



Évaluation critique des études d'intervention

Préparé par Donna Ciliska, IA, Ph.D.
Professeure, Université McMaster, et
Directrice scientifique, Centre de collaboration nationale des méthodes et outils



Comment fonctionne ce module d'apprentissage?

Durée totale prévue : environ 5 heures

Objectif : Être en mesure de déterminer si la qualité d'une étude d'intervention est suffisante pour que l'étude soit appliquée à votre propre situation. Pour ce faire, vous devez comprendre les critères d'évaluation critique d'une étude d'intervention et savoir comment les appliquer.

Processus : Dans ce module, on présente un scénario qui vous aidera à comprendre et à appliquer chacun des critères d'évaluation critique. Une fois que vous aurez lu le scénario, vous pourrez répondre aux questions dans l'ordre afin de faire l'évaluation critique et de décider d'utiliser l'étude ou non. (La durée prévue pour chaque partie du scénario est indiquée entre parenthèses.)

Liens : Chaque fois que vous voyez le mot *scénario*, vous pouvez cliquer dessus pour lire le scénario en question. De même, les termes clés sont liés à leur définition dans un glossaire.

Aperçu

1. Scénario (0,25 heure)
2. Qu'est-ce qu'une évaluation critique? À quoi est-ce que cela sert? (0,5 heure)
3. Outils et critères d'évaluation critique pour les études d'intervention/de prévention (0,5 heure)
4. Application des critères d'évaluation critique
 - a) Lire l'article et répondre au questionnaire (1 heure)
 - b) Les résultats sont-ils valables? (1 heure)
 - c) Quels sont les résultats? (1 heure)
 - d) Comment puis-je appliquer les résultats? (0,5 heure)
 - e) Conclusion du scénario (0,25 heure)
5. Pratique facultative d'évaluation
6. Références utiles
7. Glossaire

1. Scénario

(0,5 heure)



L'hiver dernier, vous avez reçu le vaccin antigrippal; cependant, vous aviez l'impression de presque toujours avoir le rhume. Vous guérissiez et n'aviez plus de symptômes pendant une semaine ou deux, puis les symptômes réapparaissaient. Vous avez vu des publicités annonçant des produits à base de ginseng destinés à prévenir et à traiter les rhumes simples. De plus, vos amis et votre famille vous ont demandé si vous pensiez que les produits à base de ginseng étaient efficaces. Vous décidez donc de chercher la réponse dans la documentation liée à la santé.

1. Vous formulez clairement la question PICO :

P atient / P opulation :	adultes en santé
I ntervention :	préparation orale à base de ginseng
C omparaison :	pas de ginseng
O utcome (Issue clinique) :	nombre de rhumes simples / durée des rhumes simples au cours d'une saison

2. Vous effectuez une recherche dans [PubMed](#). Veuillez noter que ce site n'est disponible qu'en anglais. Dans le menu de gauche, vous cliquez sur « Clinical Queries »; une boîte de dialogue apparaît et vous permet de chercher parmi différents types d'études. Vous choisissez « therapy » et sélectionnez « narrow, specific search ». (Pour en savoir plus sur la recherche de données, consultez le module 1 de cette série, qui porte sur le processus décisionnel fondé sur des faits. url) Lorsque vous tapez « ginseng and colds », une étude apparaît (dont la version complète est disponible gratuitement). Vous lisez le résumé et décidez ensuite de lire l'article au complet.

Votre lecture :

Preddy, G.N., Goel, V., Lovlin, R., Donner, A., Stitt, L., Basu, T.K. (2005). Efficacy of an extract of North American ginseng containing poly-furanosyl-pyranosyl-saccharides for preventing upper respiratory tract infections: a randomized controlled trial. *CMAJ*, 173 (9), 1043-1048.
<http://www.cmaj.ca/cgi/content/full/173/9/1043>

Questions :

Prendrez-vous une préparation à base de ginseng cet hiver pour prévenir ou traiter les rhumes simples?

Que direz-vous à votre famille et à vos amis en ce qui concerne l'efficacité du ginseng?

2. Qu'est-ce qu'une évaluation critique? À quoi est-ce que cela sert?

(0,5 heure)

Appliquer le processus décisionnel fondé sur des faits, c'est se servir des meilleures données *disponibles* pour répondre à une question précise. Il se peut que vous ayez de la chance et que vous trouviez un article dont l'évaluation critique a déjà été effectuée par quelqu'un d'autre, par exemple dans le cas d'un résumé publié dans une revue scientifique. Sinon, vous devrez évaluer vous-même les méthodes de l'étude. Ce processus est appelé « évaluation critique ». Il s'agit de juger la qualité des méthodes utilisées et de déterminer si l'étude peut être appliquée à votre propre situation, que cette situation concerne une population, un patient individuel, une politique ou vous-même. Vous devez tenter de répondre à la question suivante :

Les méthodes utilisées pour mener cette étude étaient-elles suffisamment valables pour que je puisse me fier aux résultats?

Il s'agit de la troisième étape du processus décisionnel fondé sur des faits, qui en comporte sept au total :

1. *Poser la question.* Comment formuler ma question?
2. *Acquérir.* Comment trouver les meilleures preuves en cinq minutes ou moins?
3. *Évaluer.* Comment décider si cette étude particulière est assez valable pour que je m'en serve?
4. *Intégrer.* Comment choisir laquelle des nombreuses études utiliser?
5. *Adapter.* Comment utiliser les renseignements du n° 4 pour prendre des décisions et élaborer des orientations en matière de politiques?
6. *Appliquer.* Comment élaborer le plan de mise en œuvre?
7. *Analyser.* Comment savoir si le plan a fonctionné?

(Note : Pour avoir un aperçu de chacune des étapes ci-dessus, voir le module 1 de cette série, qui porte sur le processus décisionnel fondé sur des faits. [url](#))

Que devez-vous faire lorsque, comme cela arrive souvent, vous trouvez une étude selon laquelle un médicament ou un traitement est efficace, et une autre selon laquelle ce médicament ou ce traitement est inefficace, voire dangereux? La vitamine E prévient-elle les maladies du cœur ou augmente-t-elle le risque cardiovasculaire? La vitamine C permet-elle de prévenir le rhume ou ne produit-elle que de l'urine plus coûteuse? Les asticots contribuent-ils au débridement des ulcères veineux et à la guérison subséquente des lésions ou ne font-ils qu'accroître l'anxiété et le dégoût des patients et des travailleurs de la santé?

Comment savoir à quelles études vous fier? Le meilleur moyen consiste à évaluer les méthodes utilisées dans chaque étude afin de déterminer quelles études ont été faites selon les meilleures méthodes, ce qui permet de contrôler les variables confusionnelles ou les biais éventuels. Les études en question constituent alors les meilleures preuves disponibles et vous serviront, en partie, de base pour prendre vos décisions en matière de pratiques et de politiques. On nous a appris à être des consommateurs avertis et à analyser

de façon critique les affirmations que l'on nous présente dans la publicité. Nous devons exercer cette même capacité d'analyse critique envers les résultats des études sur les soins de santé.

Il existe des critères de qualité essentiels pour évaluer tous les types d'études que vous trouverez. Lorsque vous effectuez votre recherche, vous devez toujours commencer par les examens systématiques, car ils constituent un ensemble de données sur un sujet particulier ayant déjà fait l'objet d'une évaluation critique. (Voir le module 1 sur le processus décisionnel fondé sur des faits. [url](#)) Toutefois, afin de comprendre le processus d'évaluation critique des examens systématiques (module 3 de cette série), vous devez comprendre ce processus lorsqu'il est appliqué à une étude individuelle. Le présent module d'apprentissage décrit en détail le processus d'évaluation critique pour les études individuelles d'intervention (traitement) ou de prévention.

▲ Attention! Les novices de l'évaluation critique rejettent parfois des études relativement bien faites sous prétexte qu'elles ne sont pas parfaites. Il n'existe **aucune** étude parfaite. À mesure que vous vous familiarisez avec le processus, vous verrez que certains critères concernent des points cruciaux; s'ils n'étaient pas respectés, ils constitueraient donc des « défauts fatals » qui vous feraient rejeter l'étude. Cependant, certains autres critères sont moins importants et, même si l'étude ne les respectait pas, vous pourriez toujours envisager de mettre en œuvre l'intervention.

