005;2(3):218-229; discussion 229-232.
9. Mulrow CD. Rationale for systematic reviews. *BMJ*. Sep 3 1994;309(6954):597-599.
10. Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. *PLoS Med*. Mar 27 2007;4(3):e78.
11. Glasziou P, Vandenbroucke JP, Chalmers I. Assessing the quality of research. *BMJ*. Jan 3 2004;328(7430):39-41.
12. Excellence NifHaC. *Guide to the Methods of Technology Appraisal* London: NICE 2008.
13. Elwyn G, O'Connor A, Stacey D, et al. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. *BMJ*. Aug 26 2006;333(7565):417.
14. Mays N, Pope C, Popay J. Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field. *J Health Serv Res Policy*. Jul 2005;10 Suppl 1:6-20.
15. Akobeng AK. Understanding systematic reviews and meta-analysis. *Arch Dis Child*. Aug 2005;90(8):845-848.
16. Egger M., Smith G. D., Altman D. G. *Systematic reviews in health care: meta-analysis context*: BMJ Book 2001.
17. Egger M, Smith GD. Meta-Analysis. Potentials and promise. *BMJ*. Nov 22 1997;315(7119):1371-1374.
18. Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. *BMJ*. Dec 6 1997;315(7121):1533-1537.
19. Higgins JPT, Green S, editors. *Cochrane handbook for systematic reviews of interventions* version 5.1.0
Cochrane Collaboration. 2011.
20. Deeks J. J., Higgins J. P., Altman D. G. Analysing data and undertaking meta-analyses. In: Wiley-Blackwell, ed. *Cochrane Handbook for Systematic Review of*

- Interventions*2008:243-296.
21. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ*. 2011;342:d549.
 22. Tao T, Bo L, Chen F, et al. Effect of protective ventilation on postoperative pulmonary complications in patients undergoing general anaesthesia: a meta-analysis of randomised controlled trials. *BMJ Open*. 2014;4(6):e005208.
 23. Egger M, Juni P, Bartlett C, Hoenstein F, Sterne J. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. *Health Technol Assess*. 2003;7(1):1-76.
 24. Royle P, Milne R. Literature searching for randomized controlled trials used in Cochrane reviews: rapid versus exhaustive searches. *Int J Technol Assess Health Care*. Fall 2003;19(4):591-603.
 25. Berlin JA, Golub RM. Meta-analysis as evidence: building a better pyramid. *JAMA*. Aug 13 2014;312(6):603-605.
 26. Sampson M, Barrowman NJ, Moher D, et al. Should meta-analysts search Embase in addition to Medline? *J Clin Epidemiol*. Oct 2003;56(10):943-955.
 27. McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? *Lancet*. Oct 7 2000;356(9237):1228-1231.
 28. Jüni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials In: Egger M, Smith GD, Altman DG, eds. *Systematic reviews in health care: meta-analysis in context*. London: BMJ Books; 2001:87-108.
 29. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928.
 30. Zarychanski R, Abou-Setta AM, Turgeon AF, et al. Association of hydroxyethyl starch administration with mortality and acute kidney injury in critically ill patients requiring volume resuscitation: a systematic review and meta-analysis. *JAMA*. Feb 20 2013;309(7):678-688.
 31. Jacobs WC, Kruyt MC, Verbout AJ, Oner FC. Effect of methodological quality measures in spinal surgery research: a metaepidemiological study. *Spine J*. Apr 2012;12(4):339-348.
 32. Turner RM, Bird SM, Higgins JP. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews. *PLoS One*. 2013;8(3):e59202.
 33. Ciani O, Buyse M, Garside R, et al. Comparison of treatment effect sizes associated with surrogate and final patient relevant outcomes in randomised controlled trials: meta-epidemiological study. *BMJ*. 2013;346:f457.
 34. Contopoulos-Ioannidis DG, Baltogianni MS, Ioannidis JP. Comparative effectiveness of medical interventions in adults versus children. *J Pediatr*. Aug 2010;157(2):322-330 e317.
 35. Herbison P, Hay-Smith J, Gillespie WJ. Different methods of allocation to groups in randomized trials are associated with different levels of bias. A meta-epidemiological study. *J Clin Epidemiol*. Oct 2011;64(10):1070-1075.
 36. Nuesch E, Reichenbach S, Trelle S, et al. The importance of allocation concealment and patient blinding in osteoarthritis trials: a meta-epidemiologic study. *Arthritis Rheum*. Dec 15 2009;61(12):1633-1641.
 37. Nuesch E, Trelle S, Reichenbach S, et al. The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study. *BMJ*. 2009;339:b3244.
 38. Shea B, Moher D, Graham I, Pham B, Tugwell P. A comparison of the quality of

