

8. Évaluer le degré de confiance que mérite une étude méthodique

Simon Lewin^{1}, Andrew D Oxman², John N Lavis³ et Atle Fretheim⁴*

Lewin S, Oxman AD, Lavis JN, Fretheim A: **SUPPORT Tools for evidence-informed health Policymaking (STP). 8. Deciding how much confidence to place in a systematic review.** Health Research Policy and Systems; 2009, **7(Suppl 1):S8** doi:10.1186/1478-4505-7-S1-S8.

<http://www.health-policy-systems.com/content/pdf/1478-4505-7-S1-s8.pdf>

1. Norwegian Knowledge Centre for the Health Services, P.O. Box 7004, St. Olavs plass, N-0130 Oslo, Norway;
Health Systems Research Unit, Medical Research Council of South Africa
2. Norwegian Knowledge Centre for the Health Services, P.O. Box 7004, St. Olavs plass, N-0130 Oslo, Norway
3. Centre for Health Economics and Policy Analysis, Department of Clinical Epidemiology and Biostatistics, and
Department of Political Science, McMaster University, 1200 Main St. West, HSC-2D3, Hamilton, ON,
Canada, L8N 3Z5
4. Norwegian Knowledge Centre for the Health Services, P.O. Box 7004, St. Olavs plass, N-0130 Oslo, Norway;
Section for International Health, Institute of General Practice and Community Medicine, Faculty of Medicine,
University of Oslo, Norway

* Personne-ressource principale (simon.lewin@nokc.no)

Il s'agit de la traduction d'un article publié dans le Health Research Policy and Systems, 2009; 7:Supplement 1 (www.health-policy-systems.com/supplements/7/S1). La reproduction, la distribution et l'utilisation de ces articles, par quelque moyen que ce soit, sont permises à condition d'en citer la source. Le site Web SUPPORT (www.support-collaboration.org) renferme les hyperliens vers les versions chinoise, française, portugaise et espagnole. Vous pouvez envoyer vos commentaires sur la façon d'améliorer les outils présentés dans cette série de documents, par courriel, à : STP@nokc.no.

Cette série d'articles a été préparée dans le cadre du projet SUPPORT, parrainé par le programme INCO du sixième programme cadre de la Commission européenne, numéro de contrat 031939. L'Agence norvégienne de coopération pour le développement (NORAD), l'Alliance pour la recherche sur les politiques et les systèmes de santé et le Milbank Memorial Fund ont financé une réunion visant l'examen par des pairs d'une version initiale de la série. John Lavis a touché un salaire de la Chaire de recherche du Canada sur le transfert et l'échange des connaissances. NORAD, la composante norvégienne du groupe Cochrane Effective Practice and Organisation of Care (EPOC), le Centre norvégien de connaissances pour les services de santé de santé, l'AHPSR, la Fondation canadienne de la recherche sur les services de santé (FCRSS), l'Evidence-Informed Policy Network (EVIPNet) et l'Organisation panaméricaine de la santé (OPS) ont participé à la traduction et à la diffusion des articles. Aucun des bailleurs de fonds n'a influé sur la rédaction, la révision ou l'approbation du contenu de la série.

La Fondation canadienne de la recherche sur les services de santé a traduit cet article en français avec l'appui du Centro Rosarino de Estudios Perinatales (CREP) et de l'Organisation panaméricaine de la santé (OPS).



Résumé

Le présent article fait partie d'une série de documents à l'intention des personnes chargées de prendre des décisions relativement aux politiques et aux programmes de santé ainsi que de celles qui les assistent.

Les études méthodiques sur les effets des interventions en santé ne sont pas toutes aussi fiables. En conséquence, les responsables de politiques, entre autres, doivent être en mesure de déterminer le degré de confiance que l'on peut leur accorder. Le recours à des processus systématiques et transparents pour prendre des décisions de cet ordre peut aider à éviter que ces jugements soient entachés d'erreurs ou de partialité. Dans le présent article, nous proposons cinq questions qu'il faut se poser pour déterminer le degré de fiabilité des conclusions d'une étude méthodique sur les effets d'une intervention : 1. L'étude traite-t-elle explicitement d'une question appropriée de politique ou de gestion? 2. A-t-on utilisé des critères appropriés pour sélectionner les études à examiner? 3. La recherche d'études pertinentes a-t-elle été approfondie et suffisamment exhaustive? 4. Les évaluations de la pertinence des études par rapport au sujet traité et du risque de partialité sont-elles reproductibles? 5. Les résultats sont-ils semblables d'une étude à l'autre?

À PROPOS DES OUTILS DU PROJET SUPPORT À PROPOS DES OUTILS DU PROJET SUPPORT

Le présent article fait partie d'une série de documents destinés aux personnes chargées de prendre des décisions relativement aux politiques et aux programmes de santé et à celles qui les assistent, dans le but de les aider à s'assurer que leurs décisions sont éclairées par les meilleures données de recherche disponibles. L'introduction décrit plus en profondeur les outils du projet SUPPORT et les manières dont ils peuvent être utilisés [1]. Le glossaire de la série est joint à chaque article (voir le fichier complémentaire 1). Les résumés des études méthodiques préparés dans le cadre du projet SUPPORT peuvent être consultés à www.support-collaboration.org (en anglais). Des résumés en français, espagnol, portugais et chinois seront affichés sur ce site au cours de 2010 (www.support-collaboration.org/supporttool.htm). Toute rétroaction visant l'amélioration des outils abordés dans la série est la bienvenue et doit être envoyée à STP@nokc.no.

SCÉNARIOS

Scénario 1 : Vous êtes un haut fonctionnaire qui doit soumettre au ministre une proposition sur les données probantes appuyant diverses propositions de politiques et de programmes portant sur un enjeu de santé prioritaire. Vous vous demandez dans quelle mesure vous pouvez faire confiance aux études méthodiques des données probantes relatives à chacune des propositions et désirez vous assurer qu'elles sont été convenablement évaluées par votre personnel.

Scénario 2 : Vous travaillez au ministère de la Santé où vous préparez un document sur diverses propositions possibles en vue de s'attaquer à un enjeu de santé prioritaire. On a trouvé un certain nombre d'études méthodiques sur les effets des propositions envisagées et on vous a demandé de déterminer le degré de confiance à accorder à chacune de ces études.

Scénario 3 : Vous travaillez au sein d'une unité autonome qui aide le ministère de la Santé à mettre à profit les données probantes aux fins de l'élaboration de politiques. Vous préparez à l'intention du Ministère un document sur les incidences probables des propositions visant à résoudre un enjeu de santé prioritaire. Vous avez besoin d'orientation pour déterminer le degré de fiabilité des études méthodiques sur l'impact de chacune des propositions.

CONTEXTE

Aux responsables de politiques (scénario 1), le présent article propose diverses questions à soumettre à leur personnel en vue de déterminer le degré de confiance à

accorder aux conclusions d'une étude méthodique sur les effets des interventions en santé. Aux personnes qui les assistent (scénarios 2 et 3), il suggère des questions susceptibles d'orienter une évaluation critique des études méthodiques de ces effets.

Il est généralement admis que les études méthodiques d'essais cliniques à répartition aléatoire renferment les données probantes les plus fiables sur les effets des interventions en santé [2,3]. Les études méthodiques se caractérisent par le recours à des méthodes systématiques et explicites afin de repérer, de sélectionner et d'évaluer les travaux de recherche pertinents, et de recueillir et d'analyser les données issues des études retenues [3]. On les utilise de plus en plus pour trouver, évaluer et colliger les données probantes sur les conséquences économiques des interventions [4], comme le rapport coût-efficacité de la promotion de l'allaitement maternel dans les services de néonatalité [5] ou les coûts de diverses stratégies de diffusion et de mise en œuvre de lignes directrices [6]. Elles servent également à résumer les données probantes issues d'études qualitatives portant, par exemple, sur les opinions des consommateurs ou des prestataires de services de santé au sujet des interventions en santé [7-10]. Le présent article s'intéresse en particulier aux études méthodiques des effets des politiques ou des programmes de santé, qu'il s'agisse des mécanismes de prestation de soins (le remplacement des médecins par des infirmières dans les services de première ligne, par exemple [11]) ou des stratégies de changement telles que les réunions de formation continue à l'intention des professionnels de la santé [12].