3. Outils et critères d'évaluation critique pour les études d'intervention/de prévention

(0,5 heure)

La plupart des outils actuels d'évaluation critique pour la recherche quantitative ont été élaborés selon les critères clés établis par l'Evidence-Based Medicine Working Group. Le *Journal de l'Association médicale canadienne* a d'abord publié la série « Readers Guides » à partir de 1981 puis, la série a été revue et augmentée, et elle a été publiée dans le *Journal of the American Medical Association (JAMA)* sous le titre « Users' Guides » de 1993 à 2000 et, enfin, on en a fait un livre (Guyatt et Rennie, 2002). L'Evidence-Based Medicine Working Group a produit quelque 25 outils pour différents types de questions cliniques et de méthodologies de recherche (par exemple, les études portant sur les traitements, les causes ou le diagnostic d'une maladie, les examens systématiques, les analyses économiques, les guides de prédiction clinique, les directives cliniques et la recherche sur les services de santé).

Ce module d'apprentissage porte sur l'évaluation critique des études d'intervention (aussi appelées études sur les traitements ou études thérapeutiques) et de prévention. L'héparine de faible poids moléculaire prévient-elle la thrombose veineuse profonde? Le contrôle étroit de la glycémie des patients diabétiques peut-il prévenir les complications cardiovasculaires? L'aromathérapie favorise-t-elle la relaxation? Le traitement hormonal substitutif réduit-il les « bouffées de chaleur » de la ménopause? Un programme intensif de sensibilisation peut-il diminuer le taux de grossesse chez les adolescentes? Toutes ces questions constituent des questions de recherche pouvant faire l'objet d'études d'intervention ou de prévention.

Les critères de base pour l'évaluation critique d'une étude d'intervention sont les suivants :

Encadré 1. ÉVALUATION CRITIQUE D'UNE ÉTUDE D'INTERVENTION OU DE PRÉVENTION

Les résultats sont-ils valables?

Les participants ont-ils été randomisés?

La randomisation a-t-elle été dissimulée?

Les participants ont-ils été analysés dans les groupes dans lesquels ils avaient été placés lors de la randomisation?

Les participants de chaque groupe étaient-ils semblables en ce qui a trait aux variables prévisionnelles connues?

Les participants savaient-ils dans quel groupe ils avaient été placés?

Les cliniciens connaissaient-ils la répartition des participants dans les groupes?

Les personnes qui évaluaient les résultats connaissaient-elles la répartition des participants dans les groupes?

A-t-on effectué un suivi complet?

Quels sont les résultats?

Quelle a été la portée du traitement?

Quelle était la précision de l'estimation de l'effet du traitement?

Comment puis-je appliquer les résultats?

Les participants de l'étude étaient-ils semblables à ceux de ma propre situation?

Les résultats importants sur le plan clinique (préjudices et bienfaits) ont-ils tous été pris en compte?

Traduit et adapté de Guyatt et Rennie, 2002

Vous avez sans doute remarqué que les questions d'évaluation critique parlent de randomisation et portent sur un essai contrôlé randomisé (ECR). En effet, lorsque c'est possible, un ECR est le type d'étude le plus indiqué pour répondre à des questions liées à l'intervention, car la répartition aléatoire des participants permet de distribuer équitablement les déterminants des résultats, connus et inconnus, entre les groupes. Ainsi, vous pouvez être plus sûr de la probabilité que les différences éventuelles entre les résultats des deux groupes soient dues à l'intervention elle-même et non aux différences sous-jacentes entre les groupes. Autrement dit, les essais randomisés sont le type d'étude qui permet le plus grand contrôle des variables confusionnelles et des biais. Cependant, il n'est pas toujours possible d'utiliser des essais randomisés pour répondre à des questions d'intervention ou de prévention. Par exemple, du point de vue de l'éthique, il n'est pas possible (ni faisable) de randomiser des femmes pour qu'elles allaitent leur nouveau-né ou leur donnent le biberon afin d'évaluer l'effet du lait maternel sur la prévention de

l'asthme lorsqu'on en nourrit exclusivement un nouveau-né. On utiliserait plutôt un plan avant-après non randomisé comportant deux groupes. Dans le cas d'essais non randomisés, les critères d'évaluation critique demeurent utiles, mais il est important de comprendre qu'il existe une plus grande possibilité que des différences sous-jacentes entre les groupes soient responsables des différences dans les résultats. (Pour plus d'information à ce sujet, voir le module 1, « Le processus décisionnel fondé sur des faits ». url) La section 5, « Pratique facultative d'évaluation », contient un exemple de question pour laquelle aucun ECR n'est disponible et la meilleure source de données est un plan avant-après comportant deux groupes (étude non randomisée).

Dans la section 4, on présente le scénario qui permet d'illustrer et d'appliquer les critères.



RESSOURCE RECOMMANDÉE

Parce qu'il est facile d'y accéder et de les comprendre, nous vous suggérons d'utiliser les critères d'évaluation critique du Critical Appraisal Skills Programme (CASP) de la Public Health Resources Unit du Royaume-Uni. Cet organisme a produit une série d'outils fondés sur les « Users' Guides » (Guyatt et Rennie, 2002). Leurs outils sont avantageux parce que l'explication des critères y est intégrée et que vous pouvez les consulter gratuitement en ligne pour votre usage personnel. Bien qu'ils aient été élaborés par un organisme de santé publique, ils peuvent être utilisés pour d'autres sujets.

(<http://www.phru.nhs.uk/Pages/PHD/resources.htm>)

Source :

Guyatt, G. & Rennie, D. (Eds) (2002). *Users' Guides to the Medical Literature: A manual for Evidence-Based Clinical Practice*. American Medical Association.

4. Application des critères d'évaluation critique

a) Lire l'article et répondre au questionnaire.

(1 heure)

Retournez au scénario de la section 1, qui traite de l'efficacité du ginseng pour la prévention ou le traitement du rhume.

Vous trouvez cet article :

Preddy, G.N., Goel, V., Lovlin, R., Donner, A., Stitt, L., Basu, T.K. (2005). Efficacy of an extract of North American ginseng containing poly-furanosyl-pyranosyl-saccharides for preventing upper respiratory tract infections: a randomized controlled trial. *CMAJ*, 173 (9), 1043-1048. <http://www.cmaj.ca/cgi/content/full/173/9/1043>



Maintenant, à vous d'essayer de répondre aux questions! Vous utiliserez cet article et répondrez, dans l'ordre, à chaque question du *Formulaire d'évaluation critique pour les interventions*.

NOTE : Pour faire cette partie du module, nous vous suggérons d'imprimer la version PDF de l'article ou de l'afficher sur votre écran afin de pouvoir vous y référer au fur et à mesure que vous réalisez l'activité qui suit.

1. Veuillez lire l'article en entier.
2. Répondez aux questions d'évaluation critique du guide ci-dessous.

FORMULAIRE D'ÉVALUATION CRITIQUE POUR LES INTERVENTIONS

Référence :

Guide	Commentaires
I) Les résultats sont-ils valables?	
Les participants ont-ils été randomisés?	
La randomisation a-t-elle été dissimulée?	
Les participants ont-ils été analysés dans les groupes dans lesquels ils avaient été placés lors de la randomisation?	
Les participants du groupe de traitement et ceux du groupe témoin étaient-ils semblables en ce qui a trait aux variables prévisionnelles connues?	
Les participants savaient-ils dans quel groupe ils avaient été placés?	
Les cliniciens connaissaient-ils la répartition des participants dans les groupes?	
Les personnes qui évaluaient les résultats connaissaient-elles la répartition des participants dans les groupes?	
A-t-on effectué un suivi complet?	
II. Quels sont les résultats?	
Quelle a été la portée du traitement?	
Quelle était la précision de l'estimation de l'effet du traitement?	
III. Comment puis-je appliquer les résultats?	
Les participants de l'étude étaient-ils semblables à ceux de ma propre situation?	
Les résultats importants sur le plan clinique (préjudices et bienfaits) ont-ils tous été pris en compte?	

4. b) Les résultats sont-ils valables?

(1 heure)

i) Les participants ont-ils été randomisés?