- Cochrane reviews and systematic reviews published in paper-based journals. *Eval Health Prof.* Mar 2002;25(1):116-129.
39. Wood L, Egger M, Gluud LL, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ.* Mar 15 2008;336(7644):601-605.
 40. Bailey KR. Generalizing the results of randomized clinical trials. *Control Clin Trials.* Feb 1994;15(1):15-23.
 41. Flather M, Delahunty N, Collinson J. Generalizing results of randomized trials to clinical practice: reliability and cautions. *Clin Trials.* 2006;3(6):508-512.
 42. Glasziou P, Guyatt GH, Dans AL, Dans LF, Straus S, Sackett DL. Applying the results of trials and systematic reviews to individual patients. *ACP J Club.* Nov-Dec 1998;129(3):A15-16.
 43. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". *Lancet.* Jan 1-7 2005;365(9453):82-93.
 44. Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. *BMJ.* Jul 7 2001;323(7303):42-46.
 45. Goodman S, Dickersin K. Metabias: a challenge for comparative effectiveness research. *Ann Intern Med.* Jul 5 2011;155(1):61-62.
 46. Begg CB, Berlin JA. Publication bias and dissemination of clinical research. *J Natl Cancer Inst.* Jan 18 1989;81(2):107-115.
 47. Dickersin K, Min YI. Publication bias: the problem that won't go away. *Ann N Y Acad Sci.* Dec 31 1993;703:135-146; discussion 146-138.
 48. Dickersin K, Min YI, Meinert CL. Factors influencing publication of research results. Follow-up of applications submitted to two institutional review boards. *JAMA.* Jan 15 1992;267(3):374-378.
 49. Egger M, Smith GD. Bias in location and selection of studies. *BMJ.* Jan 3 1998;316(7124):61-66.
 50. Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database Syst Rev.* 2009(1):MR000006.
 51. Sterne JA, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *J Clin Epidemiol.* Oct 2001;54(10):1046-1055.
 52. Sterne JA, Egger M, Smith GD. Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. *BMJ.* Jul 14 2001;323(7304):101-105.
 53. von Elm E, Rollin A, Blumle A, Huwiler K, Witschi M, Egger M. Publication and non-publication of clinical trials: longitudinal study of applications submitted to a research ethics committee. *Swiss Med Wkly.* Apr 5 2008;138(13-14):197-203.
 54. Ioannidis JP. Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *JAMA.* Jan 28 1998;279(4):281-286.
 55. Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *BMJ.* Sep 13 1997;315(7109):640-645.
 56. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet.* Apr 13 1991;337(8746):867-872.
 57. Gotzsche PC. Multiple publication of reports of drug trials. *Eur J Clin Pharmacol.* 1989;36(5):429-432.
 58. Tramer MR, Reynolds DJ, Moore RA, McQuay HJ. Impact of covert duplicate publication on meta-analysis: a case study. *BMJ.* Sep 13 1997;315(7109):635-640.
 59. Egger M, Zellweger-Zahner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German.

- Lancet*. Aug 2 1997;350(9074):326-329.
60. Gotzsche PC. Reference bias in reports of drug trials. *Br Med J (Clin Res Ed)*. Sep 12 1987;295(6599):654-656.
 61. Kjaergard LL, Gluud C. Citation bias of hepato-biliary randomized clinical trials. *J Clin Epidemiol*. Apr 2002;55(4):407-410.
 62. Ravnskov U. Cholesterol lowering trials in coronary heart disease: frequency of citation and outcome. *BMJ*. Jul 4 1992;305(6844):15-19.
 63. Chan AW, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *BMJ*. Apr 2 2005;330(7494):753.
 64. Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA*. May 26 2004;291(20):2457-2465.
 65. Kirkham JJ, Dwan KM, Altman DG, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ*. 2010;340:c365.
 66. Franco A, Malhotra N, Simonovits G. Social science. Publication bias in the social sciences: unlocking the file drawer. *Science*. Sep 19 2014;345(6203):1502-1505.
 67. Dickersin K. The existence of publication bias and risk factors for its occurrence. *JAMA*. Mar 9 1990;263(10):1385-1389.
 68. Song F, Parekh-Bhurke S, Hooper L, et al. Extent of publication bias in different categories of research cohorts: a meta-analysis of empirical studies. *BMC Med Res Methodol*. 2009;9:79.
 69. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med*. Jan 17 2008;358(3):252-260.
 70. Yusuf S, Teo K, Woods K. Intravenous magnesium in acute myocardial infarction. An effective, safe, simple, and inexpensive intervention. *Circulation*. Jun 1993;87(6):2043-2046.
 71. ISIS-4: a randomised factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium sulphate in 58,050 patients with suspected acute myocardial infarction. ISIS-4 (Fourth International Study of Infarct Survival) Collaborative Group. *Lancet*. Mar 18 1995;345(8951):669-685.
 72. Dickersin K, Rennie D. Registering clinical trials. *JAMA*. Jul 23 2003;290(4):516-523.
 73. De Angelis C, Drazen JM, Frizelle FA, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *Lancet*. Sep 11-17 2004;364(9438):911-912.
 74. Sterne JA, Sutton AJ, Ioannidis JP, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*. 2011;343:d4002.
 75. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. Sep 13 1997;315(7109):629-634.
 76. Ioannidis JP, Trikalinos TA. The appropriateness of asymmetry tests for publication bias in meta-analyses: a large survey. *CMAJ*. Apr 10 2007;176(8):1091-1096.
 77. Harbour R, Miller J. A new system for grading recommendations in evidence based guidelines. *BMJ*. Aug 11 2001;323(7308):334-336.
 78. Mulrow CD. The medical review article: state of the science. *Ann Intern Med*. Mar 1987;106(3):485-488.
 79. Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analyses of randomized controlled trials. *N Engl J Med*. Feb 19 1987;316(8):450-455.
 80. Sacks HS, Reitman D, Pagano D, Kupelnick B. Meta-analysis: an update. *Mt Sinai J*