Les méthodes systématiques et explicites employées dans les études méthodiques visent à minimiser le risque d'erreurs ou de partialité et à faciliter l'évaluation critique de ces synthèses [13,14]. Toutefois, comme elles ne sont pas toutes menées avec la même rigueur – autrement dit, les conclusions de ces études ne sont pas toutes également dignes de confiance. Le fait de définir une évaluation comme étant une « étude méthodique » (ou une méta-analyse) ne garantit donc pas à lui seul la fiabilité des conclusions devant servir à éclairer la prise de décisions en matière de politiques.

Ainsi, lorsqu'ils utilisent les études méthodiques sur les effets des propositions pour éclairer leurs décisions, les responsables de politiques et les autres décideurs doivent être en mesure de juger de la confiance qu'ils peuvent accorder à ces données probantes. Le recours à un processus systématique et transparent peut aider à éviter que leurs jugements ne soient entachés d'erreurs ou de partialité. Un processus de cette nature permet également à d'autres parties intéressées, notamment au grand public, de comprendre et d'apprécier ces jugements. Ce facteur est particulièrement important lorsque de telles évaluations influent sur des recommandations ou des décisions concernant des interventions ou des services cliniques [15], ou encore sur des décisions touchant la mise en œuvre ou l'abandon de programmes ou de politiques. La figure 1 présente la marche à suivre pour trouver et évaluer des études méthodiques qui serviront à éclairer l'élaboration de politiques.

La confiance accordée aux conclusions d'une étude méthodique peut être tempérée pour diverses raisons, dont le défaut de :

- Préciser la question et la méthodologie avant d'entreprendre l'examen, par exemple dans un protocole d'étude publié
- Définir des critères explicites pour l'inclusion ou l'exclusion des études
- Donner une description adéquate des études retenues
- Déterminer le risque de partialité dans les études retenues
- Mesurer le risque de biais de publication, c.-à-d. Le fait que certaines études, généralement celles qui font état de résultats positifs (« statistiquement significatifs »), sont plus susceptibles d'être publiées et donc d'être incluses dans une revue
- Faire appel à des méthodes appropriées pour combiner les résultats des études retenues (dans une méta-analyse), le cas échéant
- Prendre dûment en considération les écarts dans les conclusions des études retenues (c.-à-d. L'« hétérogénéité » des conclusions)
- Fonder les conclusions de l'examen sur les données présentées dans les études méthodiques

Les limites éventuelles des études méthodiques tiennent également aux conflits d'intérêts (qui peuvent influencer sur la fiabilité d'une étude de n'importe laquelle des manières énumérées ci-dessus) ainsi qu'à leur obsolescence.

Par exemple, des variations de la fiabilité ont été observées dans une étude comparant la méthodologie et les éléments de compte rendu des revues de la Collaboration Cochrane avec ceux des études méthodiques publiées dans des revues scientifiques sur support papier. Cette étude a révélé que les revues Cochrane comportaient des éléments qui, en général, les exposaient moins aux biais de publication. Cet état de fait est attribué principalement à la description sans équivoque des critères d'inclusion et d'exclusion ainsi qu'à l'évaluation formelle du risque de partialité de chacune des études méthodiques retenues [16]. Une autre étude comparative de la qualité de la méthodologie et des conclusions des revues Cochrane relatives aux essais cliniques sur les médicaments par rapport à celle des études sur les mêmes médicaments financées par l'industrie attribue une note plus élevée aux revues Cochrane en matière d'évaluation de la qualité parce qu'elles tiennent davantage compte du risque de partialité que les études parrainées par l'industrie. En outre, celles-ci sont significativement plus susceptibles de recommander les médicaments en cause sans aucune réserve [17]. Diverses autres études de revues font aussi état de divergences dans la qualité et les conclusions de ces dernières [18-21].

Différents outils ont été conçus pour évaluer la qualité des études méthodiques, notamment AMSTAR (A Measurement Tool to Assess Reviews [22]), CASP (Critical Appraisal Skills Programme [23]) ainsi que la grille mise au point par Oxman et Guyatt [24] (voir aussi CCOHTA [25]; West et coll. [26]), mais tous contiennent des critères semblables. (AMSTAR est l'objet du tableau 1.) Plusieurs outils comprennent

désormais des échelles de notation visant à évaluer le degré de confiance que l'on peut accorder à une étude méthodique. Cette tendance est à la hausse. De façon générale, une note élevée laisse entendre que l'on peut accorder une plus grande confiance aux conclusions d'une étude méthodique. À l'inverse, des notes faibles indiquent que les conclusions sont moins fiables et qu'il faut les examiner de plus près pour en déterminer les principales lacunes. Il convient toutefois de souligner les points suivants : d'abord, une note globale n'indique pas nécessairement quels aspects précis de l'étude méthodique ont été traités d'une façon fiable – certains peuvent avoir reçu un traitement plus rigoureux que d'autres. Ensuite, le processus de notation en soi a pour corollaire obligé l'attribution d'une pondération aux différents éléments de l'outil d'évaluation, et les écarts de pondération entre les éléments sont parfois difficiles à justifier [27]. Enfin, les outils de notation ne peuvent établir la fiabilité que des données connues. Lorsqu'une étude méthodique ne contient pas les renseignements essentiels sur les méthodes utilisées, il n'est pas aisé de savoir ce qui a été fait ou dans quelle mesure ce qui *a été fait* constitue une restriction importante.

Il importe de bien établir la distinction entre l'*évaluation* du degré de confiance que l'on peut accorder aux conclusions d'une étude méthodique et la *compréhension* des résultats de l'étude même. Le tableau 2 donne des indications quant à ce qu'il faut rechercher dans les résultats d'une étude des effets et le tableau 3 indique comment déterminer la fiabilité des conclusions des examens d'études qualitatives et d'études économiques.

Il importe également de faire la différence entre l'évaluation du degré de confiance que l'on peut accorder aux conclusions d'une étude méthodique et toute évaluation de la *pertinence* de l'étude par rapport à une question stratégique donnée. À cet égard, il faut déterminer, par exemple, la présence de données probantes sur les effets des diverses propositions de politiques ou de programmes envisagées et l'applicabilité des conclusions de l'étude dans le contexte de mise en œuvre (ce dernier élément est l'objet de l'article 9 de la présente série [28]).

QUESTIONS À ENVISAGER

Les questions suivantes peuvent orienter les décisions sur la fiabilité conclusions d'une étude méthodique portant sur les effets d'une proposition :

1. L'étude traite-t-elle explicitement d'une question appropriée de politique ou de gestion?
2. A-t-on utilisé des critères appropriés pour sélectionner les études à examiner?
3. La recherche d'études pertinentes a-t-elle été approfondie et suffisamment exhaustive?
4. Les évaluations de la pertinence des études par rapport au sujet traité et du risque de partialité sont-elles reproductibles?
5. Les résultats sont-ils semblables d'une étude à l'autre?

1. L'étude traite-t-elle explicitement d'une question appropriée de politique ou de gestion?

La première étape cruciale pour évaluer la confiance à accorder aux conclusions d'une étude méthodique consiste à examiner la *question* envisagée. La méthodologie et le déroulement d'un examen peuvent très bien être sans failles, mais ses conclusions ont peu de chances d'être utiles à la prise de décisions si elles ne traitent pas explicitement d'une question de politique ou de gestion qui est sensée, appropriée et pertinente à la question envisagée par le responsable de politiques.

Une question de politique ou de gestion appropriée doit :

- *Être explicite* : autrement dit, elle sera exposée par le menu plutôt que sous-entendue dans le texte. Si la question n'a pas été explicitement ou clairement formulée, il est difficile d'évaluer adéquatement le déroulement de l'étude, ce processus devant être, du moins en partie, mis en rapport avec la question elle-même [29]. Ainsi, l'évaluation de l'opportunité des critères de sélection des études méthodiques doit prendre en considération la question à laquelle celles-ci étaient censées répondre. De plus, une question claire aide les lecteurs à déterminer si l'étude méthodique est pertinente à leur travail [29]
- *Être établie a priori* : c'est-à-dire *avant* la réalisation de l'étude méthodique, préférablement dans un protocole ou un plan d'étude (ainsi, toutes les revues cochrane sont précédées d'un protocole d'étude publié, comme en font foi les exemples que l'on trouve dans la bibliothèque cochrane; voir www3.interscience.wiley.com/cgi-bin/mrwhome/106568753/home). Autrement, elle pourrait avoir été modifiée pour correspondre aux données probantes recensées, ce qui minerait la fiabilité des conclusions
- *Traiter d'une question en lien avec l'élaboration de politiques ou les mécanismes de gestion* : la pertinence de la question à cet égard devra être évaluée dans un contexte précis, en fonction de l'importance des problèmes à résoudre dans une province ou un pays donné et à un moment particulier. La question envisagée dans l'étude méthodique pourrait ne pas être appropriée si :
 - *Elle est trop restrictive* : si, par exemple, une étude méthodique traite des effets d'un programme sur les seuls participants d'un groupe d'âge donné dans un milieu particulier ou en fonction d'une gamme restreinte de résultats, la généralisation de ses conclusions à d'autres populations, milieux ou résultats serait impossible
 - *Elle est trop large* : dans le cas, par exemple, où une étude méthodique établit qu'un programme englobe une grande variété de pratiques, lesquelles ne sont pas toutes appropriées dans une province ou un pays donné, ou encore qu'elle soulève une question très générale qui ne peut servir à éclairer la prise de décisions. Ainsi, la question de savoir si le personnel infirmier peut effectivement assurer la prestation des programmes de promotion de la santé ne sera d'aucune utilité pour décider si une catégorie particulière d'infirmières ou

d'infirmiers (p. ex., le personnel infirmier ayant deux ans de formation, appelé *enrolled nurse* au Royaume-Uni) peut veiller à la prestation d'un tel programme pour un problème de santé précis comme la prévention du VIH/sida