Il est important d'effectuer une randomisation pour s'assurer que les groupes sont semblables sur tous les points qui peuvent influencer les résultats (par exemple, l'âge, le sexe ou le statut socioéconomique). La seule différence observée entre les groupes doit être les résultats. Cela permet de réduire les biais possibles. Parmi les types d'études individuelles, l'essai randomisé est considéré comme la source la plus fiable de preuves. Il faut toutefois se rappeler que ce ne sont pas toutes les questions qui peuvent faire l'objet d'un ECR, que ce soit sur le plan éthique ou pratique. Bien sûr, un examen systématique d'un certain nombre d'essais constitue une source de preuves encore plus fiable. (Voir le module 1 de cette série, « Le processus décisionnel fondé sur des faits ».)

Il existe différentes façons d'effectuer la randomisation. Il est important de vous assurer qu'une vraie randomisation a été faite dans le cadre de l'étude, et non une « quasi-randomisation », par exemple en fonction des jours de la semaine ou du mois de naissance. Une vraie randomisation se fait au moyen d'un tableau de nombres aléatoires ou d'un générateur de nombres aléatoires informatisé.

Q. Les participants ont-ils été randomisés?

R. OUI. À la page 1044, à la 1^{re} ligne sous « Methods », on dit que l'étude consistait en un essai randomisé.

De plus, à la page 1045, dans le 1^{er} paragraphe complet, les auteurs révèlent qu'ils ont utilisé un modèle de randomisation informatisé.

ii) La randomisation a-t-elle été dissimulée?

Pourquoi est-ce important? Parce que d'autres études ont démontré que si la personne qui recrute les participants connaît la séquence de répartition (dans quel groupe sera placé le prochain participant), elle risque de faire un choix conscient ou inconscient (c'est-à-dire de substituer une enveloppe à une autre) si elle croit qu'il serait plus avantageux pour le participant de faire partie du groupe d'intervention. Afin de garantir la dissimulation de la randomisation, on peut utiliser différentes stratégies, par exemple des enveloppes opaques numérotées de façon séquentielle ou un centre d'appels qui répartit les participants recrutés dans les groupes.

Q. La randomisation a-t-elle été dissimulée?

R. OUI.

À la page 1045, dans le 1^{er} paragraphe complet, les auteurs affirment avoir utilisé des enveloppes opaques numérotées et scellées. Cela signifie que la séquence ne pouvait pas être modifiée et que la personne effectuant le recrutement ne pouvait pas voir à travers l'enveloppe et savoir quel était le groupe dans lequel le prochain participant serait placé.

iii) Les participants ont-ils été analysés dans les groupes dans lesquels ils avaient été placés lors de la randomisation?

Le milieu de la recherche recèle des histoires d'horreur; par exemple, des participants qui avaient abandonné un groupe d'intervention et ont été « déplacés » dans le groupe témoin par les chercheurs, ou, à l'inverse, des participants du groupe témoin qui, on ne sait trop comment, ont reçu l'intervention à l'extérieur de l'étude, puis ont été déplacés dans le groupe d'intervention. Comme vous vous en doutez certainement, l'exclusion des abandons dans une étude portant sur le renoncement au tabac ou la perte de poids (des études dans lesquelles le taux d'abandons peut atteindre 50 à 60 %) peut faire paraître l'intervention plus efficace qu'elle ne l'est réellement.

Ce critère sert à garantir que les participants seront analysés dans le groupe dans lequel ils ont été placés lors de la randomisation, même s'ils cessent le traitement. Les chercheurs appellent cette façon de fonctionner l'analyse « en intention de traiter ». Comment tiennent-ils compte des abandons? Pour mesurer le résultat final, ils remplacent les données manquantes sur les personnes ayant abandonné l'intervention par leurs données de référence ou par la « dernière observation reportée » concernant ces personnes.

Q. Les participants ont-ils été analysés dans les groupes dans lesquels ils avaient été placés lors de la randomisation?

R. OUI.

Le schéma (*Fig. 1*) de la page 1044 montre que 149 personnes du groupe témoin et 130 personnes du groupe recevant du ginseng ont commencé l'intervention, et que toutes ont été incluses dans l'analyse (encadrés au bas du schéma). En outre, les chercheurs indiquent qu'ils ont effectué une analyse en intention de traiter et qu'ils se sont servi de la dernière observation reportée (page 1045, colonne de droite, 4^e paragraphe).

iv) Les participants des deux groupes étaient-ils semblables en ce qui a trait aux variables prévisionnelles connues?

Nous voulons savoir si, avant le début de l'intervention, il existait des différences entre les groupes qui pourraient avoir causé des différences dans les résultats. Normalement, la randomisation permet de garantir une répartition égale des caractéristiques des participants. Les chercheurs vérifient si la randomisation était appropriée en examinant les caractéristiques présentes au début de l'étude qui, selon eux, pourraient être liées aux résultats. Certains déséquilibres résultent d'un échantillon trop petit tandis que d'autres sont dus au hasard. Si l'échantillon était adéquat, l'analyse peut tenir compte de toutes les différences observées dans le groupe. Les chercheurs fournissent souvent les résultats « ajustés », qui tiennent compte des différences présentes au départ, et les résultats « non ajustés ».

Q. *Les participants des deux groupes étaient-ils semblables en ce qui a trait aux variables prévisionnelles connues?*

R. OUI.

Le tableau 1 contient des données sur l'âge, le sexe, le statut de fumeur ou de non-fumeur et le nombre de rhumes par année des participants, ainsi que le nombre de sujets ayant trois rhumes ou plus par année. De plus, dans la section « Results », les chercheurs affirment qu'il n'y avait pas de différences statistiquement significatives au départ (page 1046, colonne de droite, 1^{er} paragraphe). Auraient-ils dû tenir compte d'autres variables? Ils ont exclu les personnes atteintes d'une maladie chronique ou aiguë et celles prenant des médicaments; les sujets constituaient donc une population normale et en santé.

v) *Les participants savaient-ils dans quel groupe ils avaient été placés?*

L'*insu* (aussi appelé *masquage*) est le fait que différentes personnes ignorent dans quel groupe (intervention ou témoin) sont placés les participants d'une étude. Dans les rapports de recherche, on utilise parfois les expressions à *simple*, à *double* ou à *triple insu*, mais on considère maintenant qu'il est important de préciser à l'*insu* de *qui* la répartition des participants a été effectuée.

Ce critère concerne le fait que la répartition ait été faite à l'*insu* des *participants*. En effet, si ceux-ci savent de quel groupe ils font partie, ils peuvent, consciemment ou non, prêter plus attention aux aspects positifs ou négatifs de l'intervention. Dans les essais sur des médicaments, les placebos ont la même apparence que le traitement actif; il est donc difficile de les différencier. Cependant, il est beaucoup plus difficile de mener des études dans le domaine de la pédagogie curative ou de la psychothérapie à l'*insu* des participants : ils savent s'ils ont fait des exercices ou regardé des vidéos!

Q. *Les participants savaient-ils dans quel groupe ils avaient été placés?*

R. NON.

À la page 1044, il est indiqué qu'il s'agissait d'une étude à double insu. Les participants devaient prendre les préparations, qui n'étaient pas identifiées, selon les instructions. Les préparations (des capsules) étaient toutes identiques (page 1045, colonne de gauche, fin du dernier paragraphe complet). De plus, après l'étude, on a demandé aux participants s'ils pensaient avoir pris le ginseng ou le placebo (page 1045, colonne de droite, 1^{er} paragraphe); 69,8 % de ceux qui avaient pris le ginseng et 77,3 % de ceux qui avaient pris le placebo croyaient qu'on leur avait donné la préparation à base de ginseng.

vi) *Les cliniciens connaissaient-ils la répartition des participants dans les groupes?*

Ce critère consiste à déterminer si les cliniciens prenant part à l'étude savent dans quel groupe se trouvent leurs patients. S'ils le savent, ils pourraient, consciemment ou non, modifier leur plan de traitement, prodiguer plus de soins à leurs patients ou être plus vigilants quant aux résultats positifs ou négatifs. Si le clinicien est celui qui applique l'intervention, il est impossible que celle-ci se déroule à son insu; toutefois, il peut être

tenu dans l'ignorance en ce qui concerne la question de recherche ou, du moins, la comparaison effectuée.