- Med.* May-Sep 1996;63(3-4):216-224.
81. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet.* Nov 27 1999;354(9193):1896-1900.
 82. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Ann Intern Med.* Aug 18 2009;151(4):W65-94.
 83. Delaney A, Bagshaw SM, Ferland A, Manns B, Laupland KB, Doig CJ. A systematic evaluation of the quality of meta-analyses in the critical care literature. *Crit Care.* Oct 5 2005;9(5):R575-582.
 84. Wen J, Ren Y, Wang L, et al. The reporting quality of meta-analyses improves: a random sampling study. *J Clin Epidemiol.* Aug 2008;61(8):770-775.
 85. Turner L, Shamseer L, Altman DG, Schulz KF, Moher D. Does use of the CONSORT Statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review. *Syst Rev.* 2012;1:60.
 86. Cobo E, Cortes J, Ribera JM, et al. Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: masked randomised trial. *BMJ.* 2011;343:d6783.
 87. Hirst A, Altman DG. Are peer reviewers encouraged to use reporting guidelines? A survey of 116 health research journals. *PLoS One.* 2012;7(4):e35621.
 88. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol.* 2007;7:10.
 89. Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol.* Oct 2009;62(10):1013-1020.
 90. Li T, Vedula SS, Scherer R, Dickersin K. What comparative effectiveness research is needed? A framework for using guidelines and systematic reviews to identify evidence gaps and research priorities. *Ann Intern Med.* Mar 6 2012;156(5):367-377.
 91. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 8. Rating the quality of evidence--indirectness. *J Clin Epidemiol.* Dec 2011;64(12):1303-1310.
 92. Schmid CH, Stark PC, Berlin JA, Landais P, Lau J. Meta-regression detected associations between heterogeneous treatment effects and study-level, but not patient-level, factors. *J Clin Epidemiol.* Jul 2004;57(7):683-697.
 93. Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? *Stat Med.* Jun 15 2002;21(11):1559-1573.
 94. Dechartres A, Altman DG, Trinquart L, Boutron I, Ravaud P. Association between analytic strategy and estimates of treatment outcomes in meta-analyses. *JAMA.* Aug 13 2014;312(6):623-630.
 95. Dechartres A, Trinquart L, Boutron I, Ravaud P. Influence of trial sample size on treatment effect estimates: meta-epidemiological study. *BMJ.* 2013;346:f2304.
 96. Lathyris DN, Patsopoulos NA, Salanti G, Ioannidis JP. Industry sponsorship and selection of comparators in randomized clinical trials. *Eur J Clin Invest.* Feb 2010;40(2):172-182.
 97. Goldberg NH, Schneeweiss S, Kowal MK, Gagne JJ. Availability of comparative efficacy data at the time of drug approval in the United States. *JAMA.* May 4 2011;305(17):1786-1789.
 98. Estellat C, Ravaud P. Lack of head-to-head trials and fair control arms: randomized controlled trials of biologic treatment for rheumatoid arthritis. *Arch Intern Med.* Feb

- 13 2012;172(3):237-244.
99. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med*. Oct 30 2004;23(20):3105-3124.
 100. Schöttker B, Luhmann D, Boukheimair D, Raspe H. Indirekte Vergleiche von Therapieverfahren. *GMS Health Technology Assessment*. 2009.
 101. Wells GA, Sultan SA, Chen L, Khan M, Coyle D. Indirect Evidence: Indirect Treatment Comparisons in Meta-Analysis. *Ottawa: Canadian Agency for Drugs and Technologies in Health*. 2009.
 102. Dias S, Welton NJ, Sutton AJ, Ades AE. Introduction to evidence synthesis for decision making. London: NICE DSU Technical Support Document 1; 2011.
 103. Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and Technical Challenges in Network Meta-analysis. *Ann Intern Med*. Jul 16 2013;159(2):130-137.
 104. Mills EJ, Ioannidis JP, Thorlund K, Schunemann HJ, Puhan MA, Guyatt GH. How to use an article reporting a multiple treatment comparison meta-analysis. *JAMA*. Sep 26 2012;308(12):1246-1253.
 105. Salanti G, Kavvoura FK, Ioannidis JP. Exploring the geometry of treatment networks. *Ann Intern Med*. Apr 1 2008;148(7):544-553.
 106. Ioannidis JP. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. *CMAJ*. Oct 13 2009;181(8):488-493.
 107. Rizos EC, Salanti G, Kontoyiannis DP, Ioannidis JP. Homophily and co-occurrence patterns shape randomized trials agendas: illustration in antifungal agents. *J Clin Epidemiol*. Aug 2011;64(8):830-842.
 108. Mills EJ, Ghement I, O'Regan C, Thorlund K. Estimating the power of indirect comparisons: a simulation study. *PLoS One* 2011;6(1):e16237.
 109. Thorlund K, Mills EJ. Sample size and power considerations in network meta-analysis. *Syst Rev*. 2012;1:41.
 110. Kim DD, Tang JY, Ioannidis JP. Network geometry shows evidence sequestration for medical vs. surgical practices: treatments for basal cell carcinoma. *J Clin Epidemiol*. Apr 2014;67(4):391-400.
 111. Mauri D, Polyzos NP, Salanti G, Pavlidis N, Ioannidis JP. Multiple-treatments meta-analysis of chemotherapy and targeted therapies in advanced breast cancer. *J Natl Cancer Inst*. Dec 17 2008;100(24):1780-1791.
 112. Georgiev H, Tsalatsanis A. *Social Network Analysis of Research Programs In Multiple Myeloma*. 53rd ASH Annual Meeting and Exposition, San Diego 2011.
 113. Felson DT, Anderson JJ, Meenan RF. The comparative efficacy and toxicity of second-line drugs in rheumatoid arthritis. Results of two metaanalyses. *Arthritis Rheum*. Oct 1990;33(10):1449-1461.
 114. O'Brien BJ, Anderson DR, Goeree R. Cost-effectiveness of enoxaparin versus warfarin prophylaxis against deep-vein thrombosis after total hip replacement. *CMAJ*. Apr 1 1994;150(7):1083-1090.
 115. Glenny AM, Altman DG, Song F, et al. Indirect comparisons of competing interventions. *Health Technol Assess*. Jul 2005;9(26):1-134, iii-iv.
 116. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol*. Jun 1997;50(6):683-691.
 117. Alkhafaji AA, Trinquart L, Baron G, Desvarieux M, Ravaud P. Impact of evergreening on patients and health insurance: a meta analysis and reimbursement cost analysis of citalopram/escitalopram antidepressants. *BMC Med*. 2012;10:142.
 118. Song F, Harvey I, Lilford R. Adjusted indirect comparison may be less biased than

