

RÉSULTATS DE RECHERCHE POUR LE SECTEUR HUMANITAIRE

UN GUIDE PRATIQUE



REMERCIEMENTS

Nous sommes très reconnaissants envers ceux qui ont participé au salon Evidence Lounge à Genève, en Suisse, en septembre 2017, au cours duquel ce guide a été discuté, où des versions antérieures du guide ont été commentées ou les deux: Prisca Benelli (Save the Children), Kate Bingley (Christian Aid), Jonathan Breckon (Alliance for Useful Evidence), Alyoscia D'Onofrio (International Rescue Committee - IRC), Tarah Friend (UK Department for International Development - DfID), Tara Karul (International Initiative for Impact Evaluation - 3ie), Paul Knox -Clarke (Active Learning Network for Accountability and Performance - ALNAP), Flurina Kuhn (UK Science & Innovation Network), David Loquercio (Core Humanitarian Standard on Quality and Accountability - CHS), Tim Martineau (Liverpool School of Tropical Medicine - LSTM), Virginia Murray (Public Health England - PHE), Alice Obrecht (ALNAP), Aninia Nadig (Sphere Project), Adina Rom (ETH Zurich), Rodolfo Rossi (International Committee of the Red Cross - ICRC), Eva Svoboda (Overseas Development Institute - ODI, and Humanitarian Policy Group - HPG), Pierluigi Testa (Terre des hommes), Erik Von Elm (Cochrane Switzerland), Karin Wendt (Humanitarian Exchange and Research Centre - ICI).

Par ailleurs, les organisations suivantes ont contribué en fournissant des exemples (y compris des études de cas) et d'autres contenus : 3ie, ALNAP, Christian Aid, Evidence Aid, International Committee of the Red Cross, International Rescue Committee, Liverpool School of Tropical Medicine, London School of Hygiene and Tropical Medicine, Overseas Development Institute, Public Health England, Save the Children, et ShelterBox.



CITATION

Ce document devrait être cité sous la forme suivante : Blanchet K, Allen C, Breckon J, Davies P, Duclos D, Jansen J, Mthiyane H, Clarke M. (2018) Using Research Evidence in the Humanitarian Sector: A practice guide. London, UK: Evidence Aid, London School of Hygiene and Tropical Medicine and Nesta (Alliance for Useful Evidence).

AUTEURS

Ce document a été écrit par Karl Blanchet^a, Claire Allen^b, Jonathan Breckon^c, Phil Davies^b, Diane Duclos^b, Jerien Jansen^b, Helen Mthiyane^c et Mike Clarke^b (**a** Health in Humanitarian Crisis Centre, London School of Hygiene and Tropical Medicine, **b** Evidence Aid, **c** Nesta, Alliance for Useful Evidence).

FINANCEMENT

Ce document a été financé par le UK Science & Innovation Network (représenté par SIN Switzerland, Ambassade de Grande-Bretagne à Berne).

CORRESPONDANCE

La correspondance concernant ce document doit être adressée à Karl Blanchet, Directeur du Centre de santé dans les crises humanitaires, au sein de l'École d'hygiène & de médecine tropicale de Londres, située Tavistock Place à Londres, au Royaume-Uni, et Mike Clarke, directeur de la recherche chez Evidence Aid, au sein du Centre pour la santé publique, Université Queen's à Belfast, ICS Block A, Royal Hospitals, Belfast, également au Royaume-Uni.

Karl: Karl.Blanchet@lshtm.ac.uk
Mike: mclarke@qub.ac.uk

LE TRAVAIL SUR CE GUIDE A ÉTÉ COORDONNÉ PAR TROIS ORGANISATIONS :

Evidence Aid (www.evidenceaid.org) défend une approche fondée sur des preuves dans le secteur humanitaire. Sa mission est de soulager les souffrances et de sauver des vies en fournissant les meilleures preuves disponibles sur l'efficacité de l'action humanitaire et en permettant son utilisation. Evidence Aid a développé une gamme de ressources disponibles gratuitement sur son site Web. Celles-ci comprennent des examens systématiques des interventions en cas de catastrophe, de crise humanitaire et d'autres urgences majeures. Evidence Aid organise des événements pour promouvoir l'utilisation des preuves dans le secteur humanitaire, notamment une Semaine de la preuve humanitaire en novembre, en collaboration avec le Center for Evidence-Based Medicine de l'Université d'Oxford, au Royaume-Uni, parallèlement aux salons Evidence Lounges. Ces salons (Evidence Lounges) rassemblent des praticiens et des membres de la communauté de recherche universitaire, afin d'améliorer les collaborations et de développer des moyens pratiques d'utiliser les preuves dans le secteur humanitaire, y compris le présent guide.

Health in Humanitarian Crisis Centre (<http://crises.lshtm.ac.uk>), à la London School of Hygiene and Tropical Medicine, génère des recherches primaires et des formations sur la santé publique dans les situations de crise humanitaire, en étroite collaboration avec les agences humanitaires internationales et les centres de recherche des pays touchés pour relever les défis critiques en matière de santé.

RECAP, un programme quadriennal de recherche et de renforcement des capacités, a été lancé en 2018 et est axé sur la prise de décision et la responsabilisation en réponse aux crises humanitaires et aux épidémies. Le projet RECAP bénéficie du soutien du Royaume-Uni pour la recherche et l'innovation dans le cadre du Fonds de recherche pour les défis mondiaux, numéro de subvention ES/P010873/1. Le Centre propose également des cours aux praticiens via des modules sur les conflits et la santé, des cours abrégés et un MOOC en ligne gratuit sur la santé dans les crises humanitaires.

Nesta (Alliance for Useful Evidence) (www.alliance4usefulevidence.org) est une fondation mondiale pour l'innovation, basée au Royaume-Uni. Elle soutient de nouvelles idées pour relever les grands défis de notre époque, des pressions exercées par le vieillissement de la population aux services publics surchargés en passant par un marché du travail en rapide mutation. Elle a également publié plusieurs guides sur l'utilisation des données probantes pour relever ce défi. Cette publication s'appuie sur l'un de ceux-ci : « Utiliser les données de recherche: un guide pratique ». Ce guide a été produit par l'équipe des compétences en innovation de Nesta et par l'Alliance for Useful Evidence. Le présent guide pour le secteur humanitaire a été mis au point, en consultation avec eux et en remerciement envers Nesta qui nous a permis d'utiliser une partie de leur contenu.



SOMMAIRE

INTRODUCTION	4		
Qui peut utiliser ce guide ?	5		
Comment utiliser ce guide	5		
SECTION A : Qu'est-ce que la prise de décision éclairée par des preuves, et pourquoi se concentrer sur la recherche ?	6		
Qu'est-ce qu'une « preuve » et pourquoi se concentrer sur la recherche ?	7		
Le défi de ne se fier qu'à une opinion professionnelle	9		
Messages clés de la section A	11		
SECTION B : À quel moment les preuves peuvent-elles vous aider ?	12		
Cycle de réponse humanitaire	12		
Développement et mise en oeuvre d'une nouvelle intervention	14		
Adopter des idées d'autres régions ou secteurs	18		
Créer une théorie du changement	19		
Messages clés de la section B	19		
SECTION C : Quels éléments devriez-vous choisir ?	21		
Partie 1 : Différents types de méthodes de recherche, de concepts et d'approches	23		
Recherche expérimentale - Pourquoi en faire tout un plat?	28		
Modèles de recherche expérimentale	28		
Questions pratiques et éthiques liées aux études randomisées	30		
Lorsque la randomisation est impossible ou inacceptable	30		
Modèles quasi-expérimentaux	30		
		Au-delà de l'expérimentation : de nouvelles approches d'évaluation	32
		Est-ce que cela en a valu la peine ? L'intérêt des analyses coût-avantages et coût-efficacité	32
		Messages clés de la partie 1 de la section C	34
		Partie 2 : Comment évaluer la qualité d'une recherche ?	35
		Contrôle réciproque : comment aident-ils les décideurs ?	35
		Définir une recherche « de haute qualité »	35
		Recherche qualitative de haute qualité	37
		Éviter les sélections au hasard en utilisant des revues systématiques	37
		Méta-analyse	39
		Graphiques en forêt	39
		Évaluation rapide des preuves	40
		L'importance de la répétition et de la corroboration	40
		Examen des publications : comment évaluer des faisceaux d'indices de haute qualité	41
		Messages clés de la partie 2 de la section C	42
		SECTION D : Où devriez-vous chercher des preuves ?	43
		Examiner des résultats de recherche	43
		Sources d'examen systématiques et d'autres recherches pour le secteur humanitaire	44
		Messages clés de la section D	45
		NOTES EN FIN DE DOCUMENT	46
		Références et notes	46

Introduction

Les preuves issues de la recherche et de l'évaluation peuvent vous aider à comprendre ce qui fonctionne, où, pourquoi et pour qui. Elles peuvent également vous indiquer ce qui ne fonctionne pas et vous aider à éviter de répéter les échecs des autres en tirant des enseignements des évaluations de programmes humanitaires infructueux. Les données probantes peuvent également guider la conception des moyens les plus efficaces pour réaliser des interventions spécifiques.

Les preuves remettent en question ce que nous pourrions penser comme étant du sens commun, des connaissances perçues ou reçues.

Par exemple, on a longtemps cru que la malnutrition aiguë sévère nécessitait un traitement spécialisé dans les établissements hospitaliers avec des produits thérapeutiques, même si ce modèle posait de nombreux problèmes pour un traitement efficace à la fois pour les systèmes de santé et les patients. Des aliments thérapeutiques prêts à l'emploi ont été mis au point, au milieu des années 1990. En 2000, des projets pilotes initiaux ont été lancés pour tester l'approche de la gestion communautaire de la malnutrition aiguë (MCMA) lors d'urgences humanitaires.¹ Elle s'est avérée tellement efficace qu'elle a été approuvée par les agences des Nations Unies en 2007,² et est maintenant considérée comme la norme de soins pour la gestion de la malnutrition aiguë dans les contextes d'urgence et de développement.

Les ressources financières et autres étant limitées et souvent insuffisantes dans le secteur humanitaire, nous ne pouvons pas nous permettre de les gaspiller pour des politiques et des programmes inefficaces. Les interventions dans n'importe quel secteur peuvent avoir des résultats à la fois positifs et négatifs, ainsi que des conséquences voulues ou non. Même

dans les cas où l'investissement d'argent pour résoudre un problème est censé faire du bien, nous devons nous demander si cet argent pourrait être dépensé plus efficacement ailleurs, par exemple pour une intervention humanitaire différente. Les résultats de recherche peuvent vous aider à faire des choix éclairés quant aux interventions les plus efficaces à mettre en œuvre dans un contexte d'urgence.

La preuve est un domaine contesté, avec des opinions divergentes sur ce qui devrait être le plus valorisé ou jugé le plus pertinent pour les décideurs.³ Cependant, aux fins de ce guide de pratique, nous mettons l'accent sur des recherches reposant sur des notions scientifiques de preuve, de validité, de fiabilité et minimisant les biais. Une telle recherche présente l'avantage de la rigueur, de la pertinence et de l'indépendance. Nous nous concentrons sur les études de population qui visent à générer des réponses moyennes, qui doivent ensuite être contextualisées, par exemple en prenant en compte des facteurs culturels (tels que le recours à des médecins de sexe masculin pour examiner et traiter des patientes), et des questions de faisabilité (telles que la capacité de mener une intervention à la suite d'une catastrophe lorsque l'accès aux populations ou aux ressources peut être particulièrement limité).

Nous avons créé ce guide pour vous aider à exploiter au mieux les bases factuelles issues de la recherche lorsque vous vous trouvez dans une urgence humanitaire ou lorsque vous planifiez la prochaine urgence. Notre intention est de vous aider à trouver et à utiliser des données factuelles sur les interventions, actions et stratégies susceptibles de vous aider à faire des choix et des décisions éclairés. Ce guide n'explique pas comment générer davantage de données de recherche. Il s'agit d'utiliser et de comprendre les preuves existantes et de reconnaître les preuves manquantes. Cela devrait vous aider à renforcer votre confiance en la compilation, l'assimilation, la synthèse et l'interprétation d'une base de données solide de recherches existantes, et à réfléchir à la manière dont vous pourriez évaluer vos propres projets et commander une étude ou une recherche.

QUI PEUT UTILISER CE GUIDE ?

Ce guide pratique est principalement destiné aux décideurs humanitaires et aux praticiens travaillant sur le terrain ou au siège d'organisations donatrices, internationales, nationales ou non gouvernementales. Cela facilitera les décisions concernant le financement, la supervision, l'exécution ou l'évaluation des interventions humanitaires. Il ne vise pas les évaluateurs et les chercheurs formés, mais cherche plutôt à favoriser la demande de preuves de recherche auprès d'un public plus large du secteur humanitaire.

COMMENT SE SERVIR DE CE GUIDE

Le guide se divise en quatre parties principales :

SECTION A : Qu'est-ce que la prise de décision éclairée par des preuves, et pourquoi se concentrer sur la recherche ?

Cette section discute de ce que nous entendons par prise de décision éclairée par des preuves, et pourquoi la recherche en est un élément essentiel.

SECTION B : À quel moment les preuves peuvent-elles vous aider ?

Cette section explore différents scénarios dans lesquels l'utilisation de preuves peut vous aider, ainsi que les types de preuves dont vous pourriez avoir besoin à différents stades de l'élaboration ou de la mise en œuvre d'une nouvelle intervention ou d'une nouvelle politique.

SECTION C : Quels éléments devriez-vous choisir ?

Cette section examine différents types de données probantes et examine comment choisir la plus appropriée à la question qui vous intéresse. Elle traite également de la façon de juger de la qualité des données probantes.

SECTION D : Où devriez-vous chercher des preuves ?