Q. Les cliniciens connaissaient-ils la répartition des participants dans les groupes?

R. NON.

Le médecin menant l'étude ne connaissait pas la répartition des participants. Lorsque les symptômes indiquaient des complications secondaires, le médecin de l'étude recommandait un suivi par le médecin de famille, mais ce dernier n'avait aucun moyen de savoir si son patient prenait la préparation à base de ginseng ou le placebo (page 1045, colonne de gauche, dernière ligne).

vii) Les personnes qui évaluaient les résultats connaissaient-elles la répartition des participants dans les groupes?

Comme pour le critère v) ci-dessus, ce critère consiste à déterminer si les personnes qui mesuraient les résultats savaient de quel groupe faisaient partie les participants qu'elles évaluaient. Il y a plus de chances que l'évaluation soit faussée si la personne qui effectue les mesures (par exemple, prendre la tension artérielle) sait de quel groupe fait partie le participant et a une opinion sur l'efficacité probable de l'intervention. Ce biais possible est éliminé lorsque les tests sont effectués au moyen d'équipement de laboratoire ou d'outils informatisés, par exemple le recours aux échantillons de sang pour l'hémoglobine glycatée.

Q. Les personnes qui évaluaient les résultats connaissaient-elles la répartition des participants dans les groupes?

R. OUI.

Dans le cas présent, ce sont les participants qui ont fait la majeure partie de l'évaluation des résultats, au moyen de journaux dans lesquels ils inscrivait leurs symptômes (page 1045, colonne de gauche, dernier paragraphe). On vérifiait également leur observance en pesant les bouteilles de médicaments retournées. L'article révèle que même les personnes analysant les données ne connaissaient pas la répartition des groupes (page 1045, colonne de droite, 2^e paragraphe complet).

viii) A-t-on effectué un suivi complet?

Ce critère comporte deux sous-questions, et les réponses ne dépendent pas d'une norme générale, mais plutôt de la question de recherche. Premièrement, les participants ont-ils été suivis assez longtemps pour que les résultats du traitement puissent se manifester? Par exemple, si l'on étudie l'effet de la prise de vitamine C durant l'adolescence sur le taux de cancer du côlon, on devrait suivre les participants pendant plus de 25 ans afin de déterminer l'efficacité du traitement, car ce cancer est la plupart du temps diagnostiqué à un âge adulte moyen ou avancé. Cependant, on n'aurait besoin d'attendre que quelques heures ou quelques jours pour vérifier l'efficacité d'un extrait d'Aloe vera sur la douleur cutanée liée à un coup de soleil.

Deuxièmement, combien de participants ont abandonné l'étude avant la fin (ou, inversement, combien l'ont poursuivie)? Qu'est-il arrivé aux participants « perdus », et en quoi leurs résultats pourraient-ils différer de ceux des participants ayant poursuivi l'étude? Par exemple, dans le cas de personnes ayant abandonné une intervention visant la perte de poids ou le renoncement au tabac, nous pouvons supposer que le traitement a échoué.

Selon certaines personnes, pour qu'une étude soit considérée comme rigoureuse, elle doit présenter un taux d'abandon inférieur à 20 %. Une fois de plus, cela dépend de la nature du problème et des participants de l'étude. Par exemple, un taux d'abandon de 35 % serait exceptionnellement bas pour une étude à laquelle participeraient des jeunes de la rue.

Q. A-t-on effectué un suivi complet?

R. OUI.

L'étude s'est déroulée à Edmonton, en Alberta, à partir du début de la saison de la grippe (novembre), pendant quatre mois. Cette période était suffisante pour évaluer l'effet de l'intervention (le ginseng) sur les résultats (infection des voies respiratoires supérieures).

Le schéma de la page 1044 (*Fig. 1*) montre, pour toute la période de l'étude, l'entrée et la sortie des participants ainsi que les causes d'abandon. Dans cette étude, le taux de rétention était de 88 % (149/170) dans le groupe témoin et de 85 % (130/153) dans le groupe d'intervention. On pourrait aussi dire que ces groupes avaient des taux d'abandon respectifs de 12 % et de 15 %. Les auteurs ont indiqué les raisons pour lesquelles les personnes n'ont pas participé à l'étude ou l'ont abandonnée, lorsqu'elles étaient connues. Ce degré de suivi est très acceptable.

Les auteurs vont plus loin et examinent les différences possibles entre les données de départ sur les personnes ayant commencé l'intervention et celles ne l'ayant pas fait (tableau 1, page 1045). Puis, à la page 1046, dans le premier paragraphe complet de la colonne de droite, les auteurs mentionnent qu'il n'y avait pas de différences significatives dans les caractéristiques de départ des participants.

4. c) Quels sont les résultats?

(1 heure)

En répondant aux questions de 4 b), vous vous faites une idée des méthodes utilisées dans l'étude et de la probabilité que les résultats soient valables. Si vous croyez qu'ils le sont, vous poursuivez votre démarche : vous examinez les résultats obtenus et déterminez leur importance.

i) Quelle a été la portée du traitement?

Les bienfaits (et les préjudices) de toute intervention peuvent être mesurés grâce à différents résultats. Les résultats peuvent être dichotomiques (soit/soit), par exemple, mort ou vivant, infection ou absence d'infection, guéri ou non guéri; ou continus, par

exemple, le nombre d'éternuements par jour, la durée du séjour à l'hôpital, le rythme respiratoire ou la glycémie à jeun.

Lors de l'analyse des résultats d'études utilisant des résultats dichotomiques, on peut mesurer les comparaisons au moyen de taux (p. ex. 49 % d'ulcères guéris dans le groupe d'intervention contre 25 % dans le groupe témoin). Ces taux permettent ensuite d'exprimer d'autres valeurs, par exemple la différence de risque absolu, l'augmentation de l'avantage relatif (ou l'inverse, c'est-à-dire la réduction du risque relatif) ou le nombre nécessaire à traiter (ou le nombre nécessaire pour nuire). Dans les études utilisant des résultats continus (nombre de rhumes par saison, numération des spermatozoïdes), on analyse les différences entre les moyennes.

Vous devez vérifier si la différence entre les groupes était statistiquement significative. Le véritable effet d'un traitement ne peut être connu; les chercheurs font une *estimation* de l'effet. Les intervalles de confiance (IC) sont une façon de connaître le degré d'incertitude entourant une estimation. Un intervalle de confiance de 95 %, par exemple, signifie que nous sommes certains à 95 % que la vraie valeur de l'effet se trouve dans les limites données. Si, pour un IC de 95 % concernant un rapport de cotes ou un risque relatif, les limites incluent 1, il n'y a pas de différence statistiquement significative entre les groupes de traitement. De même, si l'IC de 95 % pour une différence moyenne inclut 0, il n'y a pas de différence statistiquement significative.

La signification statistique peut aussi être exprimée par la valeur de P. Par convention, on considère que si la valeur de P est inférieure à 0,05, la différence est statistiquement significative. Autrement dit, nous sommes prêts à accepter que la probabilité que le résultat soit uniquement dû au hasard soit inférieure à 1 sur 20.

Q. Quelle a été la portée du traitement?

R. Le résultat principal était le nombre de rhumes signalés (et vérifiés au moyen des critères de Jackson) par sujet. Le nombre moyen de rhumes était 0,68 dans le groupe ayant pris du ginseng et 0,93 dans le groupe témoin (la différence moyenne était de 0,25 rhume par personne; IC de 95 %, entre 0,04 et 0,45) (page 1046, tableau 2, et colonne de droite, 3^e paragraphe complet). Les IC n'incluent pas 0; vous pouvez donc en conclure que le résultat est statistiquement significatif. Dans le texte, les auteurs donnent également la valeur de P correspondante ($P = 0,017$), qui est inférieure à 0,05, donc statistiquement significative.

Les auteurs ont aussi examiné le nombre de personnes qui ont eu "un rhume" ou "deux rhumes ou plus" pendant la saison. Examinons celles qui en ont eu deux ou plus; à la page 1046, le tableau 2 montre que 10 % des personnes ayant pris du ginseng ont signalé deux rhumes ou plus, contre 22,8 % des personnes ayant pris le placebo. Cela représente une réduction du risque absolu de 12,8 % (IC de 95 %, entre 4,3 et 21,3). Vous savez que cette différence est statistiquement significative, parce que l'IC n'inclut pas 1.

ii) Quelle était la précision de l'estimation de l'effet du traitement?