- direct comparison for evaluating new pharmaceutical interventions. *J Clin Epidemiol*. May 2008;61(5):455-463.
119. Song F, Xiong T, Parekh-Bhurke S, et al. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study. *BMJ*. 2011;343:d4909.
 120. Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ*. Mar 1 2003;326(7387):472.
 121. Whitehead A. *Meta-analysis of Controlled Clinical Trials.*: Chichester, UK, John Wiley & Sons, Ltd.
 122. Higgins JP, Whitehead A. Borrowing strength from external trials in a meta-analysis. *Stat Med*. Dec 30 1996;15(24):2733-2749.
 123. Lumley T. Network meta-analysis for indirect treatment comparisons. *Stat Med*. Aug 30 2002;21(16):2313-2324.
 124. Salanti G, Higgins JP, Ades AE, Ioannidis JP. Evaluation of networks of randomized trials. *Stat Methods Med Res*. Jun 2008;17(3):279-301.
 125. Stijnen T, Hamza TH, Ozdemir P. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Stat Med*. Dec 20 2010;29(29):3046-3067.
 126. Jones B, Roger J, Lane PW, et al. Statistical approaches for conducting network meta-analysis in drug development. *Pharm Stat*. Nov-Dec 2011;10(6):523-531.
 127. Senn S, Gavini F, Magrez D, Scheen A. Issues in performing a network meta-analysis. *Stat Methods Med Res*. Apr 2013;22(2):169-189.
 128. Smith TC, Spiegelhalter DJ, Thomas A. Bayesian approaches to random-effects meta-analysis: a comparative study. *Stat Med*. Dec 30 1995;14(24):2685-2699.
 129. Chan AW, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. *Lancet*. Mar 26-Apr 1 2005;365(9465):1159-1162.
 130. Higgins J.P.T., Jackson D., Barrett J.K., Lu G., Ades A.E., White I. R. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Res Synth Methods*. 2012;3(2):98-110.
 131. Higgins JP, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *J R Stat Soc Ser A Stat Soc*. Jan 2009;172(1):137-159.
 132. Ades AE. ISPOR states its position on network meta-analysis. *Value Health*. Jun 2011;14(4):414-416.
 133. Lu G, Ades A. Modeling between-trial variance structure in mixed treatment comparisons. *Biostatistics*. Oct 2009;10(4):792-805.
 134. Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ*. 2009;338:b1147.
 135. Hoaglin DC, Hawkins N, Jansen JP, et al. Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 2. *Value Health*. Jun 2011;14(4):429-437.
 136. Jansen JP, Fleurence R, Devine B, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. *Value Health*. Jun 2011;14(4):417-428.
 137. Baker SG, Kramer BS. The transitive fallacy for randomized trials: if A bests B and B bests C in separate trials, is A better than C? *BMC Med Res Methodol*. Nov 13 2002;2:13.