- *Elle ne définit pas de groupe témoin approprié* : par exemple, lorsqu'on compare un programme à un scénario « sans programme » plutôt qu'au meilleur traitement possible pour une affection donnée

Une question bien formulée devrait contenir *tous* les éléments suivants : les types de population et les milieux visés par l'étude méthodique (p. ex., les enfants de un mois à six ans vivant dans une région où le paludisme est endémique); les types de programmes et de comparaisons pris en considération (p. ex., l'administration de médicaments antipaludiques à intervalles réguliers [l'intervention] contre un placebo ou l'absence de médicament [la comparaison]); les types de résultats (p. ex., paludisme clinique et anémie grave) [30,31]. On utilise parfois l'acronyme anglais PICO (Population, Intervention, Comparaison, Outcome [résultats]) pour résumer ces quatre éléments clés.

Même si la nécessité de bien formuler la question peut sembler aller de soi, nombre d'études narratives ne satisfont pas à cette exigence. L'examen d'un échantillon d'études de ce genre publiées dans les principales revues médicales a révélé que l'objet de l'étude n'était pas clairement défini dans 20 p. 100 des cas [32].

2. A-t-on utilisé des critères appropriés pour sélectionner les études à examiner?

Les critères d'inclusion et d'exclusion à une étude méthodique correspondent aux listes détaillées des types de populations, des interventions, des comparaisons et des résultats qu'on se propose d'examiner. Comme ces critères, précisés dans un protocole d'étude, déterminent les études méthodiques à retenir et qu'ils influent grandement sur les résultats de l'examen, ils doivent nécessairement cadrer avec la question à envisager.

La détermination de l'opportunité de ces critères reposera donc sur les questions suivantes :

- Les critères d'inclusion et d'exclusion sont-ils clairement définis? C'est là un moyen de prévenir toute partialité dans le choix des études à retenir. Ainsi, selon une récente évaluation de la qualité méthodologique d'études méthodiques en chirurgie générale, seulement 70 p. 100 de celles-ci faisaient état des critères d'admissibilité utilisés [18]
- Les critères d'inclusion et d'exclusion sont-ils explicites en ce qui a trait aux types de populations, d'interventions et de comparaisons, de résultats visés?
- Les critères d'inclusion et d'exclusion cadrent-ils avec la question à envisager [33]? Par exemple, lorsqu'une étude méthodique a pour but d'évaluer la prophylaxie et le traitement intermittent par des médicaments antipaludiques pour prévenir le

paludisme chez les jeunes enfants de régions où cette maladie est endémique, les critères indiquent-ils que sont aussi incluses des études d'enfants des milieux en cause et précisent-ils les formes de prophylaxie et de traitement qui seront envisagés [31]? Dans le même ordre d'idées, si l'étude méthodique porte sur les effets d'interventions en vue d'augmenter la proportion de professionnels de la santé œuvrant en milieu rural et dans les autres régions mal desservies, les critères font-ils état des catégories de professionnels de la santé visés et des types d'activités de formation ou d'ordre financier qui seront offertes [34]?

3. La recherche d'études pertinentes a-t-elle été approfondie et suffisamment exhaustive?

Une facette importante de l'étude méthodique consiste à effectuer une recherche documentaire approfondie et reproductible afin de repérer des études qui satisfont aux critères d'admissibilité. C'est, entre autres, ce qui distingue les études *méthodiques* des études *narratives*. La recherche systématique contribue à minimiser la partialité de l'étude méthodique, car elle prend en considération toutes les données probantes pertinentes. Il est ainsi possible d'établir des estimations fiables des effets de la politique ou du programme envisagé [35].

La partialité des études méthodiques peut être imputable, notamment, au biais de publication, c'est-à-dire la publication d'études en fonction de l'orientation et du poids de leurs résultats [36]. Selon une récente étude visant à déterminer l'influence, sur la publication d'essais cliniques à répartition aléatoire, de la présence ou de l'absence de résultats positifs et de l'importance accordée aux conclusions des essais cliniques, ceux, parmi ces derniers, qui ont obtenu des résultats positifs sont beaucoup plus susceptibles d'être publiés que les essais cliniques présentant des résultats négatifs [37]. Des travaux de recherche montrent, à l'instar cette étude, que les essais cliniques affichant des résultats positifs sont publiés plus tôt que d'autres [38]. Aussi importe-t-il, dans le cadre d'une étude méthodique, de réunir toutes les études, tant publiées qu'inédites, sous peine de surestimer les effets positifs des programmes.

Les études méthodiques ne font pas toutes l'objet d'une recherche exhaustive. Par exemple, un examen portant sur la diffusion des études méthodiques publiées sur le traitement de l'asthme révèle que seulement 52 p. 100 des 33 études retenues ont fait l'objet d'une recherche suffisamment exhaustive des données probantes relatives aux effets [20]. Il est donc essentiel de vérifier comment a été mené le repérage des études pertinentes.

Les questions suivantes devraient aider à déterminer si la recherche d'études pertinentes a été détaillée et raisonnablement exhaustive [22] :

- L'étude méthodique comporte-t-elle une description détaillée de la stratégie utilisée pour repérer les études pertinentes? Devraient être précisés : 1. la liste des sources consultées, 2. les mots-clés utilisés (s'il y a lieu) ; 3. les années sur lesquelles le

repérage de ces sources a porté. Le tableau 4 présente des exemples des sources consultées dans le cadre des études méthodiques publiées dans la Bibliothèque Cochrane

- La stratégie de recherche englobait-elle les bases de données électroniques d'études publiées? Il existe une grande variété de bases de données qui publient des études et nombre d'entre elles peuvent être consultées gratuitement ou moyennant des frais minimes. Voici les principales : PubMed/MEDLINE, constituée par la National Library of Medicine des États-Unis; le registre central d'essais cliniques de la Collaboration Cochrane (Central Register of Controlled Trials – CENTRAL); et les bases de données régionales telles que LILACS (Literatura Latinoamericana y del Caribe en Ciencias de la Salud). Les articles 4 [39] et 5 [40] de la présente série contiennent de plus amples renseignements sur la façon de trouver des comptes rendus de recherche pertinents
- La recherche documentaire a-t-elle débordé le cadre des bases de données électroniques? Ces recherches complémentaires, qu'il s'agisse de l'exploration de la liste des lectures de référence présentées dans les études pertinentes, de communications avec des auteurs et des experts du domaine ou de la consultation des registres d'études spécialisés qui traitent du sujet de l'étude méthodique, sont utiles pour repérer des études tant publiées qu'inédites (y compris les études répertoriées dans la littérature grise, c.-à-d. issues de sources autres que les revues indexées, examinées par des pairs)
- Les recherches sont-elles à jour? L'étude méthodique précise-t-elle la période visée par les recherches et ces dernières sont-elles récentes? Une étude méthodique publiée, même si elle porte sur une question de politique, pourrait reposer sur des recherches datant de plusieurs années. En ce cas, elle ne comprendra pas toutes les données probantes les plus récentes et il ne sera guère possible d'en dégager une estimation fiable des effets de la proposition de politique ou de programme

4. Les évaluations de la pertinence des études retenues au regard de la question envisagée et du risque de partialité sont-elles reproductibles?