Cette section propose des conseils et des ressources pour vous aider à trouver les preuves adéquates par rapport à votre cas.

Qu'est-ce que la prise de décision éclairée par des preuves, et pourquoi se concentrer sur la recherche ?

Cette section discute de ce que nous entendons par prise de décision éclairée par des preuves, et pourquoi la recherche en est un élément essentiel.

Pour commencer, disons clairement ce que nous ne voulons pas dire. Nous ne parlons pas de prendre des décisions et des choix en suivant servilement des conclusions de recherche rigides. Le jugement professionnel et d'autres sources d'information, telles que les réactions de vos parties prenantes, seront toujours importants. Ce guide pratique ne vise pas à remplacer le jugement professionnel, mais à accroître l'utilisation de preuves dans l'action humanitaire.

Un bon début pour définir ce que nous voulons dire trouve un exemple dans la médecine. Il y a plus de vingt ans, David Sackett et ses collègues ont proposé la définition suivante qui a résisté à l'épreuve du temps : « *La médecine fondée sur des preuves est l'utilisation consciencieuse, explicite et judicieuse des meilleures preuves actuelles pour prendre des décisions concernant les soins à apporter à des patients. La pratique de la médecine factuelle implique d'intégrer l'expertise clinique individuelle aux meilleures preuves cliniques externes disponibles issues de la recherche systématique* ». ⁴

Cette tentative de définir la médecine fondée sur des preuves n'était pas la première, ⁵ mais elle a eu une influence et est tout aussi pertinente pour le secteur humanitaire que pour d'autres secteurs. Elle souligne que la recherche peut compléter le jugement professionnel ou d'autres sources d'informations et reconnaît l'importance

des preuves sur des questions telles que la faisabilité, les préférences et la culture.

Dans un domaine tel que le secteur humanitaire où des preuves plus nombreuses et de meilleure qualité sont nécessaires, ⁶ nous pouvons supposer que tout modèle de bonne prise de décision doit éviter de s'appuyer sur le seul jugement professionnel non étayé par des preuves scientifiques. Plus loin dans cette section, vous découvrirez comment nous pouvons tous être « irrationnels de façon prévisible » et, consciemment ou inconsciemment, faire des erreurs lors de jugements importants. Nous explorons les moyens d'atténuer ces erreurs de jugement dans les sections suivantes. Cependant, d'autres modèles décisionnels ont également souligné l'importance de combiner la connaissance des preuves avec le jugement. Le secteur humanitaire est un domaine sensible dans lequel nous devons être conscients de la politique internationale et locale et de la dynamique entre les différents acteurs impliqués dans la fourniture de l'aide humanitaire. Cela déterminera parfois l'accès aux preuves et aux informations, mais aussi la manière dont l'aide humanitaire est fournie. Cependant, l'importance des preuves demeure et, comme indiqué dans un rapport d'ALNAP en 2014, « *l'incapacité à générer et utiliser des preuves dans les politiques et les réponses rend l'action humanitaire moins efficace, moins éthique et moins responsable* ». ⁷

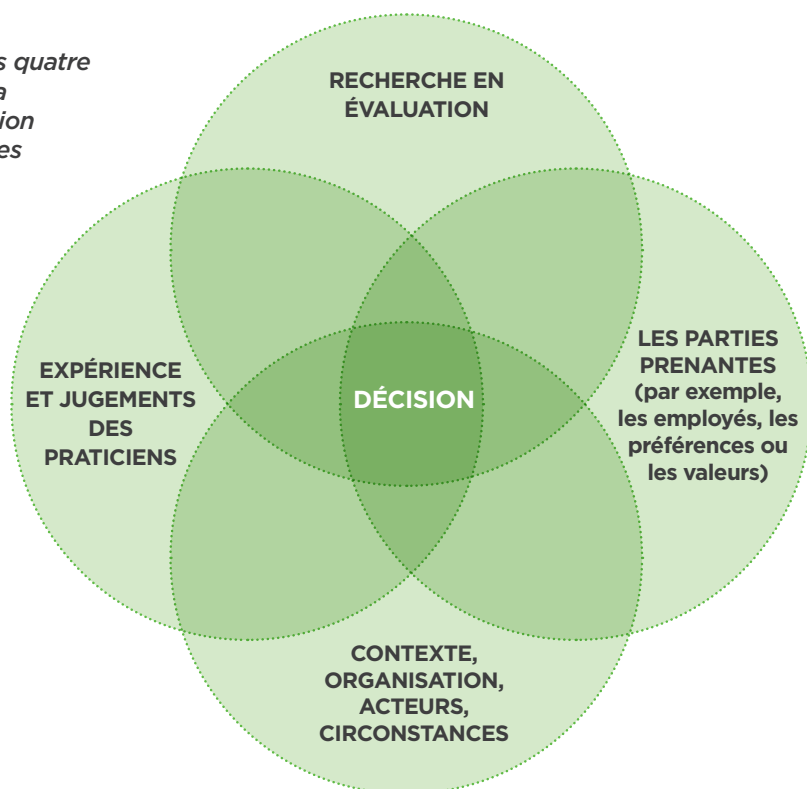
QU'EST-CE QU'UNE « PREUVE », ET POURQUOI SE CONCENTRER SUR LA RECHERCHE ?

Le dictionnaire anglais Oxford définit le terme « preuve » comme étant « l'ensemble des faits ou informations disponibles indiquant si une croyance ou une proposition est vraie ou valide »,⁸ et, de même, dans leur rapport ALNAP sur l'état des preuves dans le secteur humanitaire, Paul Knox Clarke et James Darcy le définissaient comme suit : « des informations permettant de justifier ou de prouver / réfuter la vérité d'une proposition spécifique ». ⁷ Nous suivons ces définitions car de nombreuses autres définitions ont tendance à être plutôt inutiles en étant trop inclusives (en incluant parfois presque tous les types d'informations) ou en étant trop abstraites et vagues.

La **figure A.1** montre les différents éléments qui devraient faire partie d'une prise de décision éclairée par des preuves. Notre guide de pratique se concentre sur le cercle supérieur du diagramme : recherche et évaluation.

Comme l'indiquent les auteurs du rapport «*What Counts as Good Evidence?*» de l'Alliance for Useful Evidence, «*La conduite et la publication de la recherche impliquent la documentation explicite des méthodes, l'examen par les pairs et l'examen externe, aboutissant à de la rigueur et à de l'ouverture. Ces caractéristiques contribuent à sa nature systématique et aident à fournir un moyen de juger de la fiabilité des conclusions. Elles offrent également la possibilité d'évaluer la validité d'une revendication par rapport à une autre* ». ⁹

Figure A.1: Les quatre éléments de la prise de décision éclairée par des preuves



Basé sur : Barends E, Rousseau DM, Briner RB. (2014) Evidence-based Management: The Basic Principles. Amsterdam: Center for Evidence-Based Management [www.cebma.org/wp-content/uploads/Evidence-Based-Practice-The-Basic-Principles.pdf]

Ce guide de pratique est axé sur la recherche, mais il existe de nombreux chevauchements avec le domaine de l'évaluation. Nous abordons certaines approches de l'évaluation de l'impact et des processus dans la section C. Nous accordons également plus d'attention à la recherche qui traite de l'impact - que quelque chose ait eu des résultats positifs ou négatifs - parce que les questions relatives à l'impact sont vitales pour ceux qui participent à l'action humanitaire. Ces acteurs sont soucieux de montrer leur « impact » sur les populations, leurs « résultats » en termes internationaux ou « ce qui fonctionne » pour les gouvernements et les fournisseurs locaux et internationaux. Le langage peut changer, mais l'idée de leur recherche reste la même : vérifier s'ils ont vraiment fait la différence. Par conséquent, la finalité de ce guide est de vous aider à décider de la manière dont cette recherche pourrait vous aider à choisir des interventions, des actions et des stratégies, et à adopter les politiques les plus susceptibles de faire une différence positive. Nous fournissons des exemples illustratifs tout au long du guide, et d'autres exemples d'utilisation des preuves dans les secteurs humanitaires sont disponibles dans d'autres recueils d'études de cas.¹⁰

Nous accordons une importance particulière à la recherche et à l'évaluation qui sont prêtes à l'emploi, sans devoir mener une nouvelle étude. Les décideurs disposent de peu de temps et de ressources, et beaucoup ne peuvent tout simplement pas se permettre de commander une telle étude et d'attendre que ses résultats deviennent disponibles, ce qui peut prendre

des années. Quelqu'un qui doit prendre une décision maintenant a besoin de la preuve maintenant, voire hier, pas dans un an ou plus. Les décideurs ont donc besoin de preuves qui puissent être recueillies « sur une étagère » et combinées avec des informations sur le contexte local pour éclairer leur choix. Heureusement, il est possible de trouver de telles preuves et nous couvrons certaines des manières de le faire à la section D.

La recherche est un processus engagé à des fins d'apprentissage. Elle cherche à répondre à des questions telles que « Quel était le type de blessure le plus courant après un tremblement de terre ? », « Quels sont les effets sur la violence sexiste de différentes manières pour protéger les femmes et les enfants ? » ou « Dans quelle mesure un matériau est-il imperméable lorsqu'il est utilisé comme abri ? »

L'évaluation est un processus impliquant l'évaluation des résultats et des observations par rapport à des normes, dans le but de prendre des décisions. Les évaluations posent des questions telles que « Sur quels types de premiers secours les premiers intervenants devraient-ils être formés ? », « Quel est le meilleur moyen de protéger les femmes et les enfants contre la violence sexiste ? » ou « Quel matériau faut-il utiliser pour la fabrication de tentes dans un environnement très pluvieux ? »

La recherche ne nécessite pas nécessairement d'évaluation. Cependant, faire une évaluation nécessite toujours de faire de la recherche. Une évaluation concerne une intervention qui a été effectivement mise en œuvre, alors que la

Nous accordons une importance particulière à la recherche et à l'évaluation prêtes à l'emploi, sans qu'il soit nécessaire de mener une nouvelle étude.

recherche est plus complète et, outre les évaluations, elle peut également tenter de répondre à des questions conceptuelles, telles que la planification des besoins susceptibles de suivre une catastrophe ou la création d'une nouvelle intervention. Cependant, que nous pensions à la recherche ou à l'évaluation, nous devons être conscients que les partis pris peuvent fausser les résultats. Dans la section C, nous examinons certains des partis pris qui peuvent affecter les résultats d'une étude, mais il est également important d'être conscients des partis pris dans la façon dont les gens interprètent les résultats d'une étude, ou qui influencent leur réflexion sur le besoin de recherche ou sur l'impact de cette dernière. Nous discutons ensuite certains de ces partis pris.

LE DÉFI DE COMPTER UNIQUEMENT SUR L'OPINION PROFESSIONNELLE

L'une des raisons pour lesquelles nous privilégions une recherche de qualité, plutôt que de nous fier uniquement à un jugement professionnel, est que les préjugés impliqués dans le jugement professionnel peuvent signifier qu'il est incorrect. En tant qu'êtres humains, nous sommes « naturellement irrationnels »¹¹ et sommes susceptibles de subir des biais cognitifs qui déforment notre pensée. Le **biais cognitif** fait référence à notre incapacité à être totalement objectif, ce qui peut se manifester de multiples façons - telles que la distorsion de la perception, des jugements inexacts et des interprétations illogiques ou irrationnelles.

Même avec les meilleures intentions, les professionnels peuvent se tromper. Prenons par exemple un type d'erreur cognitive: **le préjugé de confirmation**. C'est la tendance à se concentrer sur les preuves qui correspondent à ce que nous croyons et à ignorer ou à écarter ce qui ne l'est pas. Même des professionnels très expérimentés peuvent tomber dans ce piège cognitif.

Il est important de savoir comment on peut sauter aux conclusions pour se méfier des experts. Mais le préjugé de confirmation

souligne également le fait que n'importe qui - pas seulement des experts - peut utiliser de manière très sélective les résultats de recherche. Les gens ont tendance à rechercher des preuves qui correspondent à leurs croyances, intentionnellement ou non. Cela pourrait être considéré comme une « preuve fondée sur une politique » plutôt que comme une « politique fondée sur une preuve » : choisir et réajuster les preuves aux conclusions que nous avons déjà fixées dans notre esprit. Si nous sommes généreux, cela peut être une erreur inconsciente, mais dans le pire des cas, il peut s'agir d'un stratagème délibéré pour renforcer les préjugés en trouvant et en promouvant des éléments de preuve qui correspondent à ces préjugés, tout en ignorant ce qui ne l'est pas.

Un phénomène similaire mais différent est le **préjugé d'optimisme**, qui fait référence à la « *tendance prouvée des évaluateurs [de projets, programmes et politiques] à être trop optimiste quant aux paramètres clés du projet, y compris les coûts en capital, les coûts de fonctionnement, la durée du projet et la fourniture des avantages* ».¹² Cela peut aboutir à la définition d'objectifs pour les produits, les résultats et les impacts irréalistes et non livrables. Christoplos a noté que, à la suite du tsunami de 2004 dans l'océan Indien, « *Les méthodes et les directives étaient remplies d'optimisme normatif. Des termes tels que « Web homogène » ont été appliqués pour décrire l'état de la situation qui se présenterait, si les bonnes méthodes étaient appliquées. Les individus, les ménages, les communautés et les nations devaient suivre un continuum de souffrances humaines aiguës au développement « normal ». Cet optimisme s'est estompé au fil des ans. LRRD [Links between Relief, Rehabilitation and Development] s'est plus référé à une énigme qu'à un chemin doré. Des tensions politiques insupportables et l'insécurité humaine sont désormais associées à LRRD, d'abord dans les situations d'après conflit, puis de plus en plus souvent lors de catastrophes naturelles* ».¹³

Le préjugé d'optimisme concernant les anciennes et les nouvelles interventions est souvent dû à un manque de preuves de leurs véritables effets. La recherche et les preuves tirées d'évaluations de ces interventions, ou similaires, peuvent aider à réduire cette incertitude.