La précision peut être déterminée uniquement au moyen de l'intervalle de confiance (IC) (voir 4 c) i). Si l'IC est large, l'estimation de l'effet véritable manque de

précision et nous ne sommes pas certains de l'effet du traitement. Plus l'IC est étroit, plus l'estimation est précise, et plus nous pouvons être certains des résultats. L'utilisation d'échantillons plus grands produit des résultats plus précis; vous devez donc vous méfier des petits échantillons et des grands intervalles de confiance (c'est-à-dire ne pas "avoir confiance" aux résultats qui y sont liés).

Enfin, vous devez déterminer si le résultat qui est significatif sur le plan statistique l'est aussi sur le plan clinique (ou personnel). Par exemple, une perte de poids de 4 kg peut être statistiquement significative, mais elle n'est pas cliniquement ou personnellement significative pour des patients souffrant d'obésité morbide. Vous pouvez également utiliser la plus petite valeur de résultat possible (limite inférieure de l'intervalle de confiance) pour répondre à la question suivante : si le traitement produisait un effet aussi petit, est-ce que cela vaudrait toujours la peine de l'appliquer?

Q. *Quelle était la précision de l'estimation de l'effet du traitement?*

R. Comme nous l'avons mentionné dans la section précédente, le résultat quant au nombre de rhumes donnait une différence moyenne de 0,25 rhume par personne (IC de 95 %, entre 0,04 et 0,45). L'étendue de l'intervalle de confiance montre que, si les résultats étaient portés aux extrêmes, on pourrait obtenir une moyenne d'aussi peu que 0,04 rhume de moins par année, ou à l'inverse, d'environ un demi-rhume de moins par année. Ces résultats semblent d'une précision moyenne. Ils sont significatifs sur le plan statistique, mais le sont-ils sur le plan clinique? Si le traitement permettait d'obtenir une réduction moyenne aussi importante qu'un demi-rhume par année (ou un rhume chaque deux ans), cela pourrait constituer un avantage clinique en termes de qualité de vie, de réduction des jours d'absence au travail ou à l'école, ou de productivité.

Si l'on considère le nombre de récurrences du rhume, une réduction de 12,8 % (IC de 95 %, entre 4,3 et 21,3) semble également être d'une précision moyenne. Même si le traitement avait le plus petit effet possible, une réduction de 4,3 % des récurrences du rhume pourrait valoir la peine (être cliniquement significative).

Référence

Lipman, M.M. (2008). No safety in numbers. *Consumer Reports on Health*, June, 11. Presents a discussion geared to non-statisticians of 'absolute risk reduction', 'number needed to treat', and 'number needed to harm'.

4 d) Comment puis-je appliquer les résultats?

(0,5 heure)

i) Les participants de l'étude étaient-ils semblables à ceux de ma propre situation?

Vous devez déterminer la possibilité de généraliser les résultats des participants de l'étude pour les appliquer à vos propres patients ou clients ou à votre propre situation. Par convention entre les revues médicales importantes, les participants de l'étude sont habituellement décrits dans le premier tableau du rapport d'étude. Vous devez tenir compte, par exemple, des différences d'âge, de répartition des sexes, de statut socioéconomique, d'intensité de la maladie et de cas de comorbidité, car ces facteurs pourraient faire changer les résultats que vous obtiendriez dans votre propre situation. Les participants seront rarement exactement pareils à ceux de votre situation; vous devez donc vérifier s'il y a des raisons pour lesquelles vous ne devriez *pas* appliquer les résultats de l'étude.

Vous devez également évaluer la faisabilité du traitement dans votre situation. Pour ce faire, il importe de comparer les systèmes de santé, les coûts estimés de prestation du traitement, les compétences requises pour effectuer l'intervention, la disponibilité de l'équipement et du personnel spécialisé, ainsi que la probabilité que vos patients acceptent le traitement.

Q. Les participants de l'étude étaient-ils semblables à ceux de ma propre situation?

R. Les participants de l'étude étaient des adultes de la population générale, en santé, sans maladie chronique physique ou mentale, et ne prenant pas de médicaments. L'âge moyen était de 43 ans, il y avait légèrement plus de femmes que d'hommes, et il y avait légèrement moins de fumeurs que dans la population générale (page 1045, tableau 1). Les participants habitaient dans la région d'Edmonton, où l'on peut s'attendre à ce que plus de gens aient des rhumes en raison de l'effet asséchant du climat et des systèmes de chauffage sur les muqueuses. Cependant, les caractéristiques des participants sont suffisamment semblables à celles de la population canadienne générale pour que vous puissiez appliquer les résultats de l'étude à cette dernière.

ii) Les résultats importants sur le plan clinique (préjudices et bienfaits) ont-ils tous été pris en compte?

Les chercheurs évaluent plusieurs résultats différents pour vérifier les effets d'un traitement. Ils doivent aussi vérifier si les preuves démontrent que le traitement peut causer des préjudices, bien que l'échantillon utilisé dans le cadre d'un essai puisse être trop petit. Toutefois, il est important de savoir, par exemple, qu'un médicament à l'étude permet de réduire le niveau de lipides sanguins, mais que le taux de mortalité est tout de même plus élevé dans le groupe d'intervention. On doit aussi tenir compte du système de santé et des dépenses liées au traitement; une analyse économique peut être nécessaire pour établir le rapport coûts-avantages.

Q. Les résultats importants sur le plan clinique (préjudices et bienfaits) ont-ils tous été pris en compte?

R. En plus du nombre total de rhumes et du taux de récurrence, les chercheurs ont examiné les symptômes et les effets secondaires. Les scores pour les symptômes du rhume étaient, sur le plan statistique, significativement plus bas dans le groupe prenant du ginseng que dans celui prenant un placebo (page 1046, tableau 3). En outre, les effets secondaires signalés par les participants étaient très semblables, et les chercheurs n'ont noté aucune différence statistiquement significative entre les deux groupes (page 1046, colonne de droite, 5^e paragraphe complet, et page 1047, tableau 4).

4 e) Conclusion du scénario

Au début du scénario, on vous a demandé :

Prendrez-vous une préparation à base de ginseng cet hiver pour prévenir ou traiter les rhumes simples? Que direz-vous à votre famille et à vos amis en ce qui concerne l'efficacité du ginseng?

L'étude est assez crédible en ce qui concerne les méthodes utilisées. Vous pouvez donc vous fier aux résultats. Les participants ont fait l'objet d'une vraie randomisation, ils ne savaient pas de quel groupe ils faisaient partie, et un suivi complet a été effectué. Bien que le nombre moyen de rhumes par personne n'ait pas été réduit de beaucoup (différence moyenne de 0,25 rhume par personne), on a aussi observé une réduction des récurrences du rhume et des symptômes associés à la maladie. Les participants étaient des adultes en santé du Canada, et ceux du groupe prenant du ginseng n'ont pas éprouvé plus d'effets secondaires que ceux du groupe témoin. Ces résultats vous ont convaincu d'aller à la pharmacie et de calculer ce qu'il vous en coûterait de prendre deux capsules de ginseng par jour pendant toute la saison. Vous décidez que si le coût s'élève à moins de 5 \$ par semaine, cela vaudra la peine d'en faire l'expérience vous-même. Vous allez également faire part de vos conclusions à vos amis et à vos proches.

RÉPONSES : FORMULAIRE D'ÉVALUATION CRITIQUE POUR LES INTERVENTIONS

Référence : Predy, G.N. *et al.* « Efficacy of an extract of North American ginseng containing poly-furanosyl-pyranosyl-saccharides for preventing upper respiratory tract infections: a randomized controlled trial », *CMAJ*, vol. 173, n° 9, 2005, p. 1043 à 1048.

