



Trousse d'outils de l'évaluation critique pour l'évaluation de plusieurs types de données probantes

D Moralejo¹, T Ogunremi^{2*}, K Dunn²

Résumé

Les professionnels de la santé sont souvent censés effectuer une évaluation critique de données probantes issues de la recherche afin de formuler des recommandations pour l'élaboration des pratiques et des politiques. Dans le présent numéro, nous décrivons la trousse d'outils de l'évaluation critique actuellement utilisée par l'Agence de la santé publique du Canada. La trousse d'outils comprend les éléments suivants : des algorithmes pour déterminer le type de plan d'étude, trois outils distincts (aux fins d'évaluation de la qualité des études analytiques, des études descriptives et des analyses documentaires), d'autres outils pour appuyer le processus d'évaluation, ainsi que des conseils pour effectuer la synthèse des données probantes et tirer des conclusions sur un ensemble de données probantes. Bien que la trousse d'outils ait été créée pour aider à l'élaboration de lignes directrices nationales relatives à la prévention et au contrôle des infections, les cliniciens, les responsables des politiques et les étudiants peuvent l'utiliser pour guider l'évaluation de toute recherche quantitative liée à la santé. Les participants à un essai pilote ont effectué 101 évaluations critiques, au total, et ont trouvé que la trousse d'outils était conviviale et utile dans le processus d'évaluation critique. Les commentaires des participants à l'essai pilote de la trousse d'outils ont permis d'éclairer davantage les révisions avant sa publication. La trousse d'outils s'ajoute à l'arsenal d'outils accessibles et peut s'avérer particulièrement utile lorsque les meilleures données probantes accessibles proviennent d'essais ou d'études non cliniques comportant des lacunes de conception, où d'autres outils peuvent ne pas être facilement appliqués.

Citation proposée : Moralejo D, Ogunremi T, Dunn K. Trousse d'outils de l'évaluation critique pour l'évaluation de plusieurs types de données probantes. *Relevé des maladies transmissibles au Canada*. 2017;43(9):199-205. <https://doi.org/10.14745/ccdr.v43i09a02f>

Introduction

Les professionnels de la santé, les chercheurs et les responsables des politiques participent souvent à l'élaboration de politiques ou de lignes directrices en matière de santé publique. Les lignes directrices les plus précieuses fournissent une base de pratique fondée sur des données probantes ainsi que des recommandations éclairées par des preuves scientifiques actuelles, de haute qualité et évaluées par les pairs. Pour élaborer ces lignes directrices, les données probantes accessibles doivent être évaluées de manière critique afin que les recommandations soient fondées sur les « meilleures » données probantes. La capacité d'évaluation critique de la recherche est, par conséquent, une compétence essentielle pour les professionnels de la santé œuvrant dans des groupes de travail chargés de l'élaboration de politiques ou de lignes directrices.

Notre expérience auprès des groupes de travail chargés de l'élaboration de lignes directrices sur la prévention et le contrôle des infections nous a permis d'observer que, bien que l'examen des données probantes se soit déroulé en douceur, l'évaluation critique des données probantes a posé de nombreux défis. Trois principaux problèmes ont été cernés. Premièrement, bien que les membres des groupes de travail aient une grande expertise

dans le domaine de la prévention et du contrôle des infections ou d'autres secteurs pertinents au sujet des lignes directrices, ils avaient différents niveaux d'expertise sur le plan des méthodes de recherche et d'évaluation critique. Deuxièmement, les outils d'évaluation critique utilisés à ce moment-là étaient axés principalement sur les études analytiques (comme les essais cliniques) et ne comportaient pas de définitions de termes clés ni d'explications relativement aux critères utilisés dans les études. Par conséquent, l'utilisation de ces outils par les membres des groupes de travail n'a pas permis d'obtenir une méthode uniforme pour évaluer les études analytiques. De plus, les outils n'ont pas fourni des moyens d'évaluer les études descriptives et les analyses documentaires. Troisièmement, les membres des groupes de travail voulaient obtenir des conseils sur la façon de progresser de l'évaluation des études individuelles à la synthèse et à l'évaluation d'un ensemble des données probantes.