Comme dans d'autres secteurs, il existe de nombreux autres préjugés liés à la façon dont les gens pensent que cela peut affliger ceux qui travaillent dans le secteur humanitaire. Ceci inclut :

Le préjugé du recul : Tendance à considérer les événements passés comme étant plus prévisibles qu'avant leur survenue.

L'aversion aux pertes : Tendance à préférer éviter les pertes qu'à acquérir des gains.

L'effet de cadrage : Tirer des conclusions différentes à partir de la même information présentée de différentes manières (par exemple, préféreriez-vous que « 95% sont revenus au travail », ou que « 5% ne sont pas retournés au travail » ?).

La « heuristique de disponibilité » : Lorsque les gens associent la taille, la fréquence ou la probabilité d'un problème à la facilité avec laquelle il est facile de s'en souvenir ou de l'imaginer.

La « représentativité heuristique » : Quand les gens surestiment la probabilité d'événements vivants.

Le « besoin de cohérence » : L'envie d'établir des modèles et des relations de cause à effet quand ils n'existent peut-être pas.

Le préjugé métacognitif : la conviction que nous sommes immunisés contre les préjugés !

Cela ne veut pas dire que le jugement professionnel est toujours faux. Des chercheurs, tels que Gary Klein, ont vanté les mérites d'un jugement d'expert intuitif, par exemple dans ses travaux sur « la prise de décision naturaliste ». ¹⁴ Les points de vue professionnels et les instincts peuvent être très précieux, mais nous devons être conscients de leurs inconvénients. Comme l'a affirmé Daniel Kahneman dans un article publié conjointement avec le professeur Klein dans *American Psychologist*, « *l'intuition professionnelle est parfois merveilleuse et parfois imparfaite* ». ¹⁵

ÉTUDE DE CAS

Comment la recherche a favorisé la croissance de la programmation de transferts monétaires

Ces dernières années, les programmes de transferts monétaires (CTP) sont devenus l'une des innovations les plus importantes de l'aide humanitaire internationale.

En 2016, 2,8 milliards de \$ d'aide humanitaire ont été décaissés en espèces et en bons, soit une augmentation de 40% par rapport à 2015 et un doublement par rapport à 2014.¹⁶ Les décaissements en espèces ont non seulement augmenté de manière globale, mais aussi en tant que part relative de l'aide humanitaire internationale totale, de 7,8% en 2015 à 10,3% en 2016.¹⁷

Cette expansion a été appuyée par un nombre croissant d'évaluations, ce qui a permis de rassembler un ensemble de preuves sur les effets de différents programmes sur les résultats individuels et au niveau des ménages. Cet ensemble de recherches et d'évaluations a joué un rôle déterminant dans la croissance des transferts monétaires.

Cela a également permis de résoudre les problèmes liés aux transferts monétaires,

tels que la corruption et l'insécurité,¹⁸ ainsi que de mieux faire comprendre que les transferts monétaires génèrent des risques différents, mais pas nécessairement plus importants, que l'assistance en nature.¹⁹ Il a été démontré que les avantages des transferts monétaires étaient multiples dans de nombreux secteurs, notamment en ce qui concerne l'autonomisation des femmes, la sécurité alimentaire, l'éducation et les soins de santé. En conséquence, les transferts monétaires sont désormais un outil accepté dans presque toutes les interventions d'urgence. L'une des principales réussites de la recherche sur les transferts monétaires est que les preuves générées ont été largement utilisées par les institutions défendant la promotion de l'utilisation des espèces dans les crises humanitaires.

Des recherches supplémentaires sont néanmoins nécessaires pour ajuster les programmes afin d'en maximiser les avantages.

Messages clés de la section A

✓ Nous ne prôtons pas des prises de décision qui suivent servilement les conclusions de recherches aux dépens de jugements professionnels.

✓ Cependant, l'expertise professionnelle ne doit pas être placée sur un piédestal. Les experts peuvent parfois se tromper et ne sont pas à l'abri des préjugés sociaux et cognitifs, tout comme les profanes.

✓ Les résultats des recherches constituent un type de preuve parmi d'autres, mais ils présentent l'avantage d'être plus rigoureux, pertinents et indépendants par rapport à d'autres types de preuves.

✓ La génération de nouvelles preuves peut être coûteuse et longue, mais il existe des preuves de bonne qualité « prêtes à l'emploi » (voir section D). Les recherches et les évaluations des urgences actuelles peuvent fournir des informations pour de futures interventions.

À quel moment les preuves peuvent-elles vous aider ?

Cette section explore différents scénarios dans lesquels l'utilisation de preuves peut vous aider, ainsi que les types de preuves dont vous pourriez avoir besoin à différentes étapes de l'élaboration ou de la mise en œuvre d'une nouvelle intervention ou d'une nouvelle politique.

Les preuves peuvent rendre les organisations plus efficaces. Des campagnes plus convaincantes à la recherche de subventions, en passant par le développement des capacités de prise de décision d'une organisation pour s'assurer que les programmes produisent des résultats, les preuves peuvent renforcer votre travail. Peu importe que vous soyez une petite organisation bénévole ou une grande organisation internationale. Quelle que soit votre taille, vous trouverez peut-être des recherches répondant à vos besoins.

LES PREUVES PEUVENT VOUS AIDER DE MULTIPLES FAÇONS, NOTAMMENT :

Commencer à faire quelque chose

Arrêter de faire quelque chose

Proposer des options et apporter des arguments

Identifier les priorités

Évaluer les besoins

Aligner les services avec les besoins

Concevoir et mettre en œuvre des programmes plus efficaces

Développer des demandes de financement

Créer des campagnes de sensibilisation efficaces

Accroître la responsabilisation des parties prenantes

CYCLE DE RÉPONSE HUMANITAIRE

Le Bureau de la coordination des affaires humanitaires (BCAH), département du Secrétariat de l'ONU, a développé un cycle de réponse humanitaire qui prévoit une série d'actions coordonnées visant à préparer, à gérer et à exécuter l'intervention humanitaire. Il repose sur cinq éléments, chaque étape s'appuyant sur l'étape précédente et conduisant à l'étape suivante.²⁰ Les preuves peuvent vous aider à prendre des décisions et à avancer dans ce cycle.

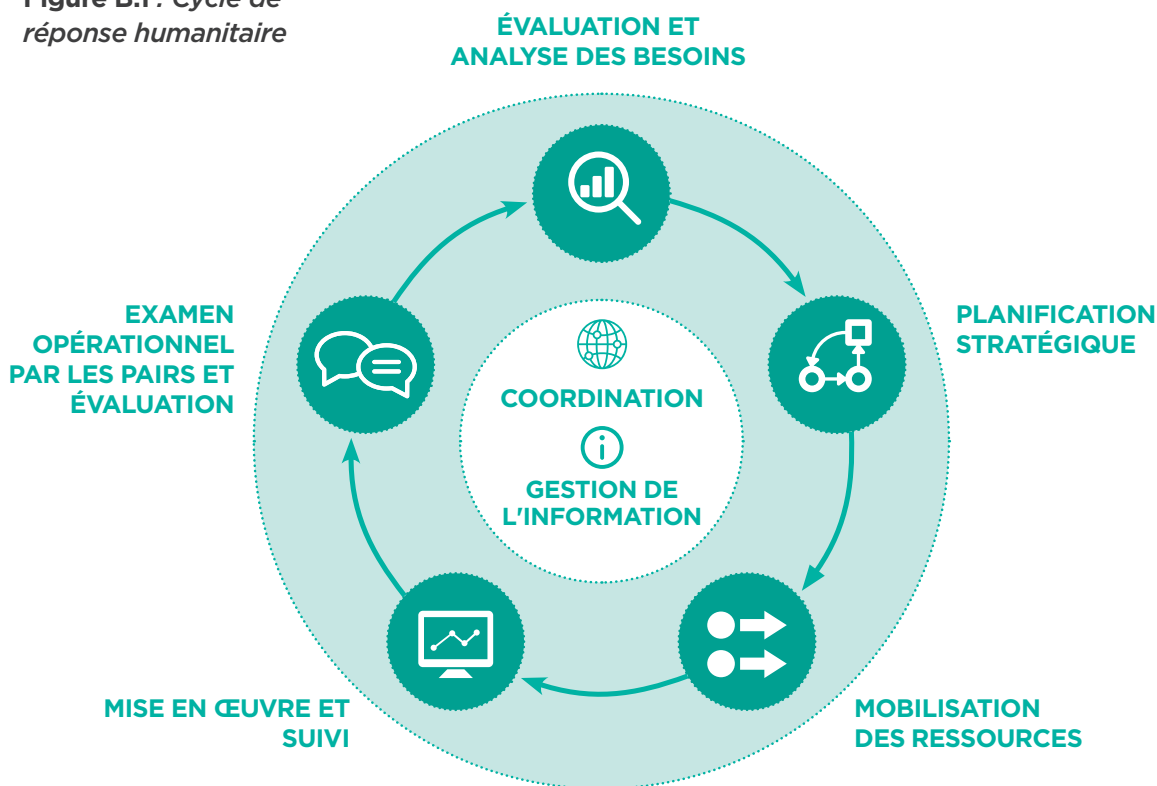
Besoins en termes d'évaluation et d'analyse :

L'analyse des recherches existantes vous fournira une base d'évaluation des besoins et vous aidera à comprendre la nature, l'importance et la dynamique de certains problèmes. Parallèlement à la collecte de données liées à l'environnement et aux circonstances auxquelles vous devez faire face (ou devrez peut-être faire face), cette analyse vous permettra d'identifier des opportunités de réponses appropriées. Dans certains cas, cette démarche peut s'avérer difficile si les opportunités de recherches pertinentes ont été peu nombreuses.²¹

Planification de réponse stratégique :

Lors de la planification de votre réponse stratégique, recherchez des preuves sur les tests et évaluations réalisés précédemment, car ces informations pourraient vous aider à prendre les bonnes décisions. Vous pouvez également utiliser des preuves sur les meilleurs moyens de communiquer les risques potentiels d'une situation aux personnes qui ont pu être impliquées.²²

Figure B.1 : Cycle de
réponse humanitaire



Mobilisation des ressources :

L'implémentation d'une nouvelle intervention ou politique nécessitera la mobilisation de ressources. Ces ressources pourront être de nature financière, afin de payer les activités ou les équipes qui seront chargées de leur exécution. L'adoption d'une approche éclairée par les preuves lors de la planification d'une réponse stratégique et le choix d'interventions et de politiques est susceptible de les rendre plus efficaces et donc plus attractives aux yeux des donateurs.²³

Mise en œuvre et suivi : Une fois que vous avez mis en œuvre une intervention ou une politique, vous devez réfléchir à la façon dont vous pouvez en réaliser le suivi et documenter son impact. Vous contribuerez ainsi à faire la différence.

Examen opérationnel par les pairs et évaluation

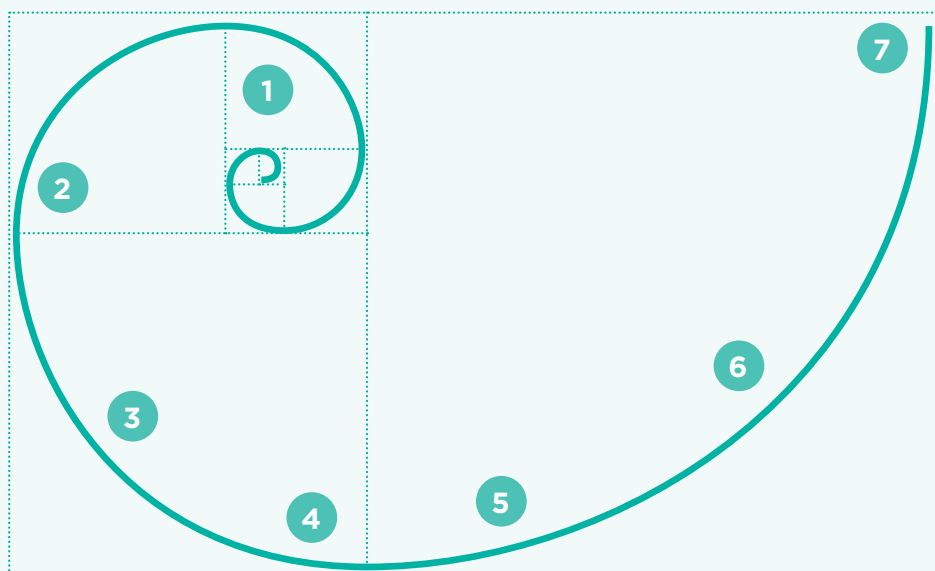
L'examen opérationnel par les pairs peut être utilisé comme un correcteur de formation afin d'identifier les domaines nécessitant des actions correctives immédiates. Parallèlement à une évaluation de votre propre réponse stratégique et des comparaisons avec des preuves issues d'interventions similaires, cet examen peut vous aider à déterminer si des ajustements ou des améliorations sont nécessaires, notamment pour l'organisation des équipes de direction, la mise en oeuvre d'autres phases du cycle, la coordination et les mécanismes de responsabilisation des personnes affectées.

DÉVELOPPEMENT ET MISE EN ŒUVRE D'UNE NOUVELLE INTERVENTION

Lors du développement ou de la mise en œuvre d'une nouvelle politique ou d'une nouvelle intervention humanitaire, il convient de commencer par prendre en compte la durée. Les preuves seront utiles à des moments différents du développement ou de la mise en œuvre et vous devrez examiner les recherches appropriées pour chaque étape du cycle de vie de la nouvelle intervention ou politique. Au tout début d'une nouvelle initiative, les recherches peuvent permettre d'identifier de nouveaux défis et d'évaluer l'étendue du problème. Elles peuvent servir à préparer le développement et la mise en œuvre d'une nouvelle intervention ou politique.