Guide	Commentaires
I) Les résultats sont-ils valables?	
Les participants ont-ils été randomisés?	OUI. À la page 1044, il est indiqué qu'il s'agissait d'un essai randomisé. À la page 1045, les auteurs révèlent qu'ils ont utilisé un modèle de randomisation informatisé.
La randomisation a-t-elle été dissimulée?	OUI. À la page 1045, les auteurs affirment avoir utilisé des enveloppes opaques numérotées et scellées.
Les participants ont-ils été analysés dans les groupes dans lesquels ils avaient été placés lors de la randomisation?	OUI. Le schéma de la page 1044 le montre. En outre, à la page 1045, les chercheurs indiquent qu'ils ont effectué une analyse en intention de traiter.
Les participants du groupe de traitement et ceux du groupe témoin étaient-ils semblables en ce qui a trait aux variables prévisionnelles connues?	OUI. Le tableau 1 montre qu'il n'y avait pas de différence significative quant à l'âge, au sexe, au statut de fumeur ou de non-fumeur, au nombre de rhumes des participants et au nombre de sujets ayant trois rhumes ou plus par année.
Les participants savaient-ils dans quel groupe ils avaient été placés?	NON. À la page 1045, on mentionne que les préparations données au groupe d'intervention et au groupe témoin étaient identiques.
Les cliniciens connaissaient-ils la répartition des participants dans les groupes?	NON. Les médecins menant l'étude ne connaissaient pas la répartition des participants.
Les personnes qui évaluaient les résultats connaissaient-elles la répartition des participants dans les groupes?	OUI. Les participants évaluaient eux-mêmes les résultats en tenant un journal dans lequel ils inscrivaient leurs symptômes.
A-t-on effectué un suivi complet?	OUI. Un suivi de quatre mois pendant l'hiver était suffisant pour évaluer si l'intervention a eu un effet sur le nombre de rhumes. À la page 1044, on indique un taux d'abandon de 15 % dans le groupe d'intervention, et de 12 % dans le groupe témoin.
II. Quels sont les résultats?	
Quelle a été la portée du traitement?	Le nombre moyen de rhumes était 0,68 dans le groupe ayant pris du ginseng et 0,93 dans le groupe témoin. La différence moyenne était de 0,25 rhume par personne (P = 0,017), ce qui est statistiquement significatif.

Guide	Commentaires
	<p>Dix pour cent des personnes ayant pris du ginseng ont signalé qu'elles avaient eu deux rhumes ou plus, contre 22,8 % des personnes ayant pris le placebo. Cela représente une réduction du risque absolu de 12,8 % (IC de 95 %, entre 4,3 et 21,3). Cette différence est statistiquement significative, étant donné que l'IC n'inclut pas 1.</p>
<p>Quelle était la précision de l'estimation de l'effet du traitement?</p>	<p>Le nombre moyen de rhumes était 0,68 dans le groupe ayant pris du ginseng et 0,93 dans le groupe témoin. La différence moyenne était de 0,25 rhume par personne (IC de 95 %, entre 0,04 et 0,45).</p> <p>Dix pour cent des personnes ayant pris du ginseng ont signalé qu'elles avaient eu deux rhumes ou plus, contre 22,8 % des personnes ayant pris le placebo. Cela représente une réduction du risque absolu de 12,8 % (IC de 95 %, entre 4,3 et 21,3).</p> <p>Ces deux résultats sont d'une précision moyenne. Une réduction de 0,25 rhume par personne n'est peut-être pas significative sur le plan clinique, mais une réduction de 12,8 % du risque d'avoir deux rhumes ou plus l'est.</p>
<p>III. Comment puis-je appliquer les résultats?</p>	
<p>Les participants de l'étude étaient-ils semblables à ceux de ma propre situation?</p>	<p>OUI. Les participants de l'étude semblent représentatifs de la population générale d'adultes en santé au Canada.</p>
<p>Les résultats importants sur le plan clinique (préjudices et bienfaits) ont-ils tous été pris en compte?</p>	<p>OUI. Les chercheurs ont examiné les symptômes et les effets secondaires : les scores pour les symptômes du rhume étaient, sur le plan statistique, significativement plus bas dans le groupe prenant du ginseng que dans celui prenant un placebo (page 1046). En outre, les effets secondaires signalés par les participants étaient très semblables, et les chercheurs n'ont noté aucune différence statistiquement significative entre les deux groupes (page 1046).</p>

4. Pratique facultative d'évaluation

1. Scénario

(0,5 heure)



Vous faites partie d'un groupe communautaire multidisciplinaire qui s'occupe de la criminalité juvénile dans votre quartier. Votre groupe est composé de représentants du domaine de la santé publique, du milieu de l'éducation, des services sociaux, des services de police et de différents groupes communautaires. À l'une de vos réunions, vous avez discuté de différentes idées de solutions possibles. L'une d'entre elles consistait à offrir, dans les écoles primaires, des interventions comprenant des activités de formation destinées aux enseignants et, peut-être, aux parents. Vous avez offert de chercher dans la documentation afin de trouver des évaluations de ce type d'intervention et de partager les résultats de vos recherches à la prochaine réunion.

1. Vous formulez clairement la question PICO :

P opulation :	enfants du primaire
I ntervention :	enrichissement du programme éducatif, formation des enseignants
C omparaison :	programme éducatif habituel
O utcome (Issue clinique) :	taux de criminalité juvénile

3. Vous effectuez une recherche dans PubMed : <http://www.ncbi.nlm.nih.gov/PubMed/>. Dans la zone de texte au haut de la page, vous tapez « crime prevention » et « school curriculum ». Vous obtenez plusieurs résultats, mais un des articles semble particulièrement intéressant, car il traite des résultats à l'âge adulte ET le texte complet est disponible gratuitement en ligne.

Votre lecture :

Hawkins, J.D., Kosterman, R., Catalano, R.F., Hill, K.G., & Abbott, R.D. (2005). Promoting positive adult functioning through social development intervention in childhood: long-term effects from the Seattle Social Development Project. *Archives of Pediatric and Adolescent Medicine*, 159, 25-31.

Questions :

Comment résumerez-vous cet article à votre groupe?

Recommanderez-vous une intervention axée sur le programme éducatif pour réduire la criminalité?



1. Veuillez lire l'article en entier. Si vous désirez le télécharger et/ou l'imprimer, veuillez le faire en format PDF afin que les pages de l'article correspondent à celles indiquées dans la feuille-réponses.
2. Répondez aux questions d'évaluation critique du *Formulaire d'évaluation critique pour les interventions*. Pour faire cet exercice, servez-vous uniquement de la comparaison entre l'intervention complète et le groupe témoin. (Il existe une autre comparaison faite avec une intervention « tardive »; n'en tenez pas compte.)
3. Comparez vos réponses avec celles déjà inscrites sur la feuille-réponses.

▲ Attention! Dans la VRAIE VIE, vous effectueriez une recherche exhaustive dans la documentation. Aux fins de cet exercice, vous « faites semblant » que cette étude est la SEULE que vous avez trouvée.

FORMULAIRE D'ÉVALUATION CRITIQUE POUR LES INTERVENTIONS

Référence :

Guide	Commentaires
I) Les résultats sont-ils valables?	
Les participants ont-ils été randomisés?	
La randomisation a-t-elle été dissimulée?	
Les participants ont-ils été analysés dans les groupes dans lesquels ils avaient été placés lors de la randomisation?	
Les participants du groupe de traitement et ceux du groupe témoin étaient-ils semblables en ce qui a trait aux variables prévisionnelles connues?	
Les participants savaient-ils dans quel groupe ils avaient été placés?	
Les cliniciens connaissaient-ils la répartition des participants dans les groupes?	
Les personnes qui évaluaient les résultats connaissaient-elles la répartition des participants dans les groupes?	
A-t-on effectué un suivi complet?	
II. Quels sont les résultats?	
Quelle a été la portée du traitement?	
Quelle était la précision de l'estimation de l'effet du traitement?	
III. Comment puis-je appliquer les résultats?	
Les participants de l'étude étaient-ils semblables à ceux de ma propre situation?	
Les résultats importants sur le plan clinique (préjudices et bienfaits) ont-ils tous été pris en compte?	

RÉPONSES : FORMULAIRE D'ÉVALUATION CRITIQUE POUR LES INTERVENTIONS

Référence : Hawkins, J.D. *et al.* « Promoting positive adult functioning through social development intervention in childhood: long-term effects from the Seattle Social Development Project », *Archives of Pediatric and Adolescent Medicine*, 159, 2005, p. 25 à 31.