Les auteurs d'études méthodiques doivent porter deux jugements cruciaux pour chaque étude primaire retenue. D'abord, l'étude répond-elle aux critères d'inclusion – autrement dit, est-elle pertinente à l'objet de l'étude méthodique? Ensuite, dans quelle mesure les résultats de l'étude présentent-ils un risque de partialité? Il s'agit, en l'occurrence, du risque d'erreur systématique ou de déviation de la vérité dans les résultats ou les inférences [27] qui peut avoir pour effet de mettre en doute l'exactitude des résultats d'une étude [27]. Comme ces jugements influenceront sur les conclusions de l'examen, il importe de les présenter de manière transparente et reproductible. D'autres parties intéressées doivent être en mesure de comprendre les fondements de ces jugements et d'appliquer ces évaluations dans d'autres contextes.

Comme nous l'avons vu précédemment, la description explicite des critères d'inclusion et d'exclusion est essentielle pour éviter toute partialité dans le processus de sélection

des études méthodiques. Ces critères et jugements influent forcément sur les conclusions de l'examen, puisqu'ils servent à déterminer le choix des études retenues. Mais il est possible de minimiser le risque de partialité ou les erreurs de jugement. D'abord en confiant à deux examinateurs le soin de décider de manière indépendante de l'inclusion des études. D'autres examinateurs peuvent aussi être mis à contribution pour résoudre les divergences d'opinion à cet égard. Ensuite, les motifs de l'inclusion d'une étude (et de l'exclusion d'une étude apparemment pertinente) devraient figurer dans la revue publiée. Les lecteurs pourront ainsi poser leur propre jugement au sujet des décisions prises. De plus, on associe de la sorte une « liste de contrôle » transparente à l'étude méthodique, « historique d'expertise », assurant du coup la reproductibilité du processus.

La validité des données obtenues de chacune des études retenues est un autre facteur permettant de tirer des conclusions d'une étude méthodique quant aux effets d'une politique ou d'un programme. Le rassemblement des résultats des études ou leur synthèse peut fausser le constat si la validité des données n'est pas éprouvée. Voilà pourquoi l'évaluation du risque de partialité des résultats est une étape aussi cruciale de l'étude méthodique, sans compter qu'elle servira également à éclairer les conclusions et leur interprétation [27].

Diverses méthodes ont été élaborées pour évaluer la qualité des essais cliniques à répartition aléatoire ou le risque de partialité qu'ils présentent [27,41,42]. Même si ces approches ne font pas l'objet du présent document, rappelons l'importance d'en donner une description explicite et de veiller à l'homogénéité de leur application.

Les questions suivantes peuvent orienter l'évaluation de la pertinence des études retenues au regard de la question envisagée et du risque de partialité :

- L'évaluation de la pertinence des études retenues au regard de la question envisagée repose-t-elle sur des critères explicites et transparents? Cet élément devrait être précisé dans l'étude méthodique, laquelle doit aussi présenter une liste des études retenues et de celles qui ont été exclues
- L'évaluation de la pertinence des études retenues au regard du risque de partialité repose-t-elle sur des critères explicites et transparents? Cet élément devrait être précisé dans l'étude méthodique, laquelle doit aussi décrire le processus d'évaluation de même que ses résultats
- L'interprétation des résultats de l'étude méthodique tient-elle compte de l'analyse de l'évaluation du risque de partialité? Cet élément est d'une importance cruciale, car la possibilité d'un risque de partialité élevé dans études retenues pourrait réduire la fiabilité des conclusions

5. Les résultats sont-ils semblables d'une étude à l'autre?

Les conclusions des études retenues peuvent présenter des similitudes – ou différer – en ce qui a trait aux effets du programme sur un résultat donné. Cette variabilité (on

parle alors d'« hétérogénéité » [27]) dépend, en partie, du champ de l'étude en cause. Si celui-ci est vaste, selon toute vraisemblance, l'écart et, par conséquent, la variabilité des études retenues le seront aussi. Par contre, lorsque ce champ est restreint, les études retenues seront probablement plus semblables les unes aux autres.

Les variations ou l'hétérogénéité résulteront d'une différence considérable entre les participants, les interventions ou les résultats des études si ces facteurs influent sur l'effet de l'intervention. L'effet réel de l'intervention étant différent d'une étude à l'autre, établir l'effet moyen sur l'ensemble des études ne sera, en l'occurrence, d'aucune utilité.

Selon le degré de variabilité, diverses méthodes seront mises à contribution pour résumer les données issues des études retenues. En voici quelques exemples:

- *Calcul de l'effet moyen (ou groupé) sur l'ensemble des études* – Cette méthode est utile quand la variabilité est faible dans l'ensemble des études. Par exemple, selon une étude méthodique des programmes « de congé précoce et de suivi à domicile » (c.-à-d. des programmes où des prestataires de services de santé assurent la prestation de traitements actifs au domicile du patient pour un problème de santé qui exigerait normalement des soins actifs en milieu hospitalier), les études retenues étaient suffisamment semblables pour qu'il soit possible d'estimer l'effet moyen du programme. Les données probantes étaient cependant trop rares pour démontrer les avantages économiques ou les bienfaits pour la santé des programmes « de congé précoce et de suivi à domicile » [43]
- *Calcul de l'effet moyen pour les sous-groupes d'études inclus dans un examen* – Cette méthode peut se révéler utile lorsque la variabilité globale des études retenues est élevée (et qu'il est donc inutile de calculer un effet moyen) mais que la variabilité est faible parmi les sous-groupes d'études. Ainsi, aux fins de l'examen des interventions de travailleurs de la santé non professionnels dans les services de première ligne et communautaires, les études ont été regroupées en fonction des problèmes de santé traités. Pour certains groupes, les travailleurs de la santé non professionnels chargés de la promotion de l'immunisation et de l'allaitement maternel par exemple, il a été possible de calculer un effet moyen sur l'ensemble des études pertinentes. Les résultats de l'examen révèlent que ce groupe peut effectivement améliorer le recours à l'immunisation et à l'allaitement maternel [44]
- *Description de la gamme de valeurs des effets* – Lorsque les études ne sont pas assez similaires pour justifier le calcul d'un effet moyen, il est quand même possible de décrire l'ensemble des effets des études. Ainsi, selon un examen des effets de la vérification et de la rétroaction sur la pratique des prestataires de services de santé, la conformité à la pratique souhaitée varie d'une diminution de 16 p. 100 à une augmentation de 70 p. 100, la médiane étant de 5 p. 100. L'examen laisse entendre que la vérification et la rétroaction peuvent accroître l'efficacité de la pratique, mais que leurs effets varient généralement de faibles à modérés [45]

- *Catalogage des types d'interventions visant à résoudre un problème particulier* – Étant donné la vaste portée de certains examens et, par conséquent, la variabilité des études sur lesquelles ils portent, il n'est pas toujours réaliste de tenter de combiner quantitativement les résultats de ces études – ou même de décrire la gamme de valeurs des effets. L'examen narratif est la solution la plus appropriée en pareil cas. Ainsi, cette étude méthodique de l'efficacité des interventions des services de santé en vue de réduire les inégalités en santé qui portait sur des études évaluant des programmes ayant pour objet de diminuer ces écarts et qui peuvent être mis en œuvre, soit uniquement à l'échelle du système de santé, soit en collaboration avec d'autres organismes. Comme la gamme d'études examinées était vaste et incluait aussi bien des programmes visant à améliorer le contrôle de la tension artérielle que des interventions de promotion de la santé, aucun regroupement statistique n'a été tenté [46]

Lorsque les résultats diffèrent d'une étude à l'autre, les questions à envisager sont les suivantes :

- *Y a-t-il une explication convaincante aux différences relevées?* Il pourrait s'agir de différences entre les participants, les interventions, les groupes témoins, les résultats, les milieux ou les périodes visés par les études retenues. Par exemple, les participants de certaines études pourraient provenir d'un groupe d'âge plus large ou souffrir d'affections préexistantes diverses
- *Dans quelle mesure une estimation groupée peut-elle être significative?* Lorsqu'un examen porte sur des études très diversifiées, une estimation groupée a peu de chances d'être significative. On pourrait alors pousser plus loin l'exploration des données, par une analyse de sous-groupes, mais il est permis de douter des résultats d'analyses exploratoires de ce genre. Étant donné l'augmentation du nombre d'études méthodiques disponibles, il est fréquent d'en trouver plus d'une qui traite d'une question de politique particulière, mais leurs résultats ou conclusions ne sont pas toujours uniformes. Le tableau 5 offre aux responsables de politiques des conseils sur la façon d'aborder de telles situations

CONCLUSION

Manifestement, les études méthodiques ne sont pas toutes menées avec la même rigueur. Aussi importe-t-il d'évaluer la fiabilité des études visant à éclairer les décisions relatives aux politiques pour être en mesure de juger du degré de confiance à leur accorder. Ces évaluations doivent être effectuées de manière systématique et transparente et mettre à contribution les outils conçus à cet effet. Cependant, ces outils ne peuvent servir à évaluer que ce qui est connu. C'est pourquoi toute évaluation à l'aide de ces outils doit être soigneusement pesée.