Le graphique en spirale ci-dessous a été élaboré par Nesta pour illustrer les différentes étapes du processus d'innovation. Ce graphique peut être utilisé pour déterminer l'avancement d'une nouvelle approche appliquée à un problème social.²⁴ Différents types de preuves seront requis à chaque étape et le processus doit être continu : une fois la fin d'une spirale atteinte, vous devez recommencer et entamer un nouveau cycle. Nous abordons dans ce document les différentes étapes de cette spirale. Des exemples spécifiques d'innovation réussie dans des contextes humanitaires sont également disponibles dans une autre étude portant sur 15 projets financés par le Fonds d'innovation humanitaire.²⁵

Figure B.2 : La spirale d'innovation de Nesta



- 1 Exploration des opportunités et des défis
- 2 Production d'idées
- 3 Développement et test

- 4 Apporter des arguments
- 5 Prestation et mise en œuvre
- 6 Croissance et évolution
- 7 Changement de systèmes

1. EXPLORATION DES OPPORTUNITÉS ET DES DÉFIS

L'utilisation des recherches actuelles vous aidera à comprendre la nature, l'importance et la dynamique de certains problèmes auxquels vous êtes confronté et les opportunités de réponses appropriées. (Voir l'étude de cas ci-dessous.)

ÉTUDE DE CAS

Présentation de preuves sur les résultats des systèmes de santé à long terme dans des situations de crise prolongée

Conformément au principe « ne laisser personne pour compte » inclus dans les objectifs de développement durable, la plupart des partenaires de développement reconnaissent que la réponse aux besoins humanitaires dans le domaine de la santé lors d'une crise doit pouvoir s'étendre à un système de santé équitable sur le plus long terme.

Le **ReBUILD Research Programme Consortium** est un partenariat de recherche internationale financé par UK Aid. Son objectif est d'améliorer l'accès à des soins efficaces pour les populations pauvres se trouvant dans des situations d'après conflit, grâce à de solides preuves sur le renforcement des systèmes de santé. Menées dans des contextes d'après conflit relativement stables, bien que fragiles, les recherches de ReBUILD ont permis d'analyser, à travers un prisme historique, les impacts à court, moyen et long terme sur les systèmes de santé des politiques mises en œuvre pendant et après le conflit et la crise. Plusieurs thèmes y étaient étudiés, notamment la rémunération des professionnels de santé, la mise en place de structures de soins destinées aux communautés les plus pauvres et les plus vulnérables, l'efficacité de l'aide et la parité homme-femme.

Ces recherches ont apporté un éclairage précieux sur l'expérience des communautés et des professionnels de santé pendant les périodes de conflit et d'après conflit, ainsi que des recommandations sur le renforcement des systèmes de santé après le conflit. Les preuves de ReBUILD présentent également un intérêt sur le plan humanitaire, par exemple pour les effets d'approches favorisant l'implication des

professionnels de santé locaux dans des interventions d'urgence sur la distribution de la main-d'œuvre à long terme, la motivation et les compétences, ou encore l'utilité de travailler avec des acteurs locaux (même si leur rôle est limité) pour soutenir à plus long terme des capacités de coordination, de planification et de flux de ressources prévisibles.

C'est pour cette raison que ReBUILD mobilise également les personnes impliquées dans la santé lors de conflits actuels ou de crises prolongées. Entre autres réalisations, ReBUILD a participé aux prises de décision et à la mise en œuvre en Sierra Leone lors de l'épidémie du virus Ebola, et a appuyé certains aspects d'un plan de donation pour la Syrie. Il est difficile de présenter des preuves sur des résultats à long terme aux personnes travaillant dans une interface développement/humanitaire, où les besoins humanitaires restent de façon compréhensible l'objectif principal. Les preuves à long terme n'occupent qu'une place très limitée dans les objectifs et les stratégies des organisations humanitaires, ce qui reflète l'écart entre approche de développement et approche humanitaire.

ReBUILD collabore avec plusieurs organisations qui considèrent les approches humanitaires basées sur des preuves (notamment Evidence Aid) et collabore étroitement avec des groupes de travail ayant pour objectif l'interface humanitaire/développement pour la santé, comme le **groupe de travail de la CSU2030 sur la CSU dans les environnements fragiles** et le **groupe de travail thématique global sur les systèmes de santé dans les états fragiles et affectés par des conflits**.

2. PRODUCTION D'IDÉES

Une fois votre objectif fixé, vous devez rechercher des interventions et des politiques qui ont été précédemment étudiées. Pouvez-vous réutiliser des idées qui ont fonctionné pour d'autres et éviter de reproduire des interventions qui ont échoué ?

3. DÉVELOPPEMENT ET TEST

Alors que de nouvelles idées sont lancées, vous devez réaliser des tests et des expériences sur les différentes approches afin d'en évaluer l'impact.

Dans la section C, nous abordons les normes en matière de preuves qui illustrent la trajectoire que doivent suivre les nouveaux programmes pour démontrer leur impact.

4. APPORTER DES ARGUMENTS

Toute preuve concernant l'impact probable de votre intervention vous aidera à apporter des arguments aux donateurs potentiels et aux acteurs sur le terrain. Vous serez mieux en mesure de les persuader de financer l'intervention et d'apporter leur soutien à son exécution et à sa mise en œuvre. (Voir l'étude de cas ci-dessous.)

ÉTUDE DE CAS



Présentation de preuves aux dirigeants des agences locales

En juin 2017, Christine Fernandes, conseillère globale en nutrition humanitaire pour Save the Children, a donné une présentation sur l'alimentation des nourrissons et des jeunes enfants dans les situations d'urgence à l'équipe de direction de Save the Children en Irak, qui passait en revue sa stratégie pour 2017-2018. Jusqu'alors, la nutrition n'avait pas été prise en compte pour la mise en œuvre en Irak. Christine a présenté sa présentation de 20 minutes sous forme de quiz : chaque diapositive commençait par une question comme :

Selon vous, quel est le moyen le plus efficace de prévenir la mortalité chez les enfants de moins de cinq ans ?

Matériaux traités avec des insecticides.

Vaccin Hib (méningite).

Allaitement maternel.

Compléments alimentaires.

Compléments en vitamine A et en zinc.

Elle a ensuite tordu le cou aux idées reçues en s'appuyant sur des données tangibles. Par exemple, en citant un article du magazine The Lancet,²⁶ elle a pu démontrer que l'allaitement maternel contribue largement (13 %) à la réduction de la mortalité chez les jeunes enfants et que les compléments alimentaires représentent le deuxième facteur le plus élevé. La méthode hypothèse/correction des idées reçues est également une façon efficace d'influencer l'opinion des gens, toujours sur la base de preuves.²⁷ Christine est parvenue à convaincre toute la direction de l'agence irakienne de l'intérêt de faire de l'alimentation des nourrissons et des jeunes enfants dans les situations d'urgence un thème prioritaire. Grâce à cette présentation, un programme d'alimentation a été lancé à Bagdad et le Ministère de la Santé a bénéficié de soutien lors de la mise en place de sa stratégie de nutrition nationale, qui a mis l'accent sur le contrôle des substituts au lait maternel, et du renforcement des capacités à l'échelle nationale dans le cadre des premiers investissements en ressources en 2018.

5. PRESTATION ET MISE EN ŒUVRE.

Une fois que vous avez mis en œuvre une intervention ou une politique, vous devez réfléchir à la façon dont vous pouvez documenter l'impact correspondant. Cela vous aidera à passer d'une position de rhétorique consistant à dire 'croyez-moi, ce projet fonctionne' à une position s'appuyant sur des preuves plus fiables issues du suivi et de l'évaluation. Cela vous aidera également à rendre des comptes en montrant que vous changez les choses, que votre projet est rentable et que vous avez procédé à une évaluation de votre travail. (Voir l'étude de cas ci-dessous.)

ÉTUDE DE CAS



Partenariats dans le domaine des évaluations d'impact

3ie (International Initiative for Impact Evaluation) a pour objectif d'améliorer les conditions de vie des populations pauvres dans les pays à faible revenu et à revenu intermédiaire, en transmettant des données probantes sur les interventions de développement qui fonctionnent (quand, pourquoi et dans quelle mesure) et en synthétisant ces données. En 2018, le dispositif thématique d'aide humanitaire de 3ie a financé sept évaluations d'impact au Tchad, en République démocratique du Congo (RDC), au Mali, au Niger, au Pakistan, au Soudan et en Ouganda. C'est ce que révèle une importante étude réalisée en Allemagne par l'Université de Mannheim, en collaboration avec l'ACTED (Agency for Technical Cooperation and Development). ACTED soutient des programmes d'aide humanitaire dans les zones sujettes aux catastrophes, et l'évaluation évalue l'efficacité de ces interventions afin de répondre à ces risques et de réduire la vulnérabilité des communautés vis-à-vis des bouleversements d'urgence. La collaboration entre ACTED et l'Université de Mannheim a inclus plusieurs exemples réussis de renforcement des capacités et de transmission :

Engagement à plusieurs niveaux :

L'étude et les données de référence ont été présentées au gouvernement pakistanais dans le but d'informer au sujet de la stratégie nutritionnelle intégrée du Pakistan, en s'appuyant sur les

informations fournies par divers acteurs expérimentés en matière de nutrition et de gestion des catastrophes. Les résultats ont également été communiqués à un groupe de travail national composé d'ONG nationales et internationales, d'universitaires et de spécialistes du secteur.

Contribution à l'amélioration de la qualité des données :

Il y a eu des améliorations significatives dans le système de surveillance et d'évaluation. Les bonnes pratiques en matière de collecte de données (vérifications ponctuelles, séances de débriefing avec les équipes de terrain et commentaires réguliers sur les divergences en matière de collecte des données) mises en place lors de l'évaluation d'impact réalisée par l'équipe de recherche, ont été reprises dans un autre projet mis en place par ACTED.

ACTED espère inclure ces pratiques dans son protocole régulier de suivi et d'évaluation, une fois qu'elles auront été testées sur le terrain.

Renforcement de la capacité de l'agence d'exécution en matière d'évaluation :

Le personnel de l'agence d'exécution a été sponsorisé pour participer à un atelier sur la mesure des effets des programmes de réduction de la pauvreté et aider à élaborer de meilleures politiques par le recours aux études d'impact. Cela lui permettra de mieux assimiler les résultats de l'évaluation et cela constituera une base pour les évaluations internes futures.

6) CROISSANCE, ÉVOLUTION ET DIFFUSION

La reproduction et le développement sont les objectifs ultimes de la plupart des innovations, afin de favoriser le changement à plus grande échelle et au quotidien, mais aussi de reproduire ces modèles dans d'autres contextes. Cela étendra l'impact positif de ces initiatives.

7) CHANGEMENT DES SYSTÈMES

Changer les systèmes est extrêmement complexe. Une bonne façon d'atteindre cette ampleur est d'obtenir plusieurs études indépendantes démontrant qu'un élément ne fonctionne pas uniquement dans les circonstances propres aux premières phases de test, mais qu'il est peut être efficace à plusieurs endroits.

ADOPTER DES IDÉES D'AUTRES RÉGIONS OU SECTEURS

Même si vous disposez de preuves sur l'adoption réussie d'une politique ou d'un projet ailleurs, est-ce que ces idées fonctionneront dans votre domaine ? Howard White donne un exemple de programme d'alimentation complémentaire (le projet intégré de nutrition au Bangladesh, BINP) qui n'a pas permis d'atteindre au Bangladesh les résultats d'un « autre programme, largement salué, *de nutrition intégré au Tamil Nadu en Inde* ». ²⁸ Dans le programme du Bangladesh, le taux de participation était 30 % inférieur à celui en Inde. Une évaluation contrefactuelle d'impact « a révélé que le programme n'avait eu aucun impact significatif sur la situation nutritionnelle, bien qu'un impact positif ait été observé sur les enfants souffrant de malnutrition grave. ». Le professeur White a souligné :

« Il faut cibler la bonne population. Ce programme était destiné aux mères de jeunes enfants, mais les mères ne prennent généralement pas de décisions (ou ne sont pas les seules à décider) en matière de santé et d'alimentation de leurs enfants. Dans les zones rurales du Bangladesh, les femmes ne vont pas au marché. Ce sont les hommes qui s'occupent des achats. Pour les femmes qui vivent en ménage commun, à savoir sous le même toit que leur belle-famille, comme c'est le cas pour

une forte minorité, ce sont les belles-mères qui prennent les décisions relevant du domaine féminin. Les taux de participation au projet sont considérablement plus faibles pour les femmes vivant avec leur belle-mère dans les régions les plus traditionalistes du pays ». ²⁸

Il observe également que, pour être efficace, le programme doit bien cibler les enfants. Dans les zones couvertes par le projet BINP, les professionnels en nutrition qui ont appliqué le programme « n'ont pas pu identifier correctement des graphiques [de croissance] quels enfants pouvaient bénéficier du programme », d'où un mauvais ciblage de la population concernée. ²⁸ Ces deux exemples de mise en œuvre inappropriée de programmes illustre l'importance d'une bonne compréhension des facteurs de contexte local, car ce sont ces facteurs qui conditionnent la réussite ou l'échec d'un programme, ainsi que la nécessité d'utiliser des preuves qualitatives et ethnographiques (voir section C) pour atteindre les résultats souhaités dans différents contextes.

Lors de la conception de nouveaux programmes destinés à changer le mode d'agissement des personnes (par exemple, en respectant le droit humanitaire international ou en acceptant l'usage de latrines plutôt que de déféquer ouvertement dans les espaces publics), les recherches en psychologie sociale, en anthropologie et sur les changements comportementaux peuvent être utiles. Cela pourrait fournir des idées s'appuyant sur les recherches approfondies réalisées sur les techniques qu'il serait bon d'adopter pour orienter les personnes vers les actions souhaitées.