138. Xiong T, Parekh-Bhurke S, Loke YK, et al. Overall similarity and consistency assessment scores are not sufficiently accurate for predicting discrepancy between direct and indirect comparison estimates. *J Clin Epidemiol*. Feb 2013;66(2):184-191.
139. Lu G, Ades AE. Assessing Evidence Inconsistency in Mixed Treatment Comparisons. *JASA*. 2006;101(474):447-459.
140. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med*. Mar 30 2010;29(7-8):932-944.
141. Dias S, Welton NJ, Sutton AJ, Ades AE. Inconsistency in Networks of Evidence Based on Randomised Controlled Trials. London: NICE DSU Technical Support Document 4; 2012.
142. Buti J, Glenny AM, Worthington HV, Nieri M, Baccini M. Network meta-analysis of randomised controlled trials: direct and indirect treatment comparisons. *Eur J Oral Implantol*. Spring 2011;4(1):55-62.
143. Cooper NJ, Sutton AJ, Morris D, Ades AE, Welton NJ. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: Application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. *Stat Med*. Jun 30 2009;28(14):1861-1881.
144. Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Res Synth Methods*. 2012;3(80-97).
145. van Valkenhoef G, Tervonen T, Zhao J, de Brock B, Hillege HL, Postmus D. Multicriteria benefit-risk assessment using network meta-analysis. *J Clin Epidemiol*. Apr 2012;65(4):394-403.
146. Madan J, Stevenson MD, Cooper KL, Ades AE, Whyte S, Akehurst R. Consistency between direct and indirect trial evidence: is direct evidence always more reliable? *Value Health*. Sep-Oct 2011;14(6):953-960.
147. Zhang WY, Li Wan Po A. Analgesic efficacy of paracetamol and its combination with codeine and caffeine in surgical pain--a meta-analysis. *J Clin Pharm Ther*. Aug 1996;21(4):261-282.
148. Donegan S, Williamson P, Gamble C, Tudur-Smith C. Indirect comparisons: a review of reporting and methodological quality. *PLoS One*. 2010;5(11):e11054.
149. Sobieraj DM, Cappelleri JC, Baker WL, Phung OJ, White CM, Coleman CI. Methods used to conduct and report Bayesian mixed treatment comparisons published in the medical literature: a systematic review. *BMJ Open*. 2013;3(7).
150. Nikolakopoulou A, Chaimani A, Veroniki AA, Vasiliadis HS, Schmid CH, Salanti G. Characteristics of networks of interventions: a description of a database of 186 published networks. *PLoS One*. 2014;9(1):e86754.
151. Coory M, Jordan S. Frequency of treatment-effect modification affecting indirect comparisons: a systematic review. *Pharmacoeconomics*. 2010;28(9):723-732.
152. Nixon RM, Bansback N, Brennan A. Using mixed treatment comparisons and meta-regression to perform indirect comparisons to estimate the efficacy of biologic treatments in rheumatoid arthritis. *Stat Med*. Mar 15 2007;26(6):1237-1254.
153. Welton NJ, Cooper NJ, Ades AE, Lu G, Sutton AJ. Mixed treatment comparison with multiple outcomes reported inconsistently across trials: evaluation of antivirals for treatment of influenza A and B. *Stat Med*. Nov 29 2008;27(27):5620-5639.
154. Salanti G, Marinho V, Higgins JP. A case study of multiple-treatments meta-analysis demonstrates that covariates should be considered. *J Clin Epidemiol*. Aug 2009;62(8):857-864.
155. Anothaisintawee T, Attia J, Nickel JC, et al. Management of chronic prostatitis/chronic pelvic pain syndrome: a systematic review and network meta-

- analysis. *JAMA*. Jan 5;305(1):78-86.
156. Tan SH, Bujkiewicz S, Sutton A, Dequen P, Cooper N. Presentational approaches used in the UK for reporting evidence synthesis using indirect and mixed treatment comparisons. *J Health Serv Res Policy*. Aug 14 2013.
 157. Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol*. Feb 2010;64(2):163-171.
 158. Li T, Puhan MA, Vedula SS, Singh S, Dickersin K. Network meta-analysis-highly attractive but more methodological research is needed. *BMC Med*. 2011;9:79.
 159. Montori V, Ioannidis J, Guyatt GH. Reporting bias. In: Guyatt GH, Rennie D, Meade MO, Cook DJ, eds. *User's guide to the medical literature: a manual for evidence-based clinical practice*. Vol Second: The McGraw-Hill Companies; 2008:543-554.
 160. Godlee F, Loder E. Missing clinical trial data: setting the record straight. *BMJ*. 2010;341:c5641.
 161. Song F, Parekh S, Hooper L, et al. Dissemination and publication of research findings: an updated review of related biases. *Health Technol Assess*. Feb 2010;14(8):iii, ix-xi, 1-193.
 162. Thorlund K, Druyts E, Avina-Zubieta JA, Wu P, Mills EJ. Why the findings of published multiple treatment comparison meta-analyses of biologic treatments for rheumatoid arthritis are different: an overview of recurrent methodological shortcomings. *Ann Rheum Dis*. Sep 1 2013;72(9):1524-1535.
 163. Mills EJ, Kanters S, Thorlund K, Chaimani A, Veroniki AA, Ioannidis JP. The effects of excluding treatments from network meta-analyses: survey. *BMJ*. 2013;347:f5195.
 164. Jansen JP, Trikalinos T, Cappelleri JC, et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. *Value Health*. Mar 2014;17(2):157-173.
 165. Puhan MA, Schunemann HJ, Murad MH, et al. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. *BMJ*. 2014;349:g5630.
 166. Lee AW. Review of mixed treatment comparisons in published systematic reviews shows marked increase since 2009. *J Clin Epidemiol*. Feb 2014;67(2):138-143.
 167. Edwards SJ, Clarke MJ, Wordsworth S, Borrill J. Indirect comparisons of treatments based on systematic reviews of randomised controlled trials. *Int J Clin Pract*. Jun 2009;63(6):841-854.
 168. Ades AE, Caldwell DM, Reken S, Welton NJ, Sutton AJ, Dias S. Evidence synthesis for decision making 7: a reviewer's checklist. *Med Decis Making*. Jul 2013;33(5):679-691.
 169. Trinquart L, Abbe A, Ravaud P. Impact of reporting bias in network meta-analysis of antidepressant placebo-controlled trials. *PLoS One*. 2012;7(4):e35219.
 170. Trinquart L, Chatellier G, Ravaud P. Adjustment for reporting bias in network meta-analysis of antidepressant trials. *BMC Med Res Methodol*. Sep 27 2012;12(1):150.
 171. Dias S, Welton NJ, Marinho V, Salanti G, Higgins JP, Ades AE. Estimation and adjustment of bias in randomized evidence by using mixed treatment comparison meta-analysis. *Journal of the Royal Statistical Society*. 2010;173(3):613-629.
 172. Vale CL, Tierney JF, Burdett S. Can trial quality be reliably assessed from published reports of cancer trials: evaluation of risk of bias assessments in systematic reviews. *BMJ*. 2013;346:f1798.
 173. Abdelhamid AS, Loke YK, Parekh-Bhurke S, et al. Use of indirect comparison methods in systematic reviews: a survey of Cochrane review authors. *Research*