Si le degré de fiabilité d'une étude méthodique est faible, les responsables de politiques ne devraient accorder qu'une confiance partielle à ces données et les utiliser avec

circonspection pour éclairer leurs décisions (la figure 2 aborde cet aspect). Outre les évaluations de la fiabilité de l'étude méthodique, les décideurs doivent aussi tenir compte d'autres facteurs, notamment de l'utilité de l'étude méthodique au regard de la question de politique et des données probantes dans le contexte local.

RESSOURCES

Documents utiles et lecture complémentaire

Higgins JPT, Altman DF. 2008. « Chapter 8: Assessing risk of bias in included studies ». dans Higgins JPT et S Green (dir.), *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.0.1 (mise à jour : septembre 2008). Cochrane Collaboration. www.cochrane-handbook.org

Counsell C. 1997. « Formulating Questions and Locating Primary Studies for Inclusion in Systematic Reviews ». *Ann Intern Med.* 127: 380-387

Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C et coll. 2007. « Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews ». *BMC Med Res Methodol* ; 7:10. www.biomedcentral.com/1471-2288/7/10

Sites Web

La base de données *Rx for Change* contient les sommaires des données probantes issues de la recherche sur l'incidence des stratégies visant à améliorer les habitudes de prescription et l'utilisation des médicaments ainsi que des évaluations de la fiabilité d'études méthodiques sur les répercussions d'interventions destinées aux professionnels, aux organismes de soins de santé, aux décideurs et aux consommateurs. www.acmts.ca/index.php/fr/compus/optimal-ther-resources/interventions/rx-for-change-fact-sheet

Le Groupe de revue des pratiques efficaces et de l'organisation des soins de la collaboration Cochrane (Cochrane Effective Practice and Organisation of Care Review Group – EPOC) fournit des lignes directrices pour évaluer la fiabilité de différents types d'études de l'efficacité. www.epoc.cochrane.org/en/index.html

Le projet SUPPORT (SUPporting POLicy relevant Reviews and Trials) présente des sommaires d'études méthodiques sur des questions hautement prioritaires pour les PRFI, notamment des évaluations de la fiabilité. www.support-collaboration.org/index.htm

CONFLIT D'INTÉRÊTS

Les auteurs déclarent n'avoir aucun conflit d'intérêts.

CONTRIBUTION DE CHAQUE AUTEUR

SL a rédigé l'ébauche initiale du présent article; ADO, JNL et AF ont contribué à ce travail en plus de réviser le document.

REMERCIEMENTS

L'introduction de la présente série comprend nos remerciements aux bailleurs de fonds et aux personnes qui y ont contribué. Nous tenons toutefois à exprimer notre reconnaissance particulière à Duff Montgomerie pour ses commentaires avisés sur une version antérieure du présent article.

RÉFÉRENCES BIBLIOGRAPHIQUES

1. Lavis JN, Oxman AD, Lewin S, Fretheim A: **SUPPORT Tools for evidence-informed health Policymaking (STP). Introduction.** *Health Res Policy Syst* 2009, **7 (Suppl 1:I1)**.
2. Lavis JN, Posada FB, Haines A, Osei E: **Use of research to inform public policymaking.** *Lancet* 2004, **364**:1615-21.
3. Oxman AD, Lavis JN, Lewin S, Fretheim A: **SUPPORT Tools for evidence-informed health Policymaking (STP). 1. What is evidence-informed policymaking.** *Health Res Policy Syst* 2009, **7 (Suppl 1:S1)**.
4. Oxman AD, Fretheim A, Lavis JN, Lewin S: **SUPPORT Tools for evidence-informed health Policymaking (STP). 12. Finding and using research evidence about resource use and costs.** *Health Res Policy Syst* 2009, **7 (Suppl 1:S12)**.
5. Renfrew MJ, Craig D, Dyson L, McCormick F, Rice S, King SE, et al: **Breastfeeding promotion for infants in neonatal units: a systematic review and economic analysis.** *Health Technol Assess* 2009, **13**:1-iv.
6. Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, et al: **Effectiveness and efficiency of guideline dissemination and implementation strategies.** *Health Technol Assess* 2004, **8**:iii-72.
7. Carlsen B, Glenton C, Pope C: **Thou shalt versus thou shalt not: a meta-synthesis of GPs' attitudes to clinical practice guidelines.** *Br J Gen Pract* 2007, **57**:971-8.
8. 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* 2005, **10 Suppl 1**:6-20.

9. Munro SA, Lewin SA, Smith HJ, Engel ME, Fretheim A, Volmink J: **Patient adherence to tuberculosis treatment: a systematic review of qualitative research.** *PLoS Med* 2007, **4**:e238.
10. Pound P, Britten N, Morgan M, Yardley L, Pope C, Daker-White G, et al: **Resisting medicines: a synthesis of qualitative studies of medicine taking.** *Soc Sci Med* 2005, **61**:133-55.
11. Laurant M, Reeves D, Hermens R, Braspenning J, Grol R, Sibbald B: **Substitution of doctors by nurses in primary care.** *Cochrane Database Syst Rev* 2005, **2**:CD001271.
12. Forsetlund L, Bjorndal A, Rashidian A, Jamtvedt G, O'Brien MA, Wolf F, et al: **Continuing education meetings and workshops: effects on professional practice and health care outcomes.** *Cochrane Database Syst Rev* 2009, **2**:CD003030.
13. Mulrow CD: **Rationale for systematic reviews.** *BMJ* 1994, **309**:597-9.
14. Oxman AD, Schunemann HJ, Fretheim A: **Improving the use of research evidence in guideline development: 8. Synthesis and presentation of evidence.** *Health Res Policy Syst* 2006, **4**:20.
15. Oxman AD, Glasziou P, Williams JW, Jr.: **What should clinicians do when faced with conflicting recommendations?** *BMJ* 2008, **337**:a2530.
16. Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, Moher M, et al: **Methodology and reports of systematic reviews and meta-analyses: a comparison of Cochrane reviews with articles published in paper-based journals.** *JAMA* 1998, **280**:278-80.
17. Jorgensen AW, Hilden J, Gotsche PC: **Cochrane reviews compared with industry supported meta-analyses and other meta-analyses of the same drugs: systematic review.** *BMJ* 2006, **333**:782.
18. Dixon E, Hameed M, Sutherland F, Cook DJ, Doig C: **Evaluating meta-analyses in the general surgical literature: a critical appraisal.** *Ann Surg* 2005, **241**:450-9.
19. Jadad AR, Cook DJ, Browman GP: **A guide to interpreting discordant systematic reviews.** *CMAJ* 1997, **156**:1411-6.
20. Jadad AR, Moher M, Browman GP, Booker L, Sigouin C, Fuentes M, et al: **Systematic reviews and meta-analyses on treatment of asthma: critical evaluation.** *BMJ* 2000, **320**:537-40.
21. Linde K, Willich SN: **How objective are systematic reviews? Differences between reviews on complementary medicine.** *J R Soc Med* 2003, **96**:17-22.
22. 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.
23. Critical Appraisal Skills Programme: *10 questions to help you make sense of reviews.* United Kingdom, Public Health Resource Unit. 2006.
24. Oxman AD, Guyatt GH: **Validation of an index of the quality of review articles.** *J Clin Epidemiol* 1991, **44**:1271-8.