Guide	Commentaires
I) Les résultats sont-ils valables?	
Les participants ont-ils été randomisés?	NON. Le résumé indique qu'il s'agit d'un essai non-randomisé. À la page 26, on explique que l'étude a d'abord été faite sous forme d'essai randomisé, puis que l'essai s'est étendu à d'autres écoles en cours de route. Ces écoles ont été réparties dans chacun des groupes de façon non aléatoire.
La randomisation a-t-elle été dissimulée?	Ne s'applique pas étant donné qu'il ne s'agissait pas d'un essai randomisé.
Les participants ont-ils été analysés dans les groupes dans lesquels ils avaient été placés lors de la randomisation?	OUI. À la page 28, les auteurs affirment avoir effectué des analyses conservatrices en intention de traiter. Les participants étaient assez âgés pour que les chercheurs puissent évaluer le taux de criminalité juvénile, le sujet qui vous intéresse. Un suivi plus long serait utile pour évaluer les effets de l'intervention sur les taux de criminalité à l'âge adulte.
Les participants du groupe de traitement et ceux du groupe témoin étaient-ils semblables en ce qui a trait aux variables prévisionnelles connues?	OUI. À la page 26, on donne seulement une brève description des caractéristiques de départ de l'échantillon, y compris l'ethnie, le sexe et l'admissibilité à des dîners gratuits. Les auteurs ne comparent pas les groupes dans cet article. Cependant, ils citent un rapport antérieur lié à cette étude (réf. 19), dans lequel ils mentionnaient que les groupes n'étaient pas différents en ce qui concerne : <ul style="list-style-type: none"> - la stabilité résidentielle, mesurée par le nombre moyen d'années vécues à Seattle à l'âge de 12 ans et le nombre moyen de résidences où les participants ont habité entre 5 et 14 ans; - le statut socioéconomique, mesuré par le nombre d'années d'éducation des parents ou la proportion de participants admissibles au programme de dîners gratuits à l'école; - la proportion de participants vivant dans une famille monoparentale; - la proportion de garçons; - la proportion de personnes blanches et non blanches.

Guide	Commentaires
	L'article indique également qu'une proportion relativement égale d'élèves du groupe d'intervention et du groupe témoin vivaient dans des quartiers désorganisés à l'âge de 16 ans, selon ce que les élèves ont eux-mêmes décrit comme des éléments présents dans leur quartier : habitations de qualité inférieure, crime, pauvreté, vente de drogue, gangs et voisins indésirables troublant l'ordre public.
Les participants savaient-ils dans quel groupe ils avaient été placés?	<p>PEU PROBABLE</p> <p>Page 26 : Les parents ont accepté que les élèves participent à l'intervention, et les participants ont accepté de participer à l'entrevue de suivi. Comme l'intervention était appliquée à toute la classe, les élèves ne savaient pas nécessairement que leur classe suivait un programme éducatif différent. Cependant, les enseignants savaient qu'ils donnaient un programme différent et qu'ils suivaient une formation particulière, ce qui peut avoir entraîné un effet Pygmalion.</p>
Les cliniciens connaissaient-ils la répartition des participants dans les groupes?	Non pertinent pour cette étude.
Les personnes qui évaluaient les résultats connaissaient-elles la répartition des participants dans les groupes?	<p>NON. Les participants ont fourni eux-mêmes les renseignements.</p> <p>Page 28 : Les taux de criminalité ont été tirés des registres de l'État et du pays.</p>
A-t-on effectué un suivi complet?	OUI. Quarante-deux pour cent des participants ont été interviewés deux ans après.
II. Quels sont les résultats?	
Quelle a été la portée du traitement?	<p>Page 29, tableau 2 (deux dernières lignes) :</p> <p>Les taux de criminalité tirés des registres concernaient les accusations portées devant les tribunaux au cours de la dernière année et les accusations au cours de la vie entière.</p> <p>La différence moyenne entre le groupe témoin et le groupe ayant reçu l'intervention complète était la suivante :</p> <ol style="list-style-type: none"> 1. accusations au cours de la dernière année : 3 % de moins dans le groupe d'intervention complète que dans le groupe témoin; 2. accusations au cours de la vie entière : 11 % de moins dans le groupe d'intervention complète que dans le groupe témoin.

Guide	Commentaires
<p>Quelle était la précision de l'estimation de l'effet du traitement?</p>	<p>1. Accusations au cours de la dernière année : 3 % de moins dans le groupe d'intervention complète (différence exprimée par -0,03, intervalle de confiance entre -0,10 et 0,04). Cet intervalle franchit la limite d'absence de différence (il inclut 0); ce résultat n'est donc pas statistiquement significatif, comme l'indique aussi la valeur de P, qui est de 0,40.</p> <p>2. Accusations au cours de la vie entière : 11 % de moins dans le groupe d'intervention complète (différence exprimée par -0,11, intervalle de confiance entre -0,21 et -0,01). Cet IC n'inclut pas 0; le résultat est donc statistiquement significatif, comme l'indique aussi la valeur de P, qui est de 0,04. Cet IC signifie que la différence dans le nombre d'accusations peut être aussi élevée que 21 % ou aussi basse que 1 %. Cet intervalle de confiance est assez large, c'est-à-dire pas très précis. La différence est-elle significative? Comme les répondants n'ont que 21 ans, ce résultat mérite d'être pris en compte.</p>
<p>III. Comment puis-je appliquer les résultats?</p>	
<p>Les participants de l'étude étaient-ils semblables à ceux de ma propre situation?</p>	<p>Il serait nécessaire d'avoir plus de renseignements sur votre propre quartier, mais cet article ne donne pas beaucoup d'information. La référence n° 19 contient beaucoup plus de renseignements sur les participants et leur statut socioéconomique.</p>
<p>Les résultats importants sur le plan clinique (préjudices et bienfaits) ont-ils tous été pris en compte?</p>	<p>OUI. L'article fait état de plusieurs autres résultats, fondés sur l'autoévaluation des participants. On a observé des différences statistiquement significatives montrant que les jeunes du groupe d'intervention ont obtenu de meilleurs résultats dans la dernière année que ceux du groupe témoin en ce qui concerne le rendement scolaire et professionnel, la régulation émotionnelle, les pensées suicidaires, la diversité des crimes et la vente de drogue.</p>

Conclusion du scénario

Au début du scénario, on vous a demandé :

1. **Comment résumerez-vous cet article à votre groupe?**
2. **Recommanderez-vous une intervention axée sur le programme éducatif pour réduire la criminalité?**

1. L'étude aurait pu être faite sous forme d'un essai randomisé, ce qui aurait permis d'éliminer toute préoccupation à propos de biais inconnus dans les groupes avant le début de l'intervention. Cependant, ce type d'étude constitue parfois la meilleure source de preuves disponible, particulièrement pour les interventions à l'échelle d'une communauté ou d'une population. Cette étude présente un taux de suivi très impressionnant, étant donné que les chercheurs ont pu suivre les enfants jusqu'à l'âge de 21 ans. De plus, le fait que les chercheurs aient complété les autoévaluations (qui sont habituellement très influencées par le facteur de désirabilité sociale) par des données provenant des registres judiciaires de l'État et du pays représente un atout.

Une différence de 11 % dans les taux de criminalité peut être assez significative à l'échelle communautaire et individuelle. Toutefois, les intervalles de confiance sont larges et, dans le pire des cas, la différence réelle pourrait n'être que de 1 %.

2. Votre recommandation : Le conseil scolaire local devrait examiner les façons d'intégrer l'intervention au programme éducatif déjà très chargé. Il devrait aussi élaborer un budget préliminaire afin de déterminer les coûts d'une telle intervention pour le conseil scolaire au cours des dix prochaines années. Enfin, il devrait prévoir un mécanisme d'évaluation de l'intervention.

▲ Attention! N'oubliez pas que dans la VRAIE VIE, vous prendriez connaissance de toutes les études (ou, idéalement, d'un examen des études existantes) et vous ne fonderiez pas votre décision uniquement sur la présente étude.

6. Références utiles



AUTRES RESSOURCES

Duke University Medical Center. *Introduction to Evidence-Based Medicine. Evaluating the Evidence.* <http://www.hsl.unc.edu/services/tutorials/ebm/Evidence.htm>

GRADE Working Group (2004). Grading quality of evidence and strength of recommendations. *British Medical Journal*, 328,1490-7.

Guyatt, G. & Rennie, D. (Eds) (2002). *Users' Guides to the Medical Literature: A manual for Evidence-Based Clinical Practice.* American Medical Association

7. Glossaire

Analyse en intention de traiter : Fait d'analyser tous les patients dans le groupe dans lequel ils ont été placés lors de la répartition aléatoire, même s'ils n'ont pas terminé l'intervention ou s'ils ont reçu la mauvaise intervention.