- Synthesis Methods*. 2012(3):71-79.
174. Donegan S, Williamson P, D'Alessandro U, Tudur Smith C. Assessing key assumptions of network meta-analysis: a review of methods. *Research Synthesis Methods*. 2013(4):291-323.
 175. Mills EJ, Thorlund K, Ioannidis JP. Demystifying trial networks and network meta-analysis. *BMJ*. 2013;346:f2914.
 176. Bafeta A, Trinquart L, Seror R, Ravaud P. Analysis of the systematic reviews process in reports of network meta-analyses: methodological systematic review. *BMJ*. 2013;347:f3675.
 177. Hutton B, Salanti G, Chaimani A, et al. The quality of reporting methods and results in network meta-analyses: an overview of reviews and suggestions for improvement. *PLoS One*. 2014;9(3):e92508.
 178. Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *J Clin Epidemiol*. Nov 2000;53(11):1119-1129.
 179. Thornton A, Lee P. Publication bias in meta-analysis: its causes and consequences. *J Clin Epidemiol*. Feb 2000;53(2):207-216.
 180. Chaimani A, Vasiliadis HS, Pandis N, Schmid CH, Welton NJ, Salanti G. Effects of study precision and risk of bias in networks of interventions: a network meta-epidemiological study. *Int J Epidemiol*. Aug 2013;42(4):1120-1131.
 181. Savovic J, Jones HE, Altman DG, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. *Ann Intern Med*. Sep 18 2012;157(6):429-438.
 182. Tan SH, Cooper NJ, Bujkiewicz S, Welton NJ, Caldwell DM, Sutton AJ. Novel presentational approaches were developed for reporting network meta-analysis. *J Clin Epidemiol*. Jun 2014;67(6):672-680.
 183. Bafeta A, Trinquart L, Seror R, Ravaud P. Reporting of results from network meta-analyses: methodological systematic review. *BMJ*. 2014;348:g1741.

Annexe

Annexe 1 : Equation de recherche

		Cochrane Database of Systematic Reviews/ DARE	Medline	EMBASE
NMA	1	"network meta-analysis" OR "network meta-regression" OR "multiple treatment meta-analysis" OR "multiple treatments meta-analysis" OR "mixed treatment comparison" OR "mixed treatment comparisons"	"network meta-analysis" OR "network meta-regression" OR "multiple treatment meta-analysis" OR "multiple treatments meta-analysis" OR "mixed treatment comparison" OR "mixed treatment comparisons"	"network meta-analysis" OR "network meta-regression" OR "multiple treatment meta-analysis" OR "multiple treatments meta-analysis" OR "mixed treatment comparison" OR "mixed treatment comparisons"
	2	MTC AND meta-analysis	MTC AND meta-analysis	(mtc AND 'meta analysis'/exp)
	3	#1 or #2	#1 or #2	#1 or #2
	4	meta-analysis	meta-analysis[sb]	'meta analysis'/exp
	5	"systematic review"	systematic [sb]	'Systematic review '/exp
	6	#4or #5	#4or #5	#4or #5
	7	"mixed treatment" OR "multiple treatment" OR "multiple treatments" OR "treatment networks" OR "multiple comparison"	"mixed treatment" OR "multiple treatment" OR "multiple treatments" OR "treatment networks" OR "multiple comparison"	"mixed treatment" OR "multiple treatment" OR "multiple treatments" OR "treatment networks" OR "multiple comparison"
	8	#6 and #7	#6 and #7	#6 and #7
	9	8 or 3	8 or 3	8 or 3
OVERVIEW OF REVIEWS	10	"overview of reviews" OR "umbrella review" OR "overview of systematic reviews" OR "overview of meta-analyses" OR "multiple systematic reviews" OR "multiple meta-analyses" OR "overview of Cochrane reviews" OR "multiple Cochrane reviews" OR "overview of Cochrane"	(overview AND reviews) OR (umbrella AND review) OR (overview AND "systematic reviews") OR (overview AND meta-analyses) OR "multiple systematic reviews" OR "multiple meta-analyses" OR (overview AND "Cochrane reviews") OR (multiple AND Cochrane AND reviews) OR (overview AND Cochrane)	"overview of reviews" OR "umbrella review" OR "overview of systematic reviews" OR "overview of meta-analyses" OR "multiple systematic reviews" OR "multiple meta-analyses" OR "overview of Cochrane reviews" OR "multiple Cochrane reviews" OR "overview of Cochrane"
	11	"treatment networks" OR "network meta-analysis" OR "mixed treatment " OR "multiple treatments" OR "multiple treatment" OR "multiple comparisons" OR "multiple comparison" OR "indirect comparison"	"treatment networks" OR "network meta-analysis" OR "mixed treatment " OR "multiple treatments" OR "multiple treatment" OR "multiple comparisons" OR "multiple comparison" OR "indirect comparison"	"treatment networks" OR "network meta-analysis" OR "mixed treatment " OR "multiple treatments" OR "multiple treatment" OR "multiple comparisons" OR "multiple comparison" OR "indirect comparison"
	12	#11and #10	#11and #10	#11and #10
EQUATION OF SEARCH		12 or 9	12 or 9	12 or 9