25. Canadian Coordinating Office for Health Technology Assessment: *Proposed Evaluation Tools for COMPUS*. Ottawa, Ottawa: Canadian Coordinating Office for Health Technology Assessment. 2005.
26. West S, King V, Carey TS, Lohr KN, McKoy N, Sutton SF, et al: *Systems to rate the strength of scientific evidence* [Evidence report/technology assessment no 47]. Publication No. 02-E016. Rockville, MD, USA, Agency for Healthcare Research and Quality. 2002.
27. Higgins JPT, Altman DF: **Chapter 8: Assessing risk of bias in included studies**. In *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1 (updated September 2008)*. Edited by Higgins JPT, Green S. The Cochrane Collaboration; 2008.
28. Lavis JN, Oxman AD, Souza NM, Lewin S, Gruen RL, Fretheim A: **SUPPORT Tools for evidence-informed health Policymaking (STP). 9. Assessing the applicability of the findings of a systematic review**. *Health Res Policy Syst* 2009, **7 (Suppl 1:S9)**.
29. Counsell C: **Formulating Questions and Locating Primary Studies for Inclusion in Systematic Reviews**. *Ann Intern Med* 1997, **127**:380-7.
30. Higgins JPT, Green S: *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1 [updated September 2008]*. The Cochrane Collaboration; 2008.
31. Meremikwu MM, Donegan S, Esu E: **Chemoprophylaxis and intermittent treatment for preventing malaria in children**. *Cochrane Database of Systematic Reviews* 2008, **2**:CD003756.
32. Mulrow CD: **The medical review article: state of the science**. *Ann Intern Med* 1987, **106**:485-8.
33. Oxman AD: **Checklists for review articles**. *BMJ* 1994, **309**:648-51.
34. Grobler LA, Marais BJ, Mabunda S, Marindi P, Reuter H, Volmink J: **Interventions for increasing the proportion of health professionals practising in underserved communities**. *Cochrane Database of Systematic Reviews* 2009, **1**:CD005314.
35. Lefebvre C, Manheimer E, Glanville J, on behalf of the Cochrane Information Retrieval Methods Group: **Searching for studies**. In *Cochrane Handbook for systematic reviews of interventions. Version 5.0.1 [updated September 2008]*. Edited by Higgins JPT, Green S. The Cochrane Collaboration; 2008.
36. Dickersin K, Min YI: **Publication bias: the problem that won't go away**. *Ann N Y Acad Sci* 1993, **703**:135-46.
37. 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.
38. Hopewell S, Clarke M, Stewart L, Tierney J: **Time to publication for results of clinical trials**. *Cochrane Database Syst Rev* 2007, **2**:MR000011.
39. Lavis JN, Wilson M, Oxman AD, Lewin S, Fretheim A: **SUPPORT Tools for evidence-informed health Policymaking (STP). 4. Using research evidence to clarify a problem**. *Health Res Policy Syst* 2009, **7 (Suppl 1:S4)**.

40. Lavis JN, Wilson MG, Oxman AD, Grimshaw J, Lewin S, Fretheim A: **SUPPORT Tools for evidence-informed health Policymaking (STP). 5. Using research evidence to frame options to address a problem.** *Health Res Policy Syst* 2009, **7 (Suppl 1:S5)**.
41. Katrak P, Bialocerkowski A, Massy-Westropp N, Kumar VS, Grimmer K: **A systematic review of the content of critical appraisal tools.** *BMC Medical Research Methodology* 2004, **4:22**.
42. Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S: **Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists.** *Control Clin Trials* 1995, **16:62-73**.
43. Shepperd S, Doll H, Broad J, Gladman J, Iliffe S, Langhorne P, et al: **Early discharge hospital at home.** *Cochrane Database Syst Rev* 2009, **1:CD000356**.
44. Lewin SA, Dick J, Pond P, Zwarenstein M, Aja G, van Wyk B, et al: **Lay health workers in primary and community health care.** *Cochrane Database Syst Rev* 2005, **1:CD004015**.
45. Jamtvedt G, Young JM, Kristoffersen DT, O'Brien MA, Oxman AD: **Audit and feedback: effects on professional practice and health care outcomes.** *Cochrane Database Syst Rev* 2006, **2:CD000259**.
46. Arblaster L, Lambert M, Entwistle V, Forster M, Fullerton D, Sheldon T, et al: **A systematic review of the effectiveness of health service interventions aimed at reducing inequalities in health.** *J Health Serv Res Policy* 1996, **1:93-103**.
47. Moher D, Jadad AR, Klassen TP: **Guides for reading and interpreting systematic reviews: III. How did the authors synthesize the data and make their conclusions?** *Arch Pediatr Adolesc Med* 1998, **152:915-20**.
48. Oxman AD, Cook DJ, Guyatt GH: **Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group.** *JAMA* 1994, **272:1367-71**.
49. Oxman AD, Lavis JN, Lewin S, Fretheim A: **SUPPORT Tools for evidence-informed health Policymaking (STP). 10. Taking equity into consideration when assessing the findings of a systematic review.** *Health Res Policy Syst* 2009, **7 (Suppl 1:S10)**.
50. Oxman AD, Lavis JN, Fretheim A, Lewin S: **SUPPORT Tools for evidence-informed health Policymaking (STP). 17. Dealing with insufficient research evidence.** *Health Res Policy Syst* 2009, **7 (Suppl 1:S17)**.
51. Dixon-Woods M, Agarwal S, Jones D, Young B, Sutton A: **Synthesising qualitative and quantitative evidence: a review of possible methods.** *J Health Serv Res Policy* 2005, **10:45-53**.
52. Noyes J, Popay J, Pearson A, Hannes K, Booth A: **Chapter 20: Qualitative research and Cochrane reviews.** In *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1 (updated September 2008)*. Edited by Higgins JPT, Green S. The Cochrane Collaboration; 2008.
53. Lavis JN: **Supporting the Use of Systematic Reviews in Policymaking.** *PLoS Med.* 2010. In Press

54. Jefferson T, Demicheli V, Vale L: **Quality of systematic reviews of economic evaluations in health care.** *JAMA* 2002, **287**:2809-12.
55. Siegfried N, Muller M, Deeks JJ, Volmink J: **Male circumcision for prevention of heterosexual acquisition of HIV in men.** *Cochrane Database Syst Rev* 2009, **2**:CD003362.
56. McGuinness B, Craig D, Bullock R, Passmore P: **Statins for the prevention of dementia.** *Cochrane Database Syst Rev* 2009, **2**:CD003160.

Figure 1. Repérage et évaluation des études méthodiques visant à éclairer les décisions au sujet de propositions de politiques et de programmes

<i>Questions</i>	<i>Raisons d'accorder moins de confiance à une étude méthodique</i>
L'étude méthodique traite-t-elle explicitement d'une question appropriée de politique ou de gestion?	<ul style="list-style-type: none"> - Les études qui traitent de questions floues... sont susceptibles d'offrir des réponses floues - Une étude traitant d'une question qui diffère de la question de politique ou de gestion qui vous intéresse... peut induire en erreur
A-t-on utilisé des critères appropriés pour sélectionner les études à examiner?	<ul style="list-style-type: none"> - Des critères implicites... peuvent fausser le choix des études à inclure - Des critères qui ne sont pas compatibles avec la question sur laquelle porte l'étude méthodique... sont susceptibles d'occasionner la sélection d'études non pertinentes pour répondre à la question
La recherche d'études pertinentes a-t-elle été approfondie et suffisamment exhaustive?	<ul style="list-style-type: none"> - Les recherches qui ne sont pas bien planifiées (et décrites)... risquent de ne pas permettre de repérer les données probantes importantes et pertinentes - Les recherches qui ne sont pas suffisamment exhaustives... risquent d'être entachées de partialité (p. ex., elles sont plus susceptibles de repérer des études dont les résultats sont « positifs ») et de ne pas trouver les données probantes importantes et pertinentes.
Les évaluations de la pertinence des études par rapport au sujet de l'examen ainsi que du risque de subjectivité sont-elles reproductibles?	<ul style="list-style-type: none"> - Si les critères utilisés ne sont pas explicites ou si tous les jugements qui s'imposent relativement à une étude sont portés par une seule personne... ces jugements ne seront pas nécessairement reproductibles et pourraient être partiels.
Les résultats sont-ils semblables d'une étude à l'autre?	<ul style="list-style-type: none"> - S'il y a des écarts importants dans les résultats d'études différentes et qu'il n'y a pas d'explication convaincante à ces divergences... il est difficile de discerner les vrais résultats ou de savoir à quoi s'attendre si l'on met en œuvre une politique semblable à celles qui ont été évaluées.