Hugh Waddington et ses collègues, ²⁹ par exemple, ont noté que l'amélioration de l'assainissement (en incitant les personnes à utiliser des latrines plutôt que de déféquer ouvertement dans les espaces publics) exigeait davantage de choses que des connaissances sur la théorie des germes. Les résultats de l'évaluation réalisée par CLTS (Community Led Total Sanitation) suggèrent qu'une meilleure connaissance technique des latrines, l'utilisation de subventions à des fins de réduction des coûts pour les ménages, l'évolution des normes sociales et l'action collective au niveau des villages sont tous des facteurs clés permettant d'amener un changement de comportement et accroître le taux d'adoption des latrines.

CRÉATION D'UNE THÉORIE DU CHANGEMENT

Lors des toutes premières phases de mise au point d'une intervention, il est important de décrire de façon logique ce que vous faites et pourquoi c'est important. Vous devez être cohérent, clair et convaincant. Cette description est souvent appelée théorie du changement et vise à donner une « *explication claire, concise et convaincante de ce que vous faites, de l'impact que vous souhaitez avoir et de la façon dont vous pensez l'obtenir. Elle constitue une composante essentielle de tout programme et une condition préalable à la réalisation d'une évaluation* ». ³⁰

Une théorie du changement est un moyen utile d'être plus explicite au sujet des preuves que vous utilisez et d'être plus clair sur la façon dont vous allez obtenir des résultats. Par exemple, si vous souhaitez mettre en œuvre un programme visant à accroître l'utilisation du lavage des mains dans un camp de réfugiés, vous aurez peut-être besoin de preuves qui répondent à une série de questions. Quelles techniques sont susceptibles d'aboutir à un changement de comportement efficace? Les nouvelles installations de lavage des mains seront-elles acceptables pour les utilisateurs potentiels ?

Où devraient se situer ces installations de lavage des mains ? En quoi ces installations doivent-elles consister ? Quels renseignements doivent être fournis? Quelles mesures incitatives doivent être prises pour influencer les décisions des ménages ?

Une théorie du changement (exemple à la page 20) vous aide à être explicite au sujet de vos objectifs et de la façon dont vous comptez les atteindre. Elle permet d'éviter d'espérer simplement que votre approche novatrice fonctionnera et que vos hypothèses sont correctes. Un autre avantage de la théorie du changement réside dans le fait qu'il s'agit d'une première étape dans la conception d'une évaluation efficace, car elle tente d'identifier tous les résultats qui seront à mesurer.

Ici, le message le plus important est de penser à la pertinence. Ce que vous voulez, c'est identifier la recherche qui correspond à vos besoins et au stade de développement où vous vous situez. Les preuves dont vous aurez besoin lors de la planification d'un nouveau programme ou au cours des premiers mois de sa mise en œuvre seront très différentes de celles dont vous aurez besoin lorsque le programme sera davantage établi. Nous reviendrons sur cette question cruciale de la pertinence à la section C.

Messages clés de la Section B

✓ Il existe un large éventail de situations dans lesquelles des données probantes peuvent vous aider. Il ne s'agit pas seulement de celles qui sont évidentes en termes de saisie des résultats des interventions et des politiques. Il peut y avoir d'autres avantages auxquels vous n'avez pas pensé (création de campagnes plus persuasives, obtention de financement ou arrêt de quelque chose qui ne fonctionne pas, par exemple).

✓ Vous devez réfléchir au calendrier d'obtention et à la pertinence des données probantes. Où en êtes-vous dans la spirale de l'innovation et quelles données probantes pourraient le mieux répondre à vos besoins ? Par exemple, dans les premiers jours de l'élaboration d'un programme novateur, il peut être plus utile d'apprendre de ce que d'autres ont déjà trouvé, en examinant les évaluations passées ou des études sociales et scientifiques de plus grande ampleur, plutôt que de faire une évaluation importante et coûteuse de votre propre travail.

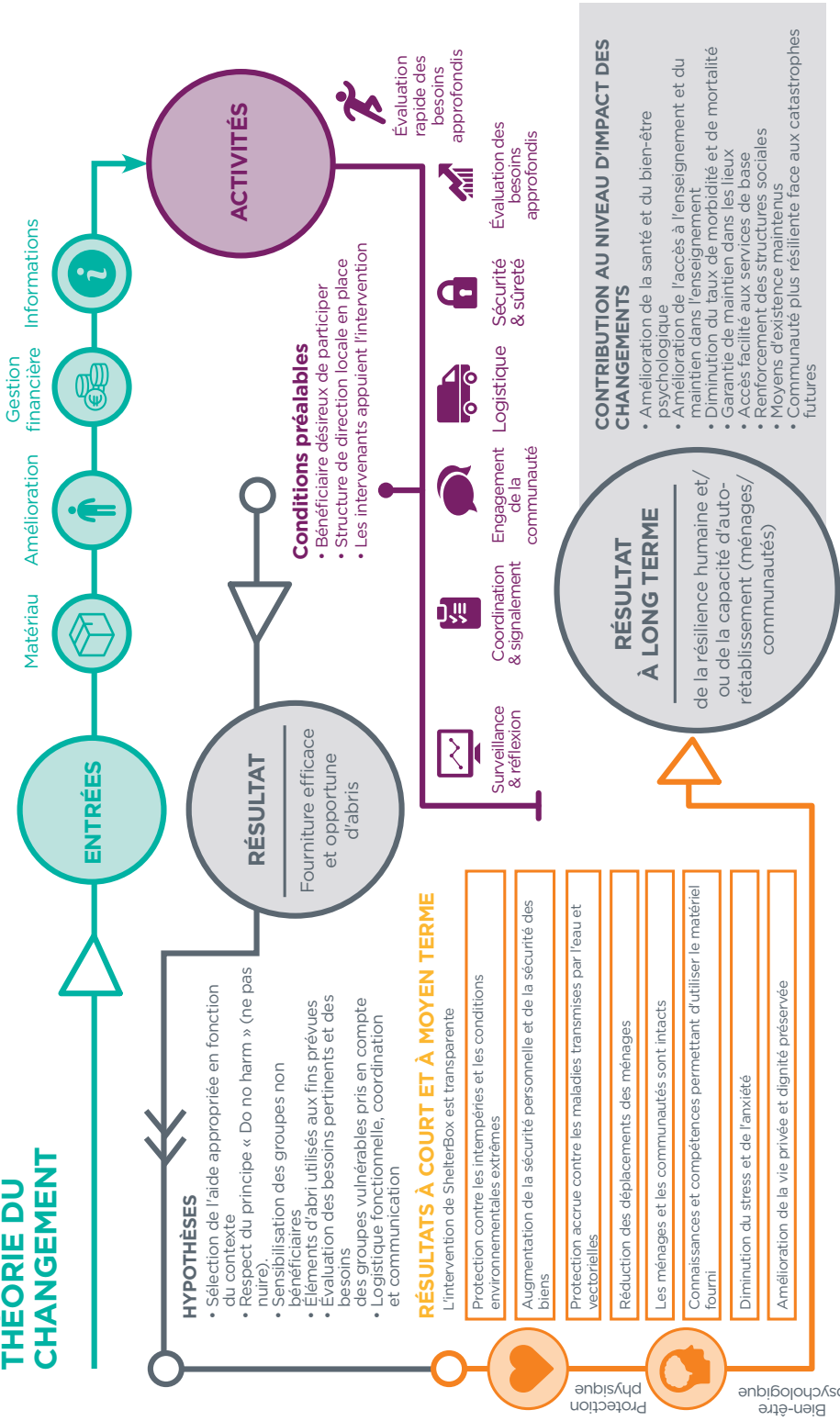
✓ Les premières étapes de l'innovation consistent à créer une théorie du changement. C'est une façon utile d'avoir une plus grande clarté quant aux preuves que vous utilisez et sur la méthode que vous allez appliquer pour obtenir des résultats. Cela vous aidera également à concevoir une évaluation efficace.

✓ Les interventions ou les politiques mieux établies permettent de mettre sur pied leurs propres expériences ou évaluations d'impact. Au fur et à mesure qu'une innovation prend de l'ampleur, vous devez également envisager des répétitions multiples pour vérifier que les avantages que vous avez identifiés ne constituent pas seulement une réussite isolée, mais qu'ils peuvent fonctionner dans d'autres lieux et contextes.

✓ Peu importe la taille de votre organisation. Il y a beaucoup de travaux de recherche qui pourraient vous aider.

Exemple de théorie du changement appliquée à la résilience des communautés

THÉORIE DU CHANGEMENT



'Outcomes and impacts - measuring the difference we make': <https://cwarham.wordpress.com/2017/02/02/shelterbox-theory-of-change>

Quelles données probantes devez-vous choisir ?

Cette section examine différents types de données probantes et examine comment choisir la plus appropriée à la question qui vous intéresse. Elle traite également de la façon de juger de la qualité des données probantes.

Toutes les preuves ne se valent pas.

L'un des messages clés de ce guide pratique consiste à réfléchir à l'utilité. Trop de personnes s'attirent des ennuis en ne réfléchissant pas clairement au type de plan de recherche, à la méthodologie ou aux méthodes de collecte de données qui répondront vraiment à leurs besoins. Elles ont tendance à « se laisser guider par les méthodes ». ³¹ En d'autres termes, nous avons tous tendance à choisir notre approche favorite (questionnaires, essais aléatoires ou interviews, notamment), parce que nous sommes plus à l'aise avec elle, plutôt que d'être plus ouverts à l'éventail des méthodes et de penser à celles qui sont les mieux adaptées pour relever notre défi et résoudre nos incertitudes.

Par exemple, la recherche expérimentale est plus adaptée aux évaluations d'impact et « de ce qui fonctionne » (c'est-à-dire de l'efficacité), mais d'autres types de recherche peuvent également apporter d'autres idées. ³² Il se peut qu'ils ne s'intéressent pas autant à l'impact, mais qu'ils soient capables de révéler pourquoi et comment les choses fonctionnent. ³³

De plus, il peut être important d'établir quels segments de la population bénéficient le plus d'une intervention, ce qui peut souvent être masqué si l'on prend comme référence l'effet moyen sur l'ensemble de la population.

ÉTUDE DE CAS

Leçons tirées de l'utilisation de l'ethnographie dans la réponse Ebola

L'épidémie d'Ebola en Afrique de l'Ouest a mis en lumière la nécessité d'encourager l'engagement communautaire dans la réponse aux épidémies, ainsi que le rôle joué par les réseaux de chercheurs en sciences sociales pour partager les informations et diffuser la réponse. C'est ce qu'illustre la Plateforme Ebola Response Anthropology, dont les membres « *travaillent de manière proactive avec les organisations sanitaires et humanitaires pour concevoir, fournir et surveiller des interventions et des travaux de recherche sur le terrain mieux adaptés et socialement informés au niveau local* ». ³⁴ Des leçons peuvent être tirées de la manière dont les anthropologues ont utilisé l'ethnographie pour apporter des dimensions socioculturelle et politique dans le cadre d'une réponse biomédicale, afin d'aborder la question des pratiques funéraires dites « dangereuses ». Dans un contexte où les cadavres transmettent le virus Ebola, des mesures exceptionnelles ont été adoptées : seuls certains travailleurs humanitaires ont été autorisés à enterrer les cadavres des personnes décédées des suites d'Ebola. Les rassemblements des communautés locales qui assistaient aux funérailles d'êtres chers

décédés des suites d'Ebola ont été perçus par les intervenants comme une résistance traditionnelle ³⁵ à la médecine moderne et comme une menace à la biosécurité. La participation d'anthropologues ayant une vaste expérience de la vie dans les pays touchés par l'épidémie a été essentielle pour identifier les pratiques acceptables et sûres (par exemple, montrer une photo du cadavre aux familles afin de répondre aux rumeurs concernant des corps disparaissant ³⁶), ainsi que pour surmonter les tensions découlant des pratiques funéraires, grâce à des approches collaboratives. L'une des principales leçons qui en est ressortie a été d'identifier les raisons pour lesquelles les communautés avaient besoin d'enterrer leurs morts plutôt que de considérer ces pratiques comme essentiellement « culturelles ». ³⁷ La réponse d'Ebola a montré « *la nécessité d'établir de nouveaux mécanismes mondiaux permettant de mobiliser rapidement tous les experts qui peuvent apporter des informations contextuelles, médicales, épidémiologiques et politiques locales pertinentes sur les urgences sanitaires mondiales* ». ³⁸

Le type de recherche que vous avez choisi comme source de données probantes pour vous aider à prendre vos décisions doit correspondre aux besoins du problème auquel vous faites face. ³⁹

La recherche expérimentale est plus adaptée aux évaluations d'impact et « de ce qui fonctionne ».

PARTIE 1 : DIFFÉRENTS TYPES DE MÉTHODES DE RECHERCHE, DE CONCEPTS ET D'APPROCHES

Alors, comment allez-vous comprendre quelle méthode de recherche est la mieux adaptée à votre cas ?

Si vous n'êtes pas chercheur, il peut être décourageant de vous confronter au jargon et aux listes interminables d'approches différentes.⁴⁰ Ces problèmes sont consignés dans une *note explicative du ministère britannique du Développement international* portant sur l'évaluation des preuves⁴¹ et sont tout aussi pertinents pour le secteur humanitaire que pour le développement international :

Certains types de recherche (tels que les conceptions expérimentales et

quasi expérimentales) conviennent mieux à l'identification de la présence d'un lien de causalité.

D'autres types de recherche sont plus appropriés pour expliquer de telles relations de causalité - voir Au-delà des expériences : nouvelles méthodes d'évaluation à la page 32.