Annexe 2 tableau : Caractéristiques générales des 121 méta-analyses en réseau

ITEMS	SUBCATEGORY	General journal, n=55	Specialized journal, n=66	Funding sources		
				Public †, n=55	Private ‡, n=48	Not reported, n=18
Journal type	General journal			32 (58)	18 (38)	5 (28)
	Specialized journal			23 (42)	30 (62)	13 (72)
Location of corresponding author	Europe	33 (60)	36 (55)	28 (51)	33 (69)	8 (44)
	North America	18 (32)	20 (30)	17 (31)	14 (29)	7 (39)
	Asia	1 (2)	7 (11)	5 (9)		3 (17)
	South America	1 (2)	2 (3)	3 (5)		
	Oceania	2 (4)	1 (1)	2 (4)	1 (2)	
Funding sources	Private	14 (26)	27 (41)			
	Public	25 (45)	15 (23)			
	None	7 (13)	8 (12)			
	Both	4 (7)	3 (5)			
	Unclear	5 (9)	13 (19)			
Medical area	Cardiology	13 (25)	14 (21)	16 (29)	6 (13)	5 (26)
	Rheumatology	4 (7)	12 (17)	5 (9)	11 (23)	
	Endocrinology	6 (11)	6 (9)	7 (13)	4 (8)	1 (6)
	Oncology	3 (5)	7 (11)	3 (5)	4 (8)	3 (16)
	Infectious disease	2 (4)	7 (11)	2 (4)	6 (13)	1 (6)
	Psychiatry/psychology	6 (11)	2 (3)	5 (9)	2 (4)	1 (6)
	Neurology	4 (7)	4 (6)	2 (4)	4 (8)	2 (11)
	Respiratology	3 (5)	3 (5)	1 (2)	3 (6)	2 (11)
	Ophthalmology	3 (5)	2 (3)	3 (5)	1 (2)	1 (6)
	Surgery	2 (4)	3 (5)	4 (7)	0 (0)	1 (6)
	Other (≤3 reviews/medical area. 11 medical areas)	9 (16)	6 (9)	7 (13)	7 (15)	1 (6)
Type of intervention assessed	Pharmacological intervention	44 (80)	56 (85)	37 (67)	46 (96)	17 (94)
	Different class ††	32 (58)	40 (61)	29 (53)	33 (69)	11 (61)
	Same class ††	15 (27)	20 (30)	12 (22)	15 (31)	8 (44)
	Any dose of same drug ††	3 (5)	9 (14)	5 (9)	5 (10)	2 (11)
	Non-pharmacological intervention	5 (9)	6 (9)	10 (18)	1 (2)	0 (0)
	Devices ††	6 (11)	4 (6)	10 (18)	0 (0)	0 (0)
	Surgery/procedure ††	4 (7)	3 (5)	7 (13)	0 (0)	0 (0)
	Therapeutic strategy/education	2 (4)	3 (5)	2 (4)	2 (4)	1 (6)
	Both	6 (11)	4 (6)	8 (15)	1 (2)	1 (6)

Typology of network	No. of interventions assessed (per network) ^{‡‡}	7 (5-10)	6 (5-9)	7 (5-10)	7 (6-10)	6 (5-9)
	No. of comparisons addressed by at least one RCT ^{††}	11 (6-15)	8 (5-14)	11 (5-16)	8 (6-15)	9 (6-13)
	NMA with at least one closed loop	48 (87)	55 (83)	48 (87)	39 (81)	16 (89)
	No. of RCTs included in network ^{‡‡}	31 (19-53)	20 (12-30)	31 (18-62)	20 (13-27)	24 (13-43)

Les données sont des nombres (%) sauf indiqué. RCT, randomized controlled trial, NMA network meta-analysis

†† Plusieurs réponses sont possible, le total n'est pas égal à 100%

‡‡ La donnée est une médiane (Q1-Q3)

† Les méta-analyses en réseau étaient classées comme ayant un financement public si elles avaient aucun financement ou un financement public

‡ Les méta-analyses en réseau étaient classées comme ayant un financement privé si elles avaient au moins une source de financement privée

Annexe 3 tableau : Evaluation des composantes essentielles des revues systématiques dans les 121 méta-analyses en réseau

Items	Subcategory (Yes if reported)	Funding sources		
		Public †, n=55	Private ‡, n=48	Not reported, n=18
INTRODUCTION				
Reporting of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design		49 (89)	43 (90)	18 (100)
METHODS				
Reporting of the existence of systematic review protocol		11 (20)	3 (6)	1 (6)
Reporting of primary outcome(s)		43 (78)	29 (60)	13 (72)
Reporting of information sources searched				
	Databases searched	54 (98)	47 (98)	17 (94)
	Electronic search strategy for each database	20 (36)	12 (25)	1 (6)
	Date of last search for each database	48 (87)	45 (94)	16 (89)
	Period covered by search for each database	40 (73)	32 (67)	11 (61)
	Search for any other sources (conference abstracts, unpublished studies, textbooks, specialized registers, contact with study authors, reviewing the references in the studies found or any relevant systematic reviews)	37 (67)	31 (65)	12 (67)
	Reviewing the references in the studies found or any relevant systematic reviews	24 (44)	23 (48)	7 (39)
	Search for ongoing studies	14 (25)	2 (4)	3 (17)
	Restriction or no restriction related to language	45 (82)	33 (69)	14 (78)
	Restriction or no restriction related to the publication status	40 (73)	31 (65)	11 (61)
Reporting of study selection and data collection process				
	Process for selecting studies	33 (60)	37 (77)	9 (50)
	Method of data extraction	44 (80)	31 (65)	14 (78)
Reporting of methods used for assessing risk of bias of individual studies		33 (60)	20 (42)	7 (39)
Reporting of methods to incorporate assessment of risk of bias of individual studies in the analysis or conclusions of review				
	Subgroup analysis	5 (9)	2 (4)	2 (11)
	Inclusion criteria	2 (4)	2 (4)	0 (0)
	GRADE	3 (5)	0 (0)	1 (6)
	Adjustment	1 (2)	1 (2)	0 (0)
	Any of these methods	11 (20)	5 (10)	3 (17)
Reporting of assessment of risk of bias that may affect the cumulative evidence (publication bias)		13 (24)	3 (6)	2 (11)
RESULTS				
Reporting of study selection				