Pour répondre à ces questions, un examen des outils d'évaluation critique a été effectué. Nous avons constaté que la plupart des outils existants étaient propres à la conception et comportaient des différences considérables sur le plan de l'intention, des critères évalués et de la création des outils. Un

Affiliations

¹ Memorial University
School of Nursing, St. John's
(Terre-Neuve-et-Labrador)

² Centre de la lutte contre les
maladies transmissibles et les
infections, Agence de la santé
publique du Canada, Ottawa
(Ontario)

*Correspondance : toju.ogunremi@phac-aspc.gc.ca



examen systématique a indiqué que moins de la moitié des outils existants avaient des lignes directrices relatives à l'utilisation de l'outil et à l'interprétation des articles (1). L'utilisation de la méthode GRADE (Grading of Recommendations Assessment, Development and Evaluation) bien connue et des outils Cochrane pour évaluer le risque de biais a été envisagée (2,3). À ce moment-là, les lignes directrices relatives à l'utilisation de ces outils étaient limitées, et les outils étaient axés principalement sur les essais randomisés contrôlés et les essais contrôlés non randomisés. Pour des raisons d'ordre éthique et de faisabilité, les essais cliniques sont rarement accessibles pour de nombreux enjeux en matière de prévention et de contrôle des infections (4,5). Par exemple, il n'y a pas d'essai d'études d'intervention pour évaluer quelles sont les restrictions en matière de pratique, le cas échéant, devraient être imposées aux travailleurs de la santé qui sont infectés par un agent pathogène à diffusion hémotogène. Les membres des groupes de travail sont préoccupés par le fait que s'ils avaient utilisé la méthode GRADE, toutes les données probantes auraient été évaluées comme étant très faibles ou de faible qualité ou degré de certitude, et les recommandations fondées sur ces données probantes auraient pu être interprétées comme étant peu convaincantes, même si elles étaient fondées sur les meilleures données probantes ou sur les seules données probantes disponibles.

L'équipe a décidé d'élaborer sa propre trousse d'outils de l'évaluation critique. Ainsi, un petit groupe de travail a été mis sur pied, dirigé par un épidémiologiste ayant une expertise en matière de recherche, de méthodologie et d'évaluation critique, dans le but de mettre au point des outils permettant d'évaluer de manière critique des études éclairant les recommandations en matière de prévention et de contrôle des infections. Le présent article fournit un aperçu de la trousse d'outils de l'évaluation critique. Le document intégral intitulé *Lignes directrices pour la prévention et le contrôle des infections : Trousse d'outils de l'évaluation critique* est accessible en ligne (6).

Aperçu

À la suite d'un examen des outils d'évaluation critique, des études éclairant les lignes directrices relatives à la prévention et au contrôle des infections qui étaient en cours d'élaboration ont été examinées afin de déterminer les types d'études qui devraient être évalués au moyen de la trousse d'outils de l'évaluation critique. Une ébauche préliminaire de la trousse d'outils a été utilisée par divers groupes de travail chargés de l'élaboration des lignes directrices, et des révisions itératives ont été effectuées au cours d'une période de deux ans. Un essai pilote de la trousse d'outils a ensuite été réalisé, ce qui a mené à la version définitive (6).

La trousse d'outils est mise en place pour aider les évaluateurs dans trois phases principales de l'évaluation critique d'un ensemble de données probantes : une évaluation d'études distinctes; le résumé des résultats des évaluations; et l'évaluation de l'ensemble de données probantes.

Outils pour l'évaluation critique d'études distinctes

La première étape de l'évaluation critique d'une étude distincte vise à déterminer la conception de l'étude; cela peut être étonnamment problématique, étant donné que bon nombre d'études de recherche publiées sont complexes. Un algorithme a été élaboré pour aider à déterminer si une étude était une étude analytique, une étude descriptive ou une analyse documentaire (pour les définitions, voir l'**encadré**). Il est essentiel d'établir tout d'abord la conception de l'étude, car les critères d'évaluation varient selon le type d'étude.

Définitions des types d'études qui peuvent être analysés à l'aide de la trousse d'outils de l'évaluation critique*

Étude analytique : Étude visant à déterminer ou à mesurer les effets d'expositions, d'interventions ou de facteurs de risque précis. Ce modèle fait appel à l'utilisation d'un groupe de comparaison adéquat pour tester les hypothèses épidémiologiques et, par conséquent, tenter de déterminer les associations ou les relations de cause à effet.

Étude descriptive : Étude qui décrit les caractéristiques d'un état de santé par rapport à des facteurs particuliers ou à une exposition d'intérêt. Cette conception fournit souvent les premiers indices importants quant à d'éventuels déterminants de la maladie, et elle est utile pour la formulation d'hypothèses qui peuvent être ensuite mises à l'essai au moyen d'un modèle analytique.