Certaines conceptions (telles que les études d'observation) sont plus utiles pour comprendre les contextes politique, social et environnemental.

Certains des avantages et des inconvénients de divers modèles de recherche sont résumés dans le **Tableau C.1**.⁴²

Tableau C.1 *Différents modèles, méthodes et approches concernant les résultats de recherche - bref aperçu*

TYPES DE RECHERCHE ET D'ÉVALUATION	DE QUOI S'AGIT-IL ?	AVANTAGES	INCONVÉNIENTS
Ethnographie	Compte rendu par une personne ayant une expérience directe d'un problème particulier ou affectée par celui-ci. L'objectif est de recueillir des informations détaillées sur l'expérience et les perceptions des personnes.	Puissant et immédiat ; peut donner des informations détaillées et vivantes sur des événements dissimulés à une grande partie de la population.	Difficile à vérifier et peut entraîner une inflation de la prévalence. Les récits émotionnels à la première personne peuvent nuire à l'évaluation critique et les anecdotes individuelles peuvent ne pas être représentatives.
Étude de cas	Examen approfondi d'un seul ou de plusieurs cas tels qu'une communauté, une famille, une organisation, des sites, un événement ou un individu.	Facile et pratique. Peut être utilisé par des praticiens et des non spécialistes. Bon pour les interventions qui ont déjà eu lieu. Pourrait identifier les conséquences néfastes d'une intervention. Aide à décrire les innovations. Génère des hypothèses pour d'autres études.	Mauvais pour déduire qu'une intervention provoque un résultat. Une petite taille signifie qu'il est difficile de généraliser au niveau national ou de la population.

TYPES DE RECHERCHE ET D'ÉVALUATION	DE QUOI S'AGIT-IL ?	AVANTAGES	INCONVÉNIENTS
Études de cas-témoins	Compare un groupe qui a, par exemple, un problème de santé avec un groupe exempt de ce trouble, et remonte le temps pour déterminer en quoi les caractéristiques des deux groupes diffèrent.	Nécessite moins de ressources que les études d'intervention ; utile lorsque le recours aux essais randomisés ou aux études de cohortes prospectives ne sont pas pratiques (études sur la mort subite du nourrisson, par exemple) ; peut générer des hypothèses qui pourront être testées dans des études plus solides.	Rare en politique sociale (voir l'évaluation « au cas par cas » ci-dessous pour plus d'informations sur les approches plus courantes en politique sociale), plus fréquente en épidémiologie et en santé ; fournit moins de données probantes pour l'inférence causale qu'un essai randomisé ; risque élevé de divers préjugés (au niveau de la sélection et de l'interviewer, par exemple).
Évaluation du processus	Approche d'évaluation qui examine ce qui se passe par rapport aux théories existantes sur le changement ou par rapport aux voies causales identifiées au cours d'une évaluation. Peut être associé à une évaluation réaliste, à une analyse comparative qualitative (AQC), à une analyse de contributions, ou à un suivi de processus.	Efficace pour l'explication des causes ; peut être utilisé dans les domaines plus confus et complexes des interventions sociales, au cours desquelles il peut y avoir plusieurs causes et où le contexte revêt une grande importance. L'évaluation du processus examinera les mécanismes de changement, les facteurs contextuels et l'évaluation de la mise en œuvre.	Manque de méthodes convenues ; possibilités de partialité ; faible pouvoir absorbant ou impacteur ; relativement peu d'évaluation des techniques utilisées par rapport à la littérature abondante critiquant les essais randomisés.
Évaluation économique	Les évaluations économiques déterminent si une intervention constitue une utilisation efficace des ressources et comment elle se compare à d'autres interventions en termes de coûts et de conséquences. Dans le secteur humanitaire, les méthodes les plus courantes sont l'analyse coûts-avantages et l'analyse coût-efficacité.	L'évaluation économique donne une bonne idée du niveau de ressources nécessaires et des coûts par un rapport bénéfice/efficacité, ce qui est un excellent indicateur pour la plupart des décideurs et peut être utilisé pour le plaidoyer.	L'évaluation économique implique plusieurs hypothèses et parfois des mesures indirectes du coût, qui reposent sur des jugements. La mesure de l'efficacité doit être fournie par un essai comparatif robuste.

TYPES DE RECHERCHE ET D'ÉVALUATION	DE QUOI S'AGIT-IL ?	AVANTAGES	INCONVÉNIENTS
<p>En coupe transversale</p>	<p>Un échantillon représentatif de personnes interrogées à un moment donné. Bien que les enquêtes telles que les questionnaires et les entretiens structurés soient couramment utilisées dans la conception transversale, elles ne sont pas le seul moyen. D'autres méthodes incluent l'analyse du contenu ou l'analyse des statistiques officielles.</p>	<p>Les données quantitatives peuvent être examinées pour détecter des modèles ou des associations ; relativement bon marché et éthique ; l'enquête peut être répétée à intervalles réguliers, illustrant l'évolution des tendances dans le temps (voir Études de cohorte / longitudinales ci-dessous) ; peut générer des hypothèses qui peuvent être testées dans des études plus robustes.</p>	<p>Établit l'association au mieux, pas la causalité. Il est rarement possible d'attribuer un changement mesuré à l'intervention ou de comprendre ce qui se serait passé en l'absence de l'intervention, car tout changement aurait été dû à des problèmes plus vastes tels que les conditions économiques, la météo ou des campagnes médiatiques, plutôt que l'intervention. Les autres inconvénients sont le risque de biais de rappel, de biais de désirabilité sociale, de biais de chercheur ; taille inégale des groupes et distribution inégale des facteurs de confusion.</p>
<p>Études de cohortes/ longitudinales</p>	<p>Le même échantillon de personnes interrogées sur plusieurs points dans le temps, parfois de l'enfance à la vieillesse.</p>	<p>Meilleure source de preuves sur l'association entre l'expérience de l'enfance et les résultats chez l'adulte ; peut apporter un soutien puissant à certaines interventions précoces ; peut être utilisé pour évaluer les programmes mis en œuvre à la suite d'essais randomisés.</p>	<p>Les données arrivent souvent trop tard pour une prise de décision efficace ; les membres de l'étude peuvent abandonner avec le temps ; approche coûteuse lorsqu'elle est maintenue pendant des décennies.</p>

TYPES DE RECHERCHE ET D'ÉVALUATION	DE QUOI S'AGIT-IL ?	AVANTAGES	INCONVÉNIENTS
<p>Conception quasi expérimentale</p>	<p>Compare un groupe qui a, par exemple, un problème de santé avec un groupe exempt de ce trouble, et remonte le temps pour déterminer en quoi les caractéristiques des deux groupes diffèrent.</p>	<p>Peut fournir une preuve raisonnablement forte de la relation entre l'intervention et les résultats mesurés ; méthode puissante d'exploration de l'impact d'une intervention lorsque la randomisation est impossible ; peut être appliqué aux grandes communautés ainsi qu'aux groupes ; pas besoin de randomisation dès le départ (ex-ante), ce qui évite certains des défis et difficultés de la randomisation.</p>	<p>L'incapacité à garantir l'équivalence des groupes et à prévenir les changements au fil du temps peut entraîner des résultats moins fiables. Les techniques d'appariement nécessitent généralement beaucoup de données dans les groupes d'intervention et de comparaison, ce qui peut prendre beaucoup de temps et être coûteux à collecter. Une bonne compréhension des facteurs à associer (sans cela, il est possible qu'il existe des différences systématiques entre les groupes non contrôlés). Ces conceptions nécessitent un travail analytique complexe et des connaissances spécialisées.</p>

TYPES DE RECHERCHE ET D'ÉVALUATION	DE QUOI S'AGIT-IL ?	AVANTAGES	INCONVÉNIENTS
Essai randomisé	Un groupe reçoit une intervention tandis qu'un autre ne reçoit pas cette intervention ou reçoit une intervention différente. Les groupes sont formés en utilisant un processus aléatoire, avec généralement la même chance d'être alloués à l'un ou l'autre groupe.	Offre les résultats les plus robustes et les plus fiables, ce qui permet de croire que toute différence mesurée entre les groupes est due à l'intervention. La répartition aléatoire devrait éviter les différences systématiques entre les groupes ; une plus grande confiance dans l'ampleur de l'effet et la relation entre l'intervention et le résultat ; approche internationalement reconnue.	Faible prise en compte du contexte (par exemple, contextes culturel, institutionnel, historique et économique) ; difficile à faire au niveau de la population nationale ; utilisé à petite échelle, cela peut ne pas être pertinent au niveau national / de la population (bien que cela représente un risque pour toutes les conceptions). Il peut être difficile de manipuler des variables pour expérimenter une politique sociale (classe, race ou lieu de résidence, par exemple). Des erreurs de randomisation peuvent invalider les résultats ; il peut être difficile de persuader les décideurs des avantages de cette conception ; problèmes politiques et éthiques potentiels liés à la randomisation (par exemple, certains groupes bénéficient aléatoirement d'une intervention bénéfique, et pas d'autres) ; peut prendre plus de temps de gestion et de temps à mettre en place que des quasi-expériences.
Revue systématique, méta-analyses	Agrégation des résultats des études éligibles, avec les critères de recevabilité définis à l'avance et les méthodologies revues	Meilleure source d'assurance que l'intervention fonctionne (ou ne fonctionne pas) ; méta-analyse met en commun les résultats statistiques ; les grandes revues ont un pouvoir statistique considérable ; est reproductible par d'autres chercheurs ; peut être appliqué à tout type de données ou d'étude.	Nécessite un nombre suffisant d'études robustes dans un domaine donné ; méthodologie moins bien développée pour la synthèse de données qualitatives et de la littérature « grise » ; peut produire des résultats trompeurs si des biais de déclaration sélectifs affectent la recherche disponible pour la revue.

RECHERCHE EXPÉRIMENTALE - POURQUOI EN FAIRE TOUT UN PLAT?

Les recherches expérimentales telles que les essais randomisés ont suscité beaucoup d'attention - certains diront peut-être trop. Alors pourquoi ont-elles suscité autant d'intérêt ?

Les praticiens et les décideurs veulent souvent savoir si une intervention a eu un impact. Y a-t-il eu un lien de causalité entre leur nouveau programme et les « impacts », « effets » et « résultats » à la fin ? Ils veulent pouvoir attribuer les effets constatés à la politique ou à l'intervention testée. Les modèles expérimentaux, même s'ils sont parfois difficiles à mettre en œuvre dans la pratique, ont plus de chances de montrer ce lien de cause à effet.

Une étude expérimentale pourrait simplement tester et comparer deux groupes de personnes avant et après l'intervention donnée à l'un des groupes. Mais si ces groupes ne sont pas formés par randomisation, il existe un risque que les groupes diffèrent en raison du biais de sélection et que les différences dans leurs résultats soient alors dues à ces différences dans leurs caractéristiques et non aux effets de l'intervention testée. Par exemple, une personne qui demande l'intervention peut être plus prédisposée à des résultats positifs. Peut-être sont-ils plus motivés, en meilleure santé, confiants et donc plus susceptibles de réussir, quelle que soit leur intervention. D'autre part, une personne qui souhaite une intervention peut être la plus vulnérable et la plus susceptible de mal se porter indépendamment de toute intervention. Ces biais fausseront les résultats de l'expérience et pourraient signifier que celle-ci fournit une estimation erronée de l'effet de l'intervention. De plus, les biais peuvent être imprévisibles et si nous ne connaissons pas le sens du biais, nous ne pouvons pas simplement ajuster les résultats de l'expérience pour essayer de les prendre en compte.⁴³

Pour atténuer ces biais, nous pouvons utiliser un processus de randomisation : une fois que la population de l'étude a été identifiée et qu'ils ont accepté de participer à l'étude, les individus à comparer se voient attribuer de manière aléatoire les groupes d'étude qui seront comparés, par exemple par tirage au sort.⁴⁴

MODÈLES DE RECHERCHE EXPÉRIMENTALE

Le choix du plan expérimental approprié et la réalisation de l'étude établissent correctement le lien de causalité entre une intervention et une réalisation. Lorsque l'allocation aléatoire est utilisée pour affecter les participants à l'intervention ou au groupe de comparaison, il s'agit du moyen le plus robuste pour déterminer si une différence dans les résultats entre les groupes est due à l'intervention ou au hasard, et le hasard est minimisé en effectuant une étude assez vaste. L'allocation aléatoire minimise la possibilité de différences systématiques entre les groupes, qu'elles soient dues à des facteurs connus ou inconnus.

L'introduction d'un groupe de contrôle avec lequel comparer l'intervention minimise plusieurs biais qui compliquent normalement le processus d'évaluation. Par exemple, si vous introduisez un système de transferts monétaires pour augmenter les taux de vaccination et l'utilisation des livraisons en établissement, comment saurez-vous si ceux qui reçoivent l'argent supplémentaire n'auraient pas utilisé les services de toute façon ? Vous devez les comparer à un groupe de personnes qui ne reçoivent pas l'argent supplémentaire.

Comme pour toute conception de recherche, la force d'un essai randomisé nécessite également que l'étude soit menée correctement. Cela implique de veiller à ce que la randomisation d'individus ou d'unités (écoles, hôpitaux, quartiers, etc.) soit correctement effectuée (« fidélité d'allocation ») et que seul le groupe expérimental ait été exposé à l'intervention

testée (« fidélité au traitement »). Un essai randomisé bien conduit requiert également un nombre suffisant de participants et une puissance statistique suffisante pour que les autres facteurs (« facteurs de confusion ») qui auraient pu provoquer le résultat soient équilibrés entre le groupe expérimental et le groupe témoin. Un faible taux d'abandon (« attrition ») est également nécessaire pour minimiser le risque de biais. Il est donc important de vérifier les rapports sur un essai randomisé pour s'assurer que ces critères de validité ont été remplis.