	No. of studies screened, assessed for eligibility and included in the review	41 (75)	36 (75)	13 (72)
	List of studies included	51 (93)	48 (100)	18 (100)
	List of studies excluded	6 (11)	2 (4)	1 (6)
Reporting of study characteristics				
	Description of network	41 (75)	27 (56)	14 (78)
	Characteristics of patients (e.g., age, female/male ratio)	27 (49)	34 (71)	9 (50)
	Duration of follow-up of patients	27 (49)	24 (50)	9 (50)
	Duration of interventions	15 (27)	13 (27)	6 (33)
	Description of interventions	24 (44)	34 (71)	6 (33)
	No. of study groups	41 (75)	40 (83)	14 (78)
	No. of patients	40 (73)	42 (88)	14 (78)
	Funding source	11 (20)	1 (2)	1 (6)
Reporting of risk of bias within studies		28 (51)	17 (35)	6 (33)
Reporting of risk of bias across studies (publication bias)		13 (24)	3 (6)	2 (11)
DISCUSSION				
Reporting of limitations at review-level (reporting or publication bias)		31 (56)	20 (42)	7 (39)
OTHER ITEMS				
Reporting of assumptions required in NMA				
	Homogeneity assumption	50 (91)	41 (85)	16 (89)
	Similarity assumption	15 (27)	23 (48)	3 (17)
	Consistency or exchangeability assumption	32 (58)	20 (42)	6 (33)
Reporting of conflict of interest		51 (93)	33 (69)	11 (61)

Les données sont les nombres de publication avec l'item correspondant (%). NMA network meta-analysis

† Les méta-analyses en réseau étaient classées comme ayant un financement public si elles avaient aucun financement ou un financement public

‡ Les méta-analyses en réseau étaient classées comme ayant un financement privé si elles avaient au moins une source de financement privée

Annexe 4 tableau: Evaluation des méthodes inadéquates dans les publications ayant une présentation adéquate

Items (Reported and inadequate)	Public sources †	Private sources ‡	Not reported
Electronic search of only one bibliographic database	6/54 (11)	5/47 (11)	2/17 (12)
Restriction of study selection based on publication status	7/40 (18)	7/31 (23)	2/11 (18)
Lack of independent duplicate study selection	12/33 (36)	18/37 (48)	4/9 (44)
Lack of independent duplicate data extraction	4/44 (9)	9/31 (29)	6/14 (43)

† Les méta-analyses en réseau étaient classées comme ayant un financement public si elles avaient aucun financement ou un financement public

‡ Les méta-analyses en réseau étaient classées comme ayant un financement privé si elles avaient au moins une source de financement privée

Annexe 5 tableau: Présentation des résultats dans les 121 méta-analyses en réseau

ITEMS, yes if reported

	Funding Public, n=55	Funding Private, n=48	Not reported, n=18
Amount of evidence in the network of RCTs			
Reporting of interventions included in the network	55 (100)	48 (100)	18 (100)
Reporting of direct comparisons between the interventions	41 (75)	27 (56)	14 (78)
Reporting of No RCTs for each comparison	33 (60)	25 (52)	11 (61)
Reporting of No patients for each comparison	15 (27)	6 (13)	4 (22)
Reporting of network graph.	37 (71)	27 (54)	14 (78)
Effect size estimates for pair-wise comparisons between interventions			
Reporting of results of all individual RCTs used in the NMA	31 (56)	26 (54)	11 (61)
Reporting of direct evidence	38 (69)	23 (48)	12 (67)
Reporting of indirect evidence	6 (11)	5 (10)	2 (11)
Reporting of mixed evidence	41 (75)	23 (48)	14 (78)
Reporting of effect estimates for some selected pair-wise comparisons between interventions from NMA	14 (25)	25 (52)	4 (22)
Format of results ††			
Table	30 (55)	33 (69)	15 (83)
Forest plot	18 (33)	20 (42)	3 (17)
Matrix	6 (11)	0 (0)	1 (6)
Figure	3 (5)	2 (4)	1 (6)
Text	3 (5)	1 (2)	0 (0)
Rank order of interventions			
Reporting of rank order of intervention	28 (51)	16 (33)	8 (44)
Reporting of rank order of intervention with incertitude	5/28 (18)	2/16 (13)	2/8 (25)

Les données sont les nombres de publication avec l'item correspondant (%).RCT, randomized controlled trial, NMA network meta-analysis

† Les méta-analyses en réseau étaient classées comme ayant un financement public si elles avaient aucun financement ou un financement public

‡ Les méta-analyses en réseau étaient classées comme ayant un financement privé si elles avaient au moins une source de financement privée

†† Plusieurs réponses sont possible, le total n'est pas égal à 100%