Analyse documentaire : Étude qui analyse les points essentiels d'un ensemble de connaissances publiées. Cela se fait par l'entremise du résumé, de la classification et de la comparaison des études antérieures. À l'exception des méta-analyses, qui, statistiquement, font une deuxième analyse de données regroupées provenant de plusieurs études, ces études constituent des sources secondaires et ne rapportent pas de travaux nouveaux ou expérimentaux.

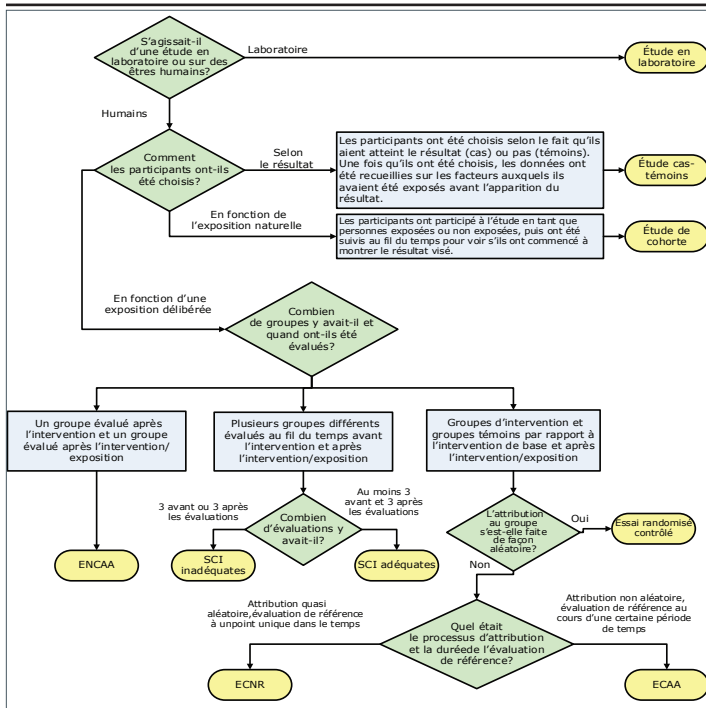
* Agence de la santé publique du Canada. *Lignes directrices pour la prévention et le contrôle des infections : Trousse d'outils de l'évaluation critique* (6)

Des algorithmes distincts ont été mis au point pour les études analytiques, les études descriptives et les analyses documentaires afin d'aider les évaluateurs à déterminer des modèles précis dans ces catégories. L'algorithme ci-dessous, par exemple, permet aux évaluateurs de déterminer quel plan d'étude a été utilisé dans la catégorie des études analytiques (**figure 1**). Il est fondé sur des points de décision clés, comme le nombre de groupes ou la répartition dans le groupe. Les légendes des algorithmes et les outils de soutien, comme le glossaire, fournissent de plus amples détails afin de différencier davantage les modèles d'études, comme le fait de savoir si une étude de cohorte était prospective ou rétrospective.

Des outils d'évaluation critique distincts ont été mis au point pour les études analytiques, les études descriptives et les analyses documentaires, avec des critères pertinents dans chaque outil. Par exemple, un résumé des points abordés dans



Figure 1 : Algorithme permettant de déterminer le type d'étude analytique



Abréviations : ECAA, étude comparative avant-après; ECNR, essai contrôlé non randomisé; ECR, essai contrôlé randomisé; ENCAA, étude non comparative avant-après, STI, séries temporelles interrompues

L'outil d'évaluation critique de l'étude analytique est présenté dans le **tableau 1**. Cet outil est utilisé pour évaluer les essais, les études d'observation et les expériences en laboratoire. Un outil de soutien pour l'évaluation de l'analyse statistique était également fourni et décrit les essais statistiques communs effectués dans le cadre des études épidémiologiques.

Tableau 1 : Aspects évalués dans l'outil d'évaluation critique d'une étude analytique

Aspect	Type d'évaluation
Échantillon et méthodes d'échantillonnage	Représentativité des participants, contrôle du biais de sélection
Validité interne	Contrôle de biais : Erreur de classification ou de renseignements Validité et fiabilité des instruments de collecte de données Justesse de la conservation et du suivi
Contrôle des facteurs de confusion	Comparabilité du groupe témoin et du groupe d'intervention/exposé Justesse du contrôle des grandes variables confusionnelles
Éthique	Justesse de la conduite éthique
Analyse	Justesse et interprétation des tests statistiques Puissance et taille de l'échantillon
Questions relatives au filtrage et à l'applicabilité	Généralisabilité des résultats Faisabilité de l'exécution

L'outil d'évaluation critique d'une étude descriptive évalue les différents aspects de l'échantillonnage, de la collecte de données, de l'analyse statistique et de l'éthique. Il sert à évaluer des études transversales, des enquêtes sur les épidémies, des séries de cas et des rapports de cas.