Figure 2. Comment déceler les éléments sujets à caution ou trompeurs dans les études méthodiques

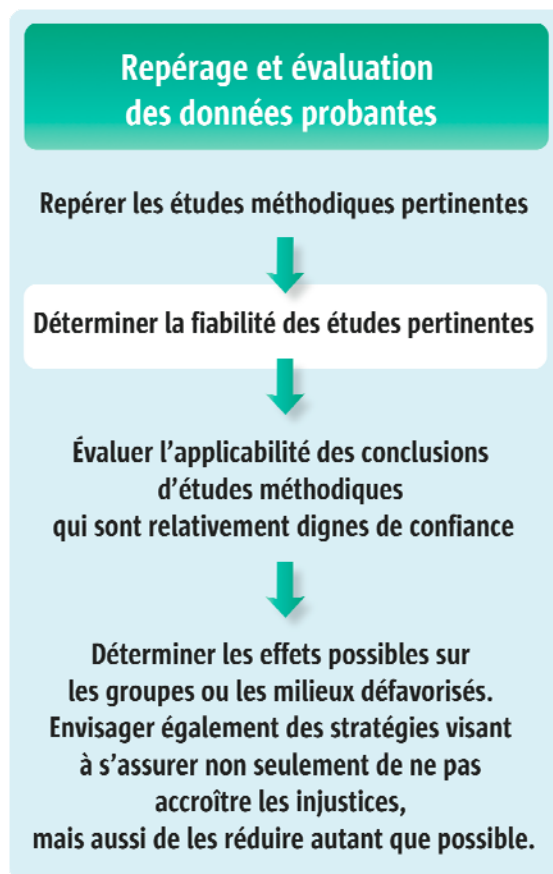


Tableau 1. AMSTAR – Instrument d'évaluation des études méthodiques [22]

Pour chacune des questions énoncées, les auteurs d'études méthodiques doivent choisir parmi les réponses suivantes : Oui / Non / Ne peut répondre / Sans objet

1. Les modalités de l'étude méthodique ont-elles été décrites a priori?

La question à l'étude et les critères d'inclusion doivent être établis avant le début de l'examen

2. Y a-t-il eu duplication de la sélection des études et de l'extraction des données?

Il doit y avoir au moins deux extracteurs de données indépendants, et la procédure consensuelle en cas de désaccord doit être énoncée

3. La recherche documentaire a-t-elle été exhaustive?

Au moins deux sources électroniques doivent être consultées. Le rapport doit préciser les années visées par l'étude et les bases de données utilisées (p. ex., CENTRAL, EMBASE et MEDLINE). Les mots-clés ou les descripteurs MeSH doivent y figurer et, dans la mesure du possible, la stratégie de recherche doit être décrite. Toutes les recherches doivent être complétées par la consultation du contenu à jour de revues, d'ouvrages, de registres spécialisés et d'experts du domaine d'étude ainsi qu'en prenant connaissance des références bibliographiques des études répertoriées

4. A-t-on considéré le statut de la publication (p. ex., littérature grise) comme étant un critère d'inclusion?

Les auteurs doivent préciser qu'ils ont effectué leur recherche de rapports sans égard au type de publication et indiquer s'ils ont exclu des rapports (de l'étude méthodique) en raison du statut de publication, de la langue, etc.

5. Une liste des études (retenues et rejetées) a-t-elle été fournie?

Il est essentiel de fournir une liste des études retenues et rejetées

6. Les caractéristiques des études retenues ont-elles présentées?

Les données relatives aux études originales portant sur les participants, les interventions et les résultats doivent être présentées sous forme groupées, p. ex., dans un tableau. Dans toutes les études analysées, les plages de valeurs doivent être précisées pour les différentes caractéristiques : l'âge, l'appartenance ethnique, le sexe, les données socioéconomiques pertinentes, le problème de santé, la durée et la gravité de la maladie et tout autre affection

7. La qualité scientifique des études retenues a-t-elle été évaluée et documentée?

Les méthodes d'évaluation établies a priori doivent être décrites (p. ex., pour les études d'efficacité, si les auteurs ont choisi de ne retenir que des études contrôlées par placebo à double insu après répartition aléatoire, ou encore le secret de l'affectation, comme critère d'inclusion); pour d'autres types d'études, des critères différents seront pris en compte

8. La qualité scientifique des études retenues a-t-elle été utilisée correctement dans la formulation des conclusions?

Les résultats ayant trait à la rigueur méthodologique et à la qualité scientifique des études doivent être pris en compte dans l'analyse et les conclusions de l'examen, et explicitement décrits dans la formulation des recommandations

9. Les méthodes employées pour combiner les résultats des études étaient-elles appropriées?

Lorsque les résultats sont groupés, un test doit être effectué pour déterminer si les études pouvaient être combinées, c.à-d. pour évaluer leur homogénéité (test du chi carré, ou χ^2). En présence d'hétérogénéité, un modèle d'effets aléatoires doit être utilisé ou la pertinence clinique du regroupement doit être prise en considération (c.-à-d., fallait-il effectuer le regroupement des résultats?)

10. A-t-on évalué la possibilité d'un biais de publication?

L'évaluation d'un biais de publication doit faire appel à une combinaison d'aides graphiques (p.ex., graphiques en entonnoir) ou de tests statistiques (p. ex., analyse de régression d'Egger)

11. Les conflits d'intérêts ont-ils été énoncés?

Les sources de soutien éventuelles doivent être clairement énoncées, tant dans l'examen que dans les études retenues

Tableau 2. Interprétation des résultats des études méthodiques sur les effets d'interventions

Les questions suivantes peuvent aider les responsables de politiques à interpréter les conclusions d'une étude méthodique sur les effets d'interventions [33,47,48]* :

- *Comment l'estimation de l'effet est-elle présentée?* Nombre d'examens présentent une estimation moyenne de l'effet sur l'ensemble des études retenues, souvent forme de risque relatif, de ratio d'incidence approché ou de différence moyenne normalisée
- *Une estimation moyenne de l'effet sur l'ensemble des études est-elle indiquée?* Dans le cadre des examens, on a recours à des méthodes statistiques pour résumer et combiner les données sur les résultats des études retenues. Afin de s'assurer que cette démarche est bien indiquée, il convient de se demander si les études retenues étaient suffisamment semblables sur les plans de la population, de l'intervention, de la comparaison et des résultats mesurés. Lorsque l'estimation moyenne de l'effet est impossible à établir, les examens systématiques présentent habituellement un survol narratif des données disponibles
- *Les limites de confiance de l'estimation de l'effet sont-elles présentées?* L'étude méthodique doit faire état des intervalles de confiance relatifs à l'estimation moyenne de l'effet. Plus l'intervalle de confiance est large, moins il est possible de confirmer l'ampleur réelle de l'effet avec certitude
- *Si les résultats d'analyses de sous-groupes sont présentés, dans quelle mesure sont-ils appropriés?* Un examen peut présenter les conclusions relatives à un sous-groupe donné de participants dans l'ensemble des essais cliniques ou pour un sous-groupe d'études [49]. Par exemple, un examen des interventions visant à réduire les maladies diarrhéiques chez les enfants de moins de cinq ans pourrait aussi prendre en considération les effets des interventions sur les enfants de moins d'un an. De même, un examen pourrait comporter une analyse de sous-groupe des études considérées comme comportant un faible risque de partialité. L'analyse de sous-groupe serait alors indiquée, tant par rapport à la question globale envisagée qu'à la connaissance préalable des facteurs susceptibles d'influencer ou d'atténuer les effets de l'intervention. Par exemple, on pourrait supposer que l'importance des effets est proportionnelle à l'intensité de l'intervention. Les analyses de sous-groupes doivent être planifiées avant le début de l'examen. Il faut se rappeler toutefois que leurs résultats sont moins fiables que ceux des analyses reposant sur *toutes* les études retenues; en outre, des analyses statistiques multiples peuvent, par simple hasard, donner des résultats positifs

- *En l'absence de données probantes sur les effets d'une intervention, a-t-on veillé à ne pas associer cette lacune à l'absence réelle de toute répercussion?* L'absence de données probantes sur les impacts d'une proposition n'est nullement signe que l'intervention est sans effet. Dans le premier cas, les données probantes sont insuffisantes pour tirer des conclusions quant aux effets de l'intervention, alors que dans le deuxième, les données des études retenues montrent clairement que l'intervention *n'a pas* les effets anticipés [50]
- *Les conclusions et les recommandations (le cas échéant) découlent-elles à la fois de la question établie a priori et des données probantes issues de l'étude méthodique?* Il est essentiel de se demander si les conclusions présentées par les auteurs de l'étude sont tirées directement des données recueillies dans le cadre de l'examen et si elles ne vont pas au-delà de ces données probantes
- *Les données probantes peuvent-elles s'appliquer à la question de politique envisagée?* En raison des différences entre les systèmes de santé, il se peut qu'un programme (ou une intervention) jugé efficace dans un milieu donné fonctionne différemment dans un autre. Les responsables de politiques doivent évaluer si les données probantes issues d'études méthodiques s'appliquent dans leur contexte. L'article 9 de la présente série présente des lignes directrices à cet égard [28]