Vous trouverez des informations supplémentaires sur la manière de déterminer si un essai randomisé a été mené et consigné correctement dans les instructions relatives à la conduite et aux rapports sur les essais.⁴⁵

L'Initiative internationale pour l'évaluation de l'impact (3ie) a identifié des évaluations d'interventions humanitaires en situation de conflit et de réconciliation qui ont utilisé des essais randomisés et des conceptions quasi expérimentales (voir ci-dessous).⁴⁶

ÉTUDE DE CAS



Utilisation de méthodes expérimentales pour déterminer l'efficacité des interventions humanitaires

Les maladies transmissibles sont un point particulièrement sensible dans les cas de populations affectées par des conflits ou catastrophes et résidant dans des camps. Lors la phase de réponse d'urgence, les maladies diarrhéiques seules représentent plus de 40 % des décès. Les maladies diarrhéiques peuvent résulter d'une variété de facteurs, mais dans un contexte d'urgence, un système d'assainissement et un approvisionnement en eau adaptés (en quantité comme en qualité) représentent le principal vecteur de réduction des diarrhées. Les technologies de traitement de l'eau sont actuellement limitées, et peu de produits sont à même de traiter une eau turbide.

Shannon Doocy et Gilbert Burnham ont présenté les conclusions d'une étude sur l'efficacité du traitement de l'eau sur place avec un floculant-désinfectant dans 400 ménages au sein de camps pour population déplacée à Monrovia au Liberia.⁴⁷

L'étude a comparé les taux de diarrhées chez les ménages utilisant un traitement de l'eau avec un floculant-désinfectant ainsi qu'un système de stockage de l'eau

amélioré (groupe d'intervention), et chez les ménages ayant seulement un stockage d'eau amélioré (groupe de contrôle). Les ménages ont été assignés au groupe expérimental ou au groupe témoin de manière aléatoire.

Dans le groupe d'intervention, la combinaison du traitement de l'eau au point d'utilisation avec un floculant-désinfectant avec un stockage amélioré a réduit l'incidence des diarrhées de 90% et leur prévalence de 83% comparativement au groupe de contrôle avec stockage de l'eau amélioré uniquement. La concentration de chlore résiduel du groupe d'intervention atteignait ou dépassait les standards Sphere dans 85% des observations (95% CI: 83.1-86.8) avec un taux de conformité à 95%.

Doocy et Burnham ont conclu que la floculation/désinfection au point d'utilisation réduisait significativement l'incidence et la prévalence des maladies diarrhéiques chez les ménages résidants dans des camps dans des contextes d'urgence.

QUESTIONS PRATIQUES ET ÉTHIQUES LIÉES AUX ÉTUDES RANDOMISÉES

Avoir un groupe de contrôle conduit à une critique commune des essais randomisés, à savoir qu'ils sont contraires à l'éthique car seule une partie de la population a accès à une intervention expérimentale qui pourrait être considérée comme supérieure. Ce « rationnement » de l'intervention est justifiable, dans le cas où nous ignorons si elle fera plus de bien que de mal, et si nous faisons l'essai pour le savoir. Cependant, s'il existe déjà des preuves convaincantes que l'intervention sera bénéfique, il faudra peut-être encore faire un essai randomisé pour déterminer son utilité et pour qui. Si tel est le cas, il existe des moyens de contourner la critique et le défi éthique. Par exemple, dans un essai contrôlé randomisé sur liste d'attente, les personnes affectées au groupe de contrôle se verront proposer l'intervention à la fin de l'essai ou à la fin de leur suivi personnel. Une autre option est la conception « en gradins » ou « en pipeline »,⁴⁸ suivant laquelle l'ordre dans lequel les participants reçoivent l'intervention est déterminé au hasard, ce qui permet de l'introduire progressivement pour que tous les participants y aient accès avant la fin de l'essai. Ce design a été utilisé pour l'étude Progres/Oportunidades, l'un des plus grands essais randomisés en matière de développement international, qui s'est intéressé aux transferts conditionnels de fonds dans la population rurale pauvre du Mexique.⁴⁹ Les études par étapes peuvent être très fonctionnelles dans un monde gagné par l'austérité, où il est difficile d'accéder à suffisamment de ressources pour déployer un nouveau programme en une seule fois. Elles permettent un déploiement graduel tout en gardant le bénéfice d'un essai randomisé pour en évaluer l'efficacité. Cependant, les études par étapes sont complexes et peuvent être difficiles à implémenter. De plus, une intervention qui s'avérerait être inefficace ou nocive aurait tout de même été appliquée à tous les participants à l'essai. Dans le cas d'interventions comportementales ou de restructuration des prestations de service, il peut alors être difficile voire impossible de revenir en arrière ou de « désactiver » l'intervention après l'essai.

LORSQUE LA RANDOMISATION EST IMPOSSIBLE OU INACCEPTABLE

Dans certaines circonstances, il peut être impossible ou inacceptable d'affecter aléatoirement des personnes à des groupes en vue de comparaison. Par exemple, il peut être impossible d'utiliser un essai randomisé pour comparer les effets relatifs de différents types de coordination des interventions d'urgence, ou l'impact de politiques nationales destinées à l'amélioration de l'intégration sociale des réfugiés.⁵⁰ Dans pareils cas, les chercheurs peuvent être amenés à utiliser des modèles quasi-expérimentaux pour examiner les liens entre interventions et résultats. C'est ce qu'ont fait Kayvan Bozorgmehr et Oliver Razum pour comparer les dépenses privées en matière de santé chez les demandeurs d'asiles et réfugiés d'Allemagne ayant un accès réduit ou bien normal aux soins.⁵¹ De même, Rodolfo Rossi et al. ont utilisé des questionnaires avant et après une campagne de vaccination pour mesurer ses effets sur la couverture vaccinale des jeunes enfants dans une zone de crise du Liban en 2015.⁵²

MODÈLES QUASI-EXPÉRIMENTAUX

Il existe une large variété de modèles quasi-expérimentaux permettant de comparer les effets d'interventions et ils sont souvent utilisés lorsque les essais randomisés ne sont pas faisables. Ce type d'étude est similaire aux essais randomisés dans la mesure où ils permettent de comparer les résultats entre individus et groupes bénéficiant d'une intervention ou programme et ceux n'en bénéficiant pas. Cependant, le fait de bénéficier ou non de l'intervention ou du programme n'est pas assigné aléatoirement. Au lieu de cela, on utilise alors d'autres types de modèles comparatifs, tels que les études de séries chronologiques interrompues, les études avant-après, ou bien des analyses utilisant la discontinuité par régression ou l'appariement des coefficients de propension.⁵³ Il convient d'observer que les conceptions de type avant-après sans groupe témoin ne peuvent établir l'impact d'une intervention ou d'un programme

donné puisqu'en l'absence de contrefactuel (c'est-à-dire un moyen de comparer avec ce qui ce serait passé sans l'intervention), tout résultat observé pourrait être dû à des facteurs autres que l'intervention elle-même. Les conceptions quasi-expérimentales présentent

donc un plus grand risque de partialité que des essais randomisés bien menés, mais elles peuvent tout de même permettre de dégager de fortes inférences causales dans des circonstances où les essais randomisés sont impossibles ou inacceptables.⁵⁴

ÉTUDE DE CAS



Déterminer l'efficacité des interventions humanitaires avec les modèles quasi-expérimentaux. Contribution de l'aide alimentaire à l'instauration de solutions durables dans les situations prolongées de réfugiés

Cette évaluation s'inscrit dans le cadre d'une série d'évaluations d'impact sollicitée conjointement par le Programme Alimentaire Mondial (PAM) et le Haut-Commissariat des Nations Unies pour les réfugiés (HCR) en 2011 et 2012 afin d'évaluer le rôle et la contribution de l'aide alimentaire en matière d'autonomie et de solutions durables pour les réfugiés et les populations d'accueil.⁵⁵

L'évaluation concernait la situation prolongée dans le Sud-Est du Bangladesh, où approximativement 30 000 réfugiés Rohingyas reçoivent de l'aide dans deux camps officiels depuis plus de vingt ans. Outre les réfugiés enregistrés, approximativement 45 000 Rohingyas non enregistrés résident dans des camps de fortune et plus de 150 000 résident dans des communautés d'accueil dans le district de Cox Bazar.

Pour évaluer l'impact aussi méthodologiquement et rigoureusement que possible que le permet le contexte, un modèle quasi expérimental a été utilisé. Le modèle d'évaluation portait principalement sur trois groupes de population :

1. Les réfugiés vivant dans les deux camps de réfugiés officiels ;
2. Les Rohingyas non enregistrés vivant sur deux sites non officiels ou dans des communautés d'accueil ;
3. Les populations d'accueil.

La principale comparaison quantitative portait sur les réfugiés enregistrés recevant de l'aide alimentaire et les réfugiés non

enregistrés n'en recevant pas. La principale question de l'évaluation était : Quels sont les impacts différentiels de l'aide alimentaire à long terme sur les différentes populations de réfugiés Rohingyas et populations d'accueil au Bangladesh ?

L'évaluation a conclu que tout comme d'autres formes d'aide extérieure, l'aide alimentaire influe sur les choix des ménages en matière d'activité économique et les stratégies de survie spécifiques qu'ils adoptent. Les réfugiés Rohingyas enregistrés s'engageaient dans des activités économiques significativement différentes de celles pratiquées par les réfugiés non enregistrés, occupant notamment des emplois plus qualifiés et moins risqués pour une rémunération moyenne supérieure. Leur situation économique en termes d'accumulation d'actifs était également significativement meilleure. L'aide alimentaire faisait partie intégrante de leurs moyens de subsistance, utilisée principalement pour la consommation et comme garantie et transfert de valeur pour les prêts et mises en gage. Le transfert de valeur de toute aide extérieure dans les camps permettait aux réfugiés de travailler moins et de se reposer sur cette assistance en période de crise.

Malgré ces différences, tous les groupes de Rohingyas réfugiés non enregistrés étaient tributaires d'une activité économique pour assurer leur subsistance. Les Rohingyas non enregistrés utilisaient un éventail plus large de mécanismes de survie, positifs comme négatifs, et constituaient une part importante de la main d'oeuvre de la région.

AU-DELÀ DE L'EXPÉRIMENTATION: DE NOUVELLES APPROCHES D'ÉVALUATION

En matière d'impact, il nous faut raisonner au-delà du seul recours aux expérimentations de type essais randomisés. Bien que les essais et les modèles quasi-expérimentaux fonctionnent bien lorsqu'il s'agit de tester des interventions simples, les interventions à tester sont souvent complexes. Les ONG travaillant avec la société civile en particulier, opèrent rarement seules et ont rarement la chance de manipuler des « traitements » expérimentaux univoques pouvant se prêter à un essai randomisé. Les évaluateurs sont à la recherche de manières alternatives de procéder qui impliquent de dépasser la notion de causalité bi-univoque et exigent une réflexion approfondie sur ce qui constitue votre « contribution ».

Dans le monde complexe de l'action humanitaire, il est peu probable que votre seul programme soit la condition nécessaire et suffisante de succès. Il peut n'être qu'un facteur parmi d'autres dans un « package de causalité ». Le succès d'un programme dépendra de tous ces autres facteurs environnants. Par le passé, une simple explication causale de « ce qui fonctionne » pouvait suffire mais désormais les chercheurs ont plutôt tendance à se demander : « est-ce que notre intervention fait la différence ? Qu'est-ce qui ce serait passé sans notre intervention ? »

Il y a toute une palette d'approches qui gagnent en popularité chez les évaluateurs, comme les modèles d'évaluations « basés sur la théorie »⁵⁶ et « basés sur les cas », la théorie de la complexité, l'évaluation réaliste,⁵⁷ l'analyse de contribution,⁵⁸ la reconstitution de processus⁵⁹ ou l'analyse comparative qualitative.⁶⁰ Ces approches ont en commun de reconnaître la complexité de la causalité, nécessitant une identification des mécanismes de causalité parmi une variété de cas ou événements spécifiques et des éléments de la chaîne de causalité fluctuants. Cela suppose « d'identifier les conditions et les autres causes en jeu, qui produisent, activent ou déclenchent les effets observés lors de

l'intervention ». ⁶¹ Par exemple, dans l'analyse de contribution, la recherche ne vise pas à démontrer qu'un facteur donné – par exemple une politique donnée – « produit » le résultat désiré, mais cherche au contraire à évaluer dans quelle mesure cette politique a contribué aux résultats observés. Elle reconnaît que les effets sont produits non par une seule cause, mais par plusieurs causes en même temps.

Ces approches ne sont pas nouvelles.⁶² En effet, la recherche d'une théorie fiable est sans aucun doute à la base de toute bonne démarche scientifique. Par conséquent, l'utilisation d'une approche « basée sur la théorie » n'est pas nouvelle, mais ces méthodes d'évaluation d'impact ont gagné en popularité, et elles permettent en effet aux évaluateurs de prendre en compte les causalités multiples.

EST-CE QUE CELA EN A VALU LA PEINE ? L'INTÉRÊT DES ANALYSES COÛT-AVANTAGE ET COÛT-EFFICACITÉ

Ayant déterminé et mesuré les effets d'une intervention, il est important pour le secteur humanitaire, comme pour tout autre secteur, de déterminer si ces programmes et interventions apportent un bon rapport qualité-prix. Lorsque les budgets sont serrés, il est nécessaire de pouvoir trancher en faveur des actions susceptibles de produire les bénéfices les plus importants. Les décideurs, commanditaires, organismes de subvention et organisations caritatives réclament une augmentation et une amélioration des mesures d'impact pour chaque livre, dollar ou euro dépensé. Les techniques d'évaluation économique permettent d'avoir cette vision sur la valeur des programmes.