L'outil d'évaluation critique d'une analyse documentaire permet d'évaluer la méthodologie, les résultats et l'applicabilité des articles-synthèses, des examens systématiques et des méta-analyses.

Après une évaluation d'éléments distincts dans chaque type d'étude, chaque outil d'évaluation critique comporte également des instructions pour tirer des conclusions au sujet de la qualité générale des données probantes tirées d'une étude, en fonction de l'évaluation par point. La qualité est évaluée comme étant élevée, modérée ou faible. Bien qu'un essai contrôlé randomisé (ECR) soit un plan d'étude solide et qu'une enquête est une conception peu efficace, il est possible d'avoir un ECR de mauvaise qualité ou une enquête de grande qualité. Par conséquent, la qualité des données probantes provenant d'une étude est distincte de la force du plan d'étude lors de l'évaluation de la qualité de l'ensemble de données probantes. Une définition de certains termes utilisés pour évaluer les données probantes dans le cadre de l'outil d'évaluation critique est présentée dans le **tableau 2**.

Tableau 2 : Définition des termes utilisés pour évaluer les données probantes

Éléments sommaires évalués	Cote	Critère
Force du plan d'étude Remarque : « x > y » signifie que le plan de x est plus fort que celui d'y.	Forte	Méta-analyse > essai contrôlé randomisé (ECR) > essai contrôlé non randomisé (ECNR) = expérience en laboratoire > étude comparative avant-après (ECAA)*
	Modérée	Cohorte > cas-témoin > séries temporelles interrompues avec collecte de données adéquate > cohorte avec groupe de comparaison non équivalent
	Faible	Étude non comparative avant-après (ENCAA) > séries temporelles interrompues avec collecte de données inadéquate > étude descriptive (transversale > lien épidémiologique > écologique ou corrélationnelle)
Qualité de l'étude	Élevée	Aucun grand obstacle à la validité interne (biais, hasard et confusion ayant fait l'objet d'un contrôle adéquat et ayant été exclus à titre d'explication concurrente des résultats)
	Moyenne	Obstacles mineurs à la validité interne ne compromettant pas gravement la capacité à tirer une conclusion à l'égard de l'estimation de l'effet



Tableau 2 : Définition des termes utilisés pour évaluer les données probantes (suite)

Éléments sommaires évalués	Cote	Critère
Qualité de l'étude (suite)	Faible	Important(s) obstacle(s) à la validité interne compromettant la capacité à tirer une conclusion à l'égard de l'estimation de l'effet
Nombre d'études	Multiples	Quatre études ou plus
	Peu	Trois études ou moins
Cohérence des résultats	Uniformes	Études ayant abouti à des résultats semblables
	Disparates	Variation des résultats, mais tendance globale claire quant à l'effet
	Contradictoires	Résultats variables et aucune tendance globale claire quant à l'effet
Caractère direct des données probantes	Données probantes directes	Proviennent d'études qui portaient précisément sur l'association en question
	Extrapolation	Déduction tirée d'une étude qui portait sur une question différente, mais connexe, ou qui portait sur la même question clé, mais dans des conditions artificielles (p. ex. certaines études de laboratoire)

* La force du plan d'étude est forte s'il y a au moins deux groupes témoins et deux groupes d'intervention. La force est moyenne s'il y a seulement un groupe témoin et un groupe d'intervention

Outils pour résumer les données probantes

La deuxième phase du processus d'évaluation critique consiste à résumer les résultats de l'évaluation critique des différentes études. Les évaluateurs sont invités à créer un modèle de tableau sommaire des données probantes, avec les principaux faits saillants de chaque étude et de leur évaluation. Les études sont énumérées par ordre décroissant de force dans le tableau. Le tableau simplifie la recherche parmi toutes les études formant le corpus de données probantes sur lequel fonder une recommandation et facilite la comparaison entre les participants, la taille d'échantillon, les méthodes, les interventions, l'ampleur et la cohérence des résultats, les mesures des critères d'évaluation et la qualité de chaque étude, tel qu'il est déterminé par l'évaluation critique. Ces tableaux sommaires des données probantes sont passés en revue par le groupe de travail afin d'établir la cote de qualité du corpus de données probantes global et de faciliter la formulation de recommandations fondées sur des données probantes.