* Le chevauchement entre les questions susmentionnées et celles destinées à l'évaluation de la fiabilité des études méthodiques s'explique par l'impérieuse nécessité de tenir compte de la fiabilité au moment d'évaluer et de comprendre les résultats d'une étude méthodique

Tableau 3. Évaluation de la confiance à accorder aux conclusions des examens systématiques d'études qualitatives et d'études économiques

De plus en plus, les examens systématiques d'études qualitatives foisonnent. Ils ont recours aux méthodes les plus diverses, dont la synthèse narrative, la méta-ethnographie et la revue réaliste. Non seulement les examens d'études qualitatives sont-ils porteurs d'une information cruciales, mais ils peuvent en outre éclairer et enrichir les études méthodiques des effets [51,52]. Toutefois, il est essentiel que le lecteur vérifie la fiabilité de ces examens. Bien qu'il existe encore peu d'outils conçus à cette fin, bon nombre des questions visant à orienter les responsables de politiques dans l'évaluation de la fiabilité des études méthodiques des effets peuvent également servir aux examens d'études qualitatives. En voici quelques exemples :

1. *L'examen traite-t-il d'une question de politique ou de gestion appropriée?* La question à envisager doit être susceptible d'être traitée au moyen de données qualitatives et devrait être appropriée à l'élaboration de politiques. Les examens d'études qualitatives peuvent aider à comprendre les opinions et les expériences des intervenants en ce qui a trait à la santé et aux services de santé, ce qui peut aider à clarifier le problème [39]. Les examens d'études qualitatives peuvent aussi permettre de comprendre comment et pourquoi une option fonctionne (par exemple, en examinant les évaluations des processus menées parallèlement à la mise en œuvre d'une politique ou d'un programme) et de percevoir les opinions des intervenants au sujet des propositions ainsi que leurs expériences à ce propos [40,53]
2. *A-t-on utilisé des critères appropriés pour sélectionner les études à examiner?* La description de la méthode de sélection des études doit être appropriée à la question de recherche
3. *Le mode de repérage des études est-il clairement expliqué?* Certains examens d'études qualitatives font l'objet d'une recherche documentaire exhaustive, tandis que d'autres procèdent par échantillonnage. La méthode choisie devrait être clairement décrite et justifiée
4. *Les critères d'évaluation de la fiabilité des études retenues sont-ils appropriés?* On devrait décrire dans l'examen comment la fiabilité des études incluses a été évaluée.
5. *A-t-on eu recours à une méthode appropriée pour analyser les conclusions des études retenues?* Dans un examen, on devrait avoir recours à une méthode de synthèse reconnue et justifier le recours à cette méthode

Questions à envisager pour évaluer la fiabilité d'un examen d'études économiques [54].

1. Est-il peu probable que des études importantes et pertinentes aient été omises?
2. Les critères d'inclusion employés pour choisir les articles sont-ils appropriés?
3. L'évaluation des études est-elle reproductible?
4. La méthodologie et le sujet des études retenues sont généralement comparables?
5. Dans quelle mesure les résultats globaux sont-ils reproductibles?
6. Les résultats contribueront-ils à l'affectation des ressources en santé?

Tableau 4. Exemples de sources consultées dans le cadre des études méthodiques

Examen	Sources consultées
<p><i>Examen sur les systèmes de santé</i> Exemple : Étude méthodique des interventions des travailleurs de la santé non professionnels dans les services de première ligne et communautaires [44]</p>	<ol style="list-style-type: none"> 1. Bases de données électroniques d'études publiées : <ul style="list-style-type: none"> • MEDLINE • Cochrane Central Register of Controlled Trials (CENTRAL) et registres spécialisés de la Collaboration Cochrane (EPOC et Consumers and Communication Review Groups) • Science Citations Index (SCI) • EMBASE • CINAHL (Cumulative Index to Nursing and Allied Health Literature) • HealthSTAR (Health Services Technology, Administration, and Research) • AMED (Allied and Complementary Medicine Database) • Leeds Health Education Effectiveness Database 2. Bibliographies des études évaluées en vue de leur inclusion dans l'examen 3. Demande d'information détaillée aux auteurs consultés sur les recherches complémentaires
<p><i>Étude sur la santé publique</i> Exemple : Étude méthodique de la circoncision masculine en vue de prévenir l'infection par le VIH chez les hétérosexuels [55]</p>	<ol style="list-style-type: none"> 1. Bases de données électroniques d'études publiées : <ul style="list-style-type: none"> • MEDLINE • EMBASE • Cochrane Central Register of Controlled Trials (CENTRAL) 2. Bases de données d'actes de conférences : <ul style="list-style-type: none"> • AIDSearch 3. Bases de données sur les essais cliniques en cours : <ul style="list-style-type: none"> • ClinicalTrials.gov • Current Controlled Trials 4. Communication avec les chercheurs et organismes pertinents du domaine. 5. Vérification des références bibliographiques de toutes les études recensées à l'aide des méthodes susmentionnées; examen de l'ensemble des études méthodiques, des méta-analyses et des lignes directrices en matière de prévention repérées dans le cadre de la recherche documentaire

Examen	Sources consultées
<p><i>Études cliniques</i> Exemple : Étude méthodique du recours aux statines pour la prévention de la démence [56]</p>	<ol style="list-style-type: none"> 1. Bases de données électroniques : <ul style="list-style-type: none"> • Registre spécialisé du Cochrane Dementia and Cognitive Improvement Group (Réseau des bases factuelles en santé) • Cochrane Central Register of Controlled Trials (CENTRAL) • MEDLINE • EMBASE • PsycINFO (base de données de l'American Psychological Association) • CINAHL • SIGLE (System of Information on Grey Literature in Europe – Système d'information sur la littérature grise européenne) • LILACS (Literatura Latinoamericana y del Caribe en Ciencias de la Salud) 2. Bases de données d'actes de conférences : <ul style="list-style-type: none"> • ISTP (Index to Scientific and Technical Proceedings – États-Unis) • British Library Inside 3. Bases de données de thèses et mémoires : <ul style="list-style-type: none"> • Index to Theses, ou ASLIB Index (thèses de recherche du Royaume-Uni et d'Irlande) • <i>Australian Digital Theses Program</i> (thèses d'étudiants au doctorat d'universités australiennes) • Thèses Canada (thèses de doctorat et mémoires de maîtrise canadiens) • DATAD – Database of African Theses and Dissertations (thèses et mémoires africains) • Dissertation Abstract Online (États-Unis; regroupe thèses et mémoires produits dans les universités américaines et britanniques) 4. Bases de données sur les essais cliniques en cours : <ul style="list-style-type: none"> • Des recherches ont été faites dans un grand nombre de bases de données sur les essais cliniques

Tableau 5. Mesures à envisager par les responsables de politiques lorsque des études méthodiques traitant des mêmes questions produisent des résultats différents

Il est fréquent que le repérage de données probantes visant à éclairer une décision particulière en matière de politique permette de recenser plus d'une étude méthodique pertinente. Il arrive aussi que les résultats de ces études diffèrent et que leurs auteurs ne tirent pas les mêmes conclusions quant aux effets d'une intervention. Le scénario est différent lorsque les conclusions de deux ou plusieurs études concordent, mais que les chercheurs ou d'autres parties intéressées ne s'entendent pas sur leur interprétation [19].

Plusieurs facteurs peuvent expliquer ces différences, notamment les questions abordées par les examens, les critères d'inclusion et d'exclusion, le type de données issues des études, le mode d'évaluation de la qualité des études et les décisions relatives à l'analyse statistique des données (de même que les méthodes employées à cette fin) [19].

Les questions suivantes, élaborées par A.R. Jadad et ses collaborateurs, peuvent orienter le repérage de conclusions divergentes et donner des pistes de solution [19].

- Les études méthodiques abordent-elles toutes la même question? Dans la négative, il faudra retenir l'étude portant sur la question qui cadre le plus avec la question de politique devant être éclairée par les données probantes ou encore évaluer les résultats correspondant le mieux à la question de politique
- Si les examens abordent la même question, incluent-ils les mêmes essais ou études primaires? Dans la négative, il convient de choisir l'examen qui inclut les études les plus pertinentes par rapport à l'enjeu stratégique à l'étude
- Si les examens systématiques portent sur les mêmes études sont-ils tous d'égale qualité? Dans la négative, il faut mettre à profit l'examen de qualité supérieure

Lorsque deux examens sont pertinents (si, par exemple, ils abordent différents aspects de la même question), il pourrait être utile de dégager des données probantes de l'un et l'autre.