Il existe tout une variété de techniques d'évaluation économique. Une analyse coût-efficacité (ACE) permet d'identifier la meilleure manière d'atteindre un résultat souhaité. Tandis que l'analyse coût-avantage (ACA) permet de savoir quelle est la meilleure utilisation des ressources disponibles. L'analyse coût-utilité (ACU) calcule quelle intervention produit le plus grand bien-être subjectif pour les populations cibles.¹²

À chaque fois que c'est possible, l'évaluation économique tente d'attribuer une valeur monétaire à chacune des causes ou activités qui contribuent à produire les effets qui à leur tour, mènent aux résultats souhaités ainsi que leurs impacts économiques, sociaux et environnementaux à long terme.

Tout ceci constitue l'analyse des coûts. Les mêmes procédures sont ensuite utilisées pour attribuer une valeur monétaire aux effets et impacts souhaités de la politique engagée. Les éventuels effets et impacts négatifs sont déduits de la valeur monétaire des effets et impacts positifs qui ont été atteints par ailleurs. Ce qui constitue l'analyse des bénéfices.

Les valeurs monétaires sont tirées dans la mesure du possible de l'activité du marché lorsqu'elle existe. Le marché du travail, par exemple, attribue une valeur monétaire aux différents types et unités de travail. Le marché du bâtiment fournit la valeur monétaire (le coût) des biens nécessaires à la construction d'un hôpital ou d'une école – briques, acier, verre, ciment, etc. Les coûts des soins médicaux peuvent être déduit des charges de fonctionnement fixes et variables des hôpitaux, centres médicaux, cliniques communautaires, mais aussi des coûts des médicaments, fournitures médicales, et ainsi de suite. C'est ce que l'on appelle la « monétisation » des coûts et bénéfices.

Certains coûts et (surtout) certains bénéfices ne sont pas négociés sur un marché et sont donc plus difficiles à monétiser. Dans ce type de cas, des économistes peuvent estimer les bénéfices (et parfois les coûts) en simulant une activité marchande à l'aide de sondages évaluant les consentements à payer et à

recevoir. Le consentement à payer (CAP) est défini comme « le montant maximum qu'un individu est prêt à payer en échange d'un bien ». ¹² De même, le consentement à recevoir (CAR) « est la somme minimum qu'un individu exigerait en compensation d'un bien ». ¹² Les résultats d'une enquête de CAP ou CAR fournissent ce que les économistes appellent les « préférences exprimées », c'est-à-dire ce que les gens déclarent être prêts à payer ou à recevoir en échange d'un bien ou d'un service. À défaut, les économistes tentent d'observer les comportements sur le marché réel. Autrement dit, ils créent une situation dans laquelle on demande à des individus de payer différentes sommes, ou de recevoir différentes sommes en échange d'un bien ou d'un service. Les comportements effectifs des individus sont ensuite enregistrés. On obtient ainsi ce que les économistes appellent les « préférences révélées ».

En plus des procédures mentionnées ci-dessus, l'évaluation économique permet également de savoir qui reçoit les bénéfices, qui supporte les coûts, et pour quelle durée. La valeur économique des bénéfices qui ne seront réalisés qu'à moyen ou long terme est actualisée grâce à ce qu'on appelle un taux d'actualisation. D'autres ajustements au calcul des coûts et bénéfices incluent les variations des prix relatifs, les différences matérielles en termes d'options fiscales, le biais d'optimisme, et la prise en compte des contingences de type inondations, sécheresses, ou les effets du changement climatique. L'évaluation économique est un travail assez technique, qui requiert habituellement l'expertise d'un économiste.

Les modèles quasi-expérimentaux peuvent tout de même permettre de dégager de fortes inférences causales dans des circonstances où les essais randomisés sont impossibles ou inacceptables.

Messages clés pour la partie 1 de la section C

- ✔ Toutes les preuves ne se valent pas. Certaines sont de meilleure qualité et seront plus appropriées aux défis que vous avez à relever.
- ✔ Le type d'étude doit être adaptée à vos besoins particuliers et la méthode utilisée doit avoir été appliquée et analysée de manière adéquate.
- ✔ Ne soyez pas découragé par la longue liste de méthodologies, modèles et approches de recherche. L'essentiel est de comprendre les hypothèses qui sous-tendent ces méthodes.
- ✔ Certains modèles d'évaluation (comme les modèles expérimentaux et quasi-expérimentaux) sont plus adaptés pour démontrer la présence d'un lien de causalité.
- ✔ D'autres méthodes de recherche sont plus appropriées pour expliquer comment se forment ces liens de causalité ou pour répondre à d'autres questions.
- ✔ Les évaluations et techniques basées sur la théorie telles que l'analyse des contributions gagnent en popularité auprès des évaluateurs. Ces techniques sont utiles lorsqu'une expérimentation s'avère difficile ou qu'il est impossible d'associer spécifiquement votre politique ou programme à un résultat clairement défini.
- ✔ Il ne s'agit pas seulement de déterminer si votre intervention fonctionne ou non, mais plutôt si elle a un « bon rapport qualité/prix ».
- ✔ Les techniques telles que l'analyse coûts-bénéfices peuvent vous aider à appréhender la valeur financière de l'impact d'une intervention.
- ✔ Il peut s'avérer nécessaire de déterminer quelles parties de la population bénéficient le plus d'une intervention. Ce qui peut souvent être occulté si on ne se fie qu'à l'effet moyen sur l'ensemble d'une population.

Il est important pour le secteur humanitaire de savoir si ses programmes et interventions sont optimales sur le plan économique.

PARTIE 2 : COMMENT ÉVALUER LA QUALITÉ D'UNE RECHERCHE ?

Une autre manière de choisir le type d'étude dont vous avez besoin est de vous demander : quelles études sont dignes de confiance ? De quel niveau de preuve ai-je besoin ? Est-ce que j'ai besoin d'un article publié dans une revue scientifique à impact élevé ? Et que dire d'une évaluation interne conduite par ma propre organisation ? Est-ce que son niveau de preuve serait valable ?

La publication dans une revue scientifique soumise à un comité de lecture peut vous donner confiance dans une étude, mais ce n'est pas la garantie d'une qualité supérieure. Dans un article célèbre, John Ioannidis de l'Université de Stanford aux États-Unis avait fait sensation en soutenant que « la plupart des conclusions publiées sont probablement fausses ». Il avait examiné les articles les plus fréquemment cités (au moins 1000 citations) dans certaines des revues médicales les plus prestigieuses au monde – provenant essentiellement de *The Lancet*, *le New England Journal of Medicine* et *le Journal of the American Medical Association*.⁶³ Parmi ceux alléguant une valeur thérapeutique positive et dont les résultats avaient été testés lors d'études ultérieures, 41% se sont avérés être erronés ou avoir un impact bien inférieur à ce que suggérait l'étude originelle.

EN QUOI L'EXAMEN PAR LES PAIRS AIDE LES DÉCIDEURS ?

Nous devons également être conscients du fait que l'examen par les pairs, qui est un fondement de la publication scientifique, est loin d'être parfait. Il peut y avoir des partis pris inconscients, comme le mimétisme (ou comportement grégaire), par lequel le comportement des examinateurs est influencé par le comportement de leurs pairs. Et le principe même des revues scientifiques peut être faussé par le biais de publication : des résultats positifs ont plus de chances d'être publiés,⁶⁴ là où les

résultats négatifs ont tendance à rester « au fond du tiroir ». Par exemple, une étude effectuée par des chercheurs de l'Université de Stanford a constaté que la plupart des études aboutissant à un résultat non significatif ne sont jamais publiées : seulement 20% avaient été publiées, et pour 65% d'entre elles, aucun article n'avait même été rédigé.⁶⁵ Cela peut avoir de graves conséquences pour les décideurs examinant les données. Si vous n'avez jamais eu connaissance des résultats négatifs d'une nouvelle intervention, vous pourriez conclure - à tort, que tout est pour le mieux et que l'intervention est efficace. Malgré ces problèmes, l'examen par les pairs reste pour beaucoup le modèle de référence lorsqu'il s'agit de décider ce qui mérite d'être publié dans la littérature scientifique : un garde-fou contre les travaux mal effectués.⁶⁶

Se baser sur un article ou une étude qui a fait l'objet d'un examen indépendant par d'autres experts permet d'avoir une certaine confiance dans la recherche effectuée. Bien qu'il faille parfois des années avant qu'une étude puisse être publiée (ce qui peut être trop long si vous devez prendre une décision rapidement), l'examen par les pairs apporte à la littérature scientifique une forme de contrôle de la qualité qui inspire la confiance.

DÉFINIR UNE RECHERCHE « DE HAUTE QUALITÉ »

L'examen par les pairs procure un minimum d'assurance. Mais que faire si vous devez inclure des données qui n'ont pas été vérifiées par d'autres experts ? Comment décider si elles constituent une base décisionnelle suffisante ?

Ce serait clairement une erreur d'ignorer des données issues de la recherche simplement parce qu'elles n'ont pas été publiées dans des revues scientifiques soumises à comité de lecture.

Ou de passer à côté du formidable gisement d'évaluations provenant de donateurs ou d'ONG qui ne sont jamais publiées dans des revues académiques, mais sont seulement disponibles dans la « littérature grise ». Il existe de nombreuses définitions de la littérature grise, mais ce terme se réfère généralement à la documentation non publiée ou bien publiée sans examen par les pairs. Elle peut également se référer à de la recherche qui est toujours en cours ou bien en attente de publication. Les rapports gouvernementaux, déclarations de principes, documents d'orientation politique, actes de conférences, font également partie de la littérature grise. La littérature grise est importante car elle peut contenir des résultats de recherche non concluants ou rapporter des interventions infructueuses, ce qui est important pour une prise de décision informée basée sur des preuves équilibrées. On peut consulter la littérature grise via des bases de données électroniques comme Open Grey (www.opengrey.eu), des actes de conférence et des dossiers financiers émanant d'organismes de financement de la recherche. Les sites webs d'organisations ayant un intérêt ou une expertise dans un domaine particulier peuvent être une autre source de littérature grise.

Ces organisations peuvent être contactées pour identifier des chercheurs et décideurs ayant une expertise dans le domaine recherché. Le mieux lorsqu'on cherche à évaluer la qualité de la preuve, est de commencer par définir ce qu'on entend par qualité. L'un des problèmes est que des expressions telles que « qualité », « standards », « robustesse », « biais » ou « solidité » sont souvent utilisés comme s'ils étaient interchangeables, et sans clairement les définir.

Ce qui génère beaucoup de confusion. Par exemple, pour certaines directives⁶⁷ la «

qualité » de la recherche se réfère à l'utilisation de modèles et méthodes particuliers, comme les essais randomisés. Cette accent mis sur la minimisation des partis pris comme moyen de garantir la qualité prend son origine dans une approche médicale et clinique de la qualité de la preuve, comme le système GRADE⁶⁸ ou l'échelle de mesure Maryland Scientific Methods Scale.⁶⁹ Ces approches de l'évaluation de la qualité sont habituellement basées sur la validité interne des études, leur validité externe et la qualité de la présentation des résultats de la recherche.

La qualité se réfère à la manière dont les études ont été conduites, présentées et analysées,⁷⁰ ainsi qu'à l'intégrité des chercheurs qui ne doivent pas détourner ni falsifier leurs données.⁷¹ Pour certains, la qualité peut aussi être liée à la pertinence d'une étude au regard des politiques et pratiques concernées.⁷²

Lorsqu'on cherche à déterminer une causalité, il est nécessaire de se demander si le modèle de recherche utilisé pour une étude permet effectivement de déterminer la causalité et si le modèle a été correctement mis en oeuvre. Une évaluation d'impact de bonne qualité permet d'attribuer une causalité : de montrer que l'intervention est la cause des résultats. Cela requiert une comparaison avec un groupe de contrôle qui soit aussi similaire que possible au groupe d'intervention à tous point de vue en dehors de l'intervention elle-même. Si c'est bien le cas et que l'étude a été correctement conduite, il en ressort une confiance élevée dans le lien de causalité entre votre intervention et les effets sur, par exemple, la prévention de la violence, la réduction du stress familial, ou un retour plus rapide à l'emploi. Il est également important de se demander si les effets observés lors de l'étude peuvent être reproduits ailleurs. C'est ce qui alimente la demande pour les méthodes mixtes de recherche et d'évaluation, ou à tout le moins pour une partie de recherche qualitative.

Ce serait clairement une erreur d'ignorer des données issues de la recherche simplement parce qu'elles n'ont pas été publiées dans des revues scientifiques soumises à comité de lecture.

RECHERCHE QUALITATIVE DE HAUTE QUALITÉ

Des données qualitative de bonne qualité⁷³ peuvent souvent s'avérer précieuses pour les décideurs politiques. La crédibilité scientifique des données qualitatives est aussi bonne que celle des données quantitatives, et une combinaison des deux est susceptible d'aider les décideurs à utiliser des données probantes sur les effets d'une intervention pour décider de l'applicabilité des données en fonction de leur situation.

L'essentiel est que les données issues de la recherche soient adaptées aux problèmes que vous avez à résoudre sur le terrain. Elles doivent être issues d'une recherche qui a utilisé des méthodes appropriées, collecté et analysé les données (qualitatives comme quantitatives) selon des méthodes clairement définies et reproductibles, et présenté ses conclusions de manière transparente en veillant à minimiser les biais.

Comme pour la recherche quantitative, il existe des standards et des recommandations portant sur la qualité des évaluations et recherches qualitatives.⁷⁴ Alors que la qualité des études expérimentales est habituellement basée sur la validité interne, la validité externe et une présentation adéquate des résultats, la qualité d'une étude qualitative dépendra typiquement de son caractère contributif, justifiable, rigoureux et crédible. Une étude contributive fait progresser les connaissances et la compréhension concernant une certaine politique, théorie, ou un domaine particulier. Une étude est justifiable si sa stratégie de recherche est pertinente au