Évaluation de la qualité du corpus de données probantes global

La troisième phase du processus d'évaluation critique est l'évaluation de la qualité du corpus de données probantes global. L'évaluation globale dépend de cinq éléments résumés au tableau 2 : la force des plans d'étude, la qualité des études, le nombre d'études, la cohérence des résultats et le caractère

direct des données probantes. Les différentes combinaisons de ces facteurs permettent d'évaluer le corpus de données probantes comme étant solide, modérément solide ou faible, comme le résume le tableau 3.

Tableau 3 : Critères d'évaluation des données probantes sur lesquelles fonder les recommandations

Force des données probantes	Catégories	Critère
Forte	AI	Données probantes directes provenant d'une méta-analyse ou de multiples études à modèle fort de haute qualité, avec résultats cohérents
	All	Données probantes directes provenant de multiples études à modèle fort de qualité moyenne, avec résultats cohérents OU Au moins une étude à modèle fort étayée par de multiples études à modèle modéré de haute qualité, avec résultats cohérents OU Au moins une étude à modèle fort de qualité moyenne étayée par une extrapolation de multiples études à modèle fort de haute qualité, avec résultats cohérents
Modérée	BI	Données probantes directes provenant de multiples études à modèle modéré de haute qualité, avec résultats cohérents OU Extrapolation de multiples études à modèle fort de haute qualité, avec résultats cohérents
	BII	Données probantes provenant de l'association d'études à modèle fort ou modéré de qualité élevée ou moyenne, avec une tendance claire, mais certains résultats incohérents OU Extrapolation de multiples études à modèle fort de qualité moyenne ou d'études à modèle modéré de qualité élevée ou moyenne, avec résultats cohérents OU Une étude à modèle fort étayée par de multiples études à modèle faible de qualité élevée ou moyenne, avec résultats cohérents
Faible	CI	Données probantes directes provenant de multiples études à modèle faible de qualité élevée ou moyenne, avec résultats cohérents OU Extrapolation de l'association d'études à modèle fort ou modéré de qualité élevée ou moyenne, avec résultats incohérents
	CII	Étude de faible qualité, quel que soit le modèle OU Résultats contradictoires, quel que soit le modèle OU Études de série de cas ou exposés de cas OU Opinion d'un expert



Un aspect unique de cette trousse d'outils est que les recommandations ne sont pas classées, mais formulées en fonction du corpus de données évalué. Il y a deux actions : recommandé ou non recommandé. C'est la force des données probantes disponibles qui varie, et non la fermeté de la recommandation. La trousse d'outils souligne toutefois la nécessité de réévaluer les nouvelles données probantes à mesure qu'elles sont obtenues, en particulier lorsque les recommandations sont fondées sur des données probantes faibles.

Essai pilote de la trousse d'outils d'évaluation critique

Sur les 34 personnes ayant manifesté leur intérêt envers l'essai pilote, 17 l'ont terminé. Plusieurs études évaluées par des pairs représentant des études analytiques, des études descriptives et des analyses documentaires ont été retenues. Les mêmes études ont été attribuées à des participants ayant une expertise semblable par rapport au contenu. Chaque participant devait évaluer trois études analytiques, deux études descriptives et une analyse documentaire au moyen de l'outil d'évaluation critique que le participant jugeait approprié. Pour chaque étude évaluée, un outil d'évaluation critique et le formulaire de rétroaction propre à l'outil connexe devaient être remplis. Chaque participant remplissait également un seul formulaire de rétroaction générale. Au total, 101 des 102 évaluations critiques ont été réalisées et retournées, dont 81 formulaires de rétroaction propres à l'outil et 14 formulaires de rétroaction générale.

Pour la plupart (> 85 %), les participants ont trouvé que chaque outil avait un enchaînement logique et une durée acceptable, mais ils ont indiqué avoir encore de la difficulté à mettre en évidence les plans d'étude (**tableau 4**).