Evidence-Based Nursing (glossaire).

<http://ebn.bmj.com/cgi/reprint/11/2/65?maxtoshow=&HITS=10&hits=10&RESULTFORMAT=&titleabstract=Glossary&searchid=1&FIRSTINDEX=0&resourcetype=HWCIT>

Augmentation de l'avantage relatif (AAR) : Augmentation proportionnelle des taux de résultats positifs chez les participants du groupe expérimental par rapport à ceux du groupe témoin; cette valeur est exprimée en pourcentage.

On prend le taux de résultats positifs dans le groupe expérimental (TEE), on soustrait le taux de résultats positifs dans le groupe témoin (TCE), et on divise la différence par le taux de résultats positifs dans le groupe témoin : $TEE - TCE / TCE$.

DiCenso, A., Guyatt, G., & Ciliska, D. (2005). *Evidence-Based Nursing: A guide to clinical practice*. St Louis: Mosby.

Biais : Erreur systématique ou éloignement de la vérité dans les résultats.

Différence de risque absolu : Différence arithmétique entre les taux de survenue d'un événement dans un groupe d'intervention et un groupe témoin (obtenue en soustrayant un taux de l'autre). Cette différence est habituellement exprimée en pourcentage. Si le risque présent dans le groupe d'intervention est plus bas que celui du groupe témoin, on parle alors de **réduction du risque absolu**.

Cullum, N., Ciliska, D., Haynes, R.B., & Marks, S. (2008). *Evidence-Based Nursing. An Introduction*. Oxford: Blackwell.

Données continues : Données pouvant prendre un nombre infini de valeurs le long d'un continuum (poids, tension artérielle).

Cullum, N., Ciliska, D., Haynes, R.B., & Marks, S. (2008). *Evidence-Based Nursing. An Introduction*. Oxford: Blackwell.

Données dichotomiques : Données qui peuvent prendre une seule de deux valeurs (p. ex. mort ou vivant, présence ou absence de symptômes). Aussi appelées données binaires.

Cullum, N., Ciliska, D., Haynes, R.B., & Marks, S. (2008). *Evidence-Based Nursing. An Introduction*. Oxford: Blackwell.

Essais contrôlés randomisés (ECR) : Méthodologie de recherche selon laquelle les sujets sont répartis aléatoirement dans des groupes qui reçoivent des interventions préventives, thérapeutiques ou diagnostiques différentes, et sont ensuite suivis pour que l'on détermine les effets des interventions (un des groupes peut ne recevoir aucune intervention).

Cullum, N., Ciliska, D., Haynes, R.B., & Marks, S. (2008). *Evidence-Based Nursing. An Introduction*. Oxford: Blackwell.

Étude cas-témoin : Étude d'observation dans le cadre de laquelle on compare les patients atteints d'un problème de santé (les cas) aux participants qui n'en sont pas atteints (le groupe témoin), puis on fait un examen rétrospectif afin de déceler de possibles facteurs de causalité. Par exemple, on peut prendre des patients atteints du cancer du poumon et d'autres qui n'en sont pas atteints et examiner leur histoire pour vérifier s'ils ont déjà fumé ou été exposés à la fumée du tabac.

Dawson-Saunders, B., Trapp, R.G. (1994). *Basic and Clinical Biostatistics*. Norwalk: Appleton & Lange

Étude de cohorte : Groupe de personnes possédant un ensemble de caractéristiques communes et qui sont suivies pendant une certaine période afin de déterminer la fréquence d'apparition d'un résultat, sans comparaison avec un groupe témoin.

Cullum, N., Ciliska, D., Haynes, R.B., & Marks, S. (2008). *Evidence-Based Nursing. An Introduction*. Oxford: Blackwell

Examen systématique : Résumé de recherche portant sur toutes les preuves liées à un sujet particulier. Il peut s'agir d'une question sur l'efficacité d'une intervention, ou encore la cause, le diagnostic ou le pronostic d'un événement. Le processus d'examen systématique respecte une méthodologie rigoureuse pour la recherche et la récupération des données, l'évaluation de leur pertinence et de leur qualité, ainsi que l'extraction, la synthèse et l'interprétation des données.

Cullum, N., Ciliska, D., Haynes, R.B., & Marks, S. (2008). *Evidence-Based Nursing. An Introduction*. Oxford: Blackwell

Insu (masquage) : Dans une étude expérimentale, le fait que les patients, les cliniciens qui effectuent l'intervention, les personnes qui évaluent les résultats et/ou les personnes qui analysent les données connaissent ou non la répartition des patients dans les groupes.

Cullum, N., Ciliska, D., Haynes, R.B., & Marks, S. (2008). *Evidence-Based Nursing. An Introduction*. Oxford: Blackwell.

Intervalle de confiance (IC) : Valeur qui quantifie le degré d'incertitude d'une mesure, habituellement exprimée de la façon suivante : intervalle de confiance de 95 %. Cela indique qu'on est certain à 95 % que la valeur réelle pour l'ensemble de la population se situe à l'intérieur des limites indiquées.

Nombre nécessaire à traiter (NNT) : Nombre de patients qu'il est nécessaire de traiter pour prévenir la survenue d'un événement négatif supplémentaire (ou pour favoriser la survenue d'un événement positif supplémentaire). Ce nombre est calculé de la façon suivante : $1/\text{réduction du risque absolu}$ (arrondi au nombre entier le plus près), accompagné par l'intervalle de confiance de 95 %.

Evidence-Based Nursing (glossaire).

<http://ebn.bmj.com/cgi/reprint/11/2/65?maxtoshow=&HITS=10&hits=10&RESULTFORMAT=&titleabstract=Glossary&searchid=1&FIRSTINDEX=0&resourcetype=HWCIT>

Processus décisionnel fondé sur des faits : Utilisation de preuves pour prendre des décisions sur des problèmes ou sujets particuliers ayant trait à la meilleure utilisation des ressources dans des établissements et dans tout le système de santé.

Canadian Health Services Research (2006). *Weighing Up the Evidence. Making evidence-informed guidance accurate, achievable, and acceptable. A summary of the workshop held on September 29, 2005.* http://www.chsrf.ca/other_documents/evidence_e.php#definition, last downloaded May 2008).

Rapport de cotes : Probabilité qu'un événement survienne chez un patient du groupe expérimental divisée par la probabilité que cet événement survienne chez un patient du groupe témoin, ou probabilité qu'un patient du groupe expérimental ait été exposé à un facteur de risque donné divisée par la probabilité qu'un patient du groupe témoin y ait été exposé.

Cullum, N., Ciliska, D., Haynes, R.B., & Marks, S. (2008). *Evidence-Based Nursing. An Introduction.* Oxford: Blackwell

Risque relatif (RR) : Proportion de patients du groupe de traitement (groupe exposé) chez qui on observe un résultat divisée par la proportion de patients du groupe témoin (groupe non exposé) chez qui on observe ce résultat.

Cullum, N., Ciliska, D., Haynes, R.B., & Marks, S. (2008). *Evidence-Based Nursing. An Introduction.* Oxford: Blackwell

Signification statistique : Indique qu'il est peu probable que les résultats d'une analyse résultent du hasard, et que l'hypothèse nulle est écartée (ce qui signifie qu'il existe une différence entre les résultats des deux groupes). Lorsqu'un résultat est statistiquement significatif, la probabilité qu'il soit dû au hasard est inférieure à une limite donnée (la plupart du temps $P < 0,05$).

Valeur de P : Valeur statistique exprimant la probabilité que les résultats obtenus soient exclusivement dus au hasard. Une valeur de $P < 0,05$ signifie qu'il y a moins de 1 chance sur 20 que le résultat soit dû exclusivement au hasard.

Cullum, N., Ciliska, D., Haynes, R.B., & Marks, S. (2008). *Evidence-Based Nursing. An Introduction.* Oxford: Blackwell

Variationnelles : Variables qui ont un effet sur la relation observée entre deux autres variables. Par exemple, la consommation d'alcool est liée au cancer du poumon, sans toutefois causer la maladie. Cependant, l'alcool et le cancer du poumon sont tous les deux liés au tabagisme (la variable confusionnelle), qui cause le cancer du poumon.

Crombie, I.K. (1996). *The pocket guide to critical appraisal: A handbook for Healthcare Professionals.* London: BMJ Publishing Group.