Tableau 4 : Rétroaction de l'essai pilote sur la convivialité

Éléments	Outil d'évaluation critique analytique (%) n = 39 sur 51*	Outil d'évaluation critique descriptif (%) n = 28 sur 34*	Outil d'évaluation critique d'une analyse documentaire (%) n = 14 sur 17*
Enchaînement logique	89,7	96,4	100
Durée acceptable	97,4	100	100
Formulation et explications claires	72,2	88,5	76,9
Outil utile dans le processus d'évaluation critique	92,3	85,7	92,9

* Nombre d'évaluations propres à l'outil retournées pour le nombre total d'évaluations critiques réalisées

Selon la vaste majorité des formulaires de rétroaction (86-93 %), les différents outils ont facilité le processus d'évaluation critique. Dans l'évaluation de la constance, toutefois, seules quatre des dix études analytiques prises en compte (40 %) ont été évaluées de la même manière par les participants en ce qui a trait à la qualité générale de l'étude; les six autres études présentaient des différences relevées comme des discordances. Quatre des six études ayant des discordances étaient des études d'observation. Les différences étaient mineures. Aucune des discordances ne concernait une étude ayant été évaluée comme étant de grande qualité et de faible qualité par différents participants. En fonction des commentaires formulés par les participants, la plupart des discordances auraient probablement pu être résolues par des discussions avec les pairs. Les évaluations discordantes n'étaient pas un problème pour les études descriptives et les analyses documentaires. En résumé, l'essai pilote nous a fourni une rétroaction utile sur les différents aspects de la trousse d'outils. Suite à l'essai pilote, des modifications ont été faites pour résoudre les problèmes signalés et, par conséquent, a renforcé la trousse d'outils d'évaluation critique.

Discussion

La *Trousse d'outils de l'évaluation critique des Lignes directrices pour la prévention et le contrôle des infections* a été élaborée en réponse aux besoins des professionnels en matière de prévention des infections amenés à analyser des publications ne comprenant généralement pas de données tirées d'essais cliniques. La trousse a été conçue pour répondre aux besoins cernés en matière de formation en évaluation critique, avec des directives exhaustives et des dictionnaires, de même que des outils qui s'appliquent aux trois types d'études (études analytiques, études descriptives et analyses documentaires). La trousse d'outils procure une méthode permettant de progresser de l'évaluation d'études individuelles à la synthèse et à l'évaluation de la force du corpus de données probantes et à l'attribution d'une classification. Les recommandations sont ensuite formulées en fonction du corpus de données probantes évalué. Ce système de classification a été utilisé par l'Agence de la santé publique du Canada dans l'élaboration des récentes lignes directrices pour la prévention et le contrôle des infections (5),(7). La trousse d'outils a aussi été utilisée pour effectuer une évaluation critique à d'autres fins, pour la résolution d'un problème pratique et comme outil éducatif (8,9).

La trousse d'outils d'évaluation critique a un certain nombre de points forts. Elle est applicable à un large éventail de plans d'étude. Les critères qui sont évalués permettent une évaluation globale de différentes études et facilitent l'évaluation critique d'un corpus de données probantes. Les dictionnaires offrent aux évaluateurs un langage et des critères communs aux fins de discussion et de prise de décisions.

La trousse d'outils d'évaluation critique a un certain nombre de limitations. Les outils ne correspondent pas à tous les plans d'étude (p. ex. études de modélisation) et la trousse d'outils fournit peu de renseignements sur les types de biais. Comme la plupart des outils d'évaluation critique (10,11), la validité et la fiabilité de ces outils n'ont pas été évaluées. Néanmoins, les critères évalués sont ceux indiqués comme étant importants dans



les manuels et les publications (12,13). L'échelle de classification utilisée dans cette trousse d'outils ne permet pas de comparer les données probantes entre différents organismes ou à l'échelle internationale, mais la plupart des évaluateurs n'ont pas besoin d'une telle comparabilité. Il est plus important que des données solides obtiennent une meilleure classification que des données plus faibles et que les évaluateurs justifient leurs conclusions, ce que la trousse leur permet de faire.

Globalement, l'essai pilote a confirmé que la trousse d'outils d'évaluation critique peut aider à la formation à l'évaluation critique et permettre d'accroître le niveau de confort des personnes ayant une expérience limitée. Une évaluation plus approfondie de la trousse permettrait d'évaluer l'efficacité des révisions apportées et d'en tester la validité et la fiabilité.

Il a souvent été demandé en quoi cette trousse diffère de la méthode GRADE puisque toutes deux distinguent les données probantes plus solides de celles qui sont plus faibles et qu'elles utilisent des concepts et une terminologie semblables. Les principales différences entre la méthode GRADE et la trousse d'outils d'évaluation critique sont présentées au **tableau 5**. Les différences clés sont les suivantes : la trousse d'outils d'évaluation critique est axée sur l'évaluation de la qualité des différentes études, donne des instructions détaillées et offre des outils de soutien qui aident les personnes ayant une expérience limitée de l'évaluation critique. En présence d'essais cliniques et d'études d'intervention bien contrôlées, la méthode GRADE et ses outils connexes de Cochrane seraient plus appropriés

Tableau 5 : Comparaison des caractéristiques de la trousse d'outils de l'évaluation critique et de la méthode GRADE

Caractéristiques	Trousse d'outils d'évaluation critique	GRADE
Plans d'étude évalués	Peut servir à tous les types d'études (essais contrôlés randomisés ou non, autres études analytiques, y compris des études d'observation, des études descriptives et des examens systématiques). Des outils sont fournis pour déterminer les plans d'étude.	Met l'accent sur les types de données probantes les plus solides (essais contrôlés randomisés ou non; études d'observation).
Type d'évaluateurs	Personnes ayant moins d'expérience dans la recherche.	Personnes ayant plus d'expérience dans la recherche.
Évaluation d'études individuelles	Des outils sont fournis pour l'évaluation critique des études individuelles et une cote de qualité est attribuée à chaque étude.	Chaque étude est évaluée individuellement, mais les différentes études ne reçoivent pas de cote de qualité.
Évaluation du corpus de données	Le corpus de données global est évalué en fonction des critères fournis.	Le corpus de données global est évalué en fonction des critères fournis.

Tableau 5 : Comparaison des caractéristiques de la trousse d'outils de l'évaluation critique et de la méthode GRADE (suite)

Caractéristiques	Trousse d'outils d'évaluation critique	GRADE
Évaluation et critères	L'évaluation qualitative est réalisée en fonction de la force des plans d'étude, de la qualité des études, du nombre d'études, de la cohérence des résultats et du caractère direct des données probantes. Une cote est attribuée en fonction de l'évaluation.	Une note numérique est calculée en fonction du fait que les données probantes soient tirées d'essais randomisés ou non, du risque de biais, de l'incohérence, du caractère indirect, de l'imprécision et du biais de publication. La cote est traduite en classification.
Classification des données probantes	Les données probantes sont évaluées comme étant de qualité élevée, modérée ou faible.	Les données probantes sont évaluées comme ayant un degré de certitude élevé, modéré, faible ou très faible.
Classification des recommandations	Les recommandations ne sont pas classées; les actions sont recommandées ou non.	Les recommandations sont classées comme étant fortes, faibles ou conditionnelles.
Orientation pour les évaluateurs	Les critères et explications détaillés à utiliser sont fournis dans une seule boîte à outils.	Les critères et explications détaillés à utiliser sont fournis dans divers documents et dans la formation disponible.

Abréviation : GRADE, Grading of Recommendations Assessment, Development and Evaluation

(2,3). Lorsqu'on ne dispose que d'études descriptives, la trousse d'outils d'évaluation critique est très utile.

Conclusion

La *Trousse d'outils de l'évaluation critique des Lignes directrices pour la prévention et le contrôle des infections* a été élaborée en réponse aux besoins en matière de formation sur l'évaluation critique, d'évaluation des données probantes issues d'un vaste éventail de méthodologies de recherche et d'une méthode pour passer de l'évaluation d'études individuelles à la caractérisation de la force d'un corpus de données probantes. Les cliniciens-chercheurs, les responsables des politiques et les étudiants peuvent utiliser ces outils pour l'évaluation critique des études, qu'ils cherchent à élaborer des politiques, à trouver une solution potentielle à un problème pratique ou à critiquer un article dans le cadre d'un club de lecture. La trousse d'outils s'ajoute à l'arsenal d'outils d'évaluation critique actuellement disponibles et est particulièrement utile pour évaluer les données probantes tirées d'un vaste éventail de plans d'étude.



Déclaration des auteurs

D.M. – Conceptualisation, méthodologie, enquête, collecte et conservation des données et rédaction — ébauche originale, examen et révision

Déclaration des auteurs (suite)

T.O. – Conceptualisation, méthodologie, enquête, collecte et conservation des données et rédaction — ébauche originale, examen et révision

K.D. – Conceptualisation, examen et révision, supervision et gestion du projet

Conflit d'intérêt

Aucun.

Collaborateur

Jennifer Kruse, Agence de la santé publique du Canada — Conceptualisation et gestion du projet

Remerciements

Nous tenons à remercier le Groupe de travail d'experts en prévention et en contrôle des infections de l'Agence de la santé publique du Canada pour sa rétroaction dans l'élaboration de la trousse d'outils, Lisa Marie Wasmund, pour la saisie des résultats de l'essai pilote, Katherine Defalco, pour l'examen des données et la révision croisée du contenu et de la terminologie technique dans la version française de la trousse d'outils, Laurie O'Neil, pour l'examen des premières versions de la trousse d'outils et sa rétroaction à leur sujet, Frédéric Bergeron, pour le soutien avec les algorithmes de la trousse, et le Centre de la lutte contre les maladies transmissibles et les infections de l'Agence de la santé publique du Canada, pour l'examen de la trousse d'outils, la rétroaction à ce sujet et l'utilisation continue de la trousse. Nous tenons à remercier D^{re} Patricia Huston, rédactrice en chef, Relevé des maladies transmissibles au Canada, pour son examen approfondi de l'article provisoire et sa rétroaction constructive à son sujet.

Financement

Ce travail a été appuyé par l'Agence de la santé publique du Canada.

Références

1. Katrak P, Bialocerkowski AE, Massy-Westropp N, Kumar VSS, Grimmer KA. A systematic review of the content of critical appraisal tools. *BMC Med Res Methodol.* 2004;4:22. DOI (<http://dx.doi.org/10.1186/1471-2288-4-22>).
2. GRADE Working Group. Criteria for applying or using GRADE. www.gradeworkinggroup.org [Consulté le 25 juillet 2017].
3. Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0. The Cochrane Collaboration, 2011. <http://handbook.cochrane.org>
4. Khan AR, Khan S, Zimmerman V, Baddour LM, Tleyjeh IM. Quality and strength of evidence of the Infectious Diseases Society of America Clinical Practice Guidelines. *Clin Infect Dis* 2010;51(10):1147-56. DOI (<http://dx.doi.org/10.1086/656735>). PubMed (https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=20946067&dopt=Abstract)
5. Agence de la santé publique du Canada. Pratiques de base et précautions additionnelles visant à prévenir la transmission des infections dans les milieux de soins. <https://www.canada.ca/fr/sante-publique/services/maladies-infectieuses/infections-nosocomiales-professionnelles/pratiques-base-precautions-additionnelles-visant-a-prevenir-transmission-infections-milieux-soins.html> [Consulté le 4 décembre 2015].
6. Agence de la santé publique du Canada. Lingés directrices pour la prévention et le contrôle des infections : Trousse d'outils de l'évaluation critique. http://publications.gc.ca/collections/collection_2014/aspc-phac/HP40-119-2014-fra.pdf [Consulté le 4 décembre 2015].
7. Agence de la santé publique du Canada. Pratiques en matière d'hygiène des mains dans les milieux de soins. <https://www.canada.ca/fr/sante-publique/services/maladies-infectieuses/infections-nosocomiales-professionnelles/pratiques-matiere-hygiene-mains-milieux-soins.html> [Consulté le 4 décembre 2015].
8. Ha S, Paquette D, Tarasuk J, Dodds J, Gale-Rowe M, Brooks JI, Kim J, Wong T. A systematic review of HIV testing among Canadian populations. *Can J Public Health.* 2014;105(1):e53-e62. DOI (<http://dx.doi.org/10.17269/cjph.105.4128>). PubMed (https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=24735698&dopt=Abstract).
9. Stevens LK, Ricketts ED, Bruneau JEE. Critical appraisal through a new lens. *Nursing Leadership.* 2014;27(2):10-3. DOI (<http://dx.doi.org/10.12927/cjnl.2014.23843>). PubMed (https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=25073050&dopt=Abstract)
10. Lohr KN. Rating the strength of scientific evidence: relevance for quality improvement programs. *Int J Qual Health Care.* 2004;16(1):9-18. DOI (<http://dx.doi.org/10.1093/intqhc/mzh005>). PubMed (https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=15020556&dopt=Abstract).
11. Crowe M, Sheppard L. A review of critical appraisal tools show they lack rigor: Alternative tool structure is proposed. *J Clin Epidemiol.* 2011;64:79-89. DOI (<http://dx.doi.org/10.1016/j.jclinepi.2010.02.008>). PubMed (https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=21130354&dopt=Abstract).
12. Young JM, Solomon MJ. How to critically appraise an article. *Nat Clin Pract Gastroenterol Hepatol.* 2009;6(2):82-91. DOI (<http://dx.doi.org/10.1038/ncpgasthep1331>). PubMed (https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=19153565&dopt=Abstract).
13. Polit DF and Beck CT. *Nursing Research: Generating and Assessing Evidence for Nursing Practice.* 9th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008. Chapter XX, Literature reviews: Finding and critiquing the evidence p. 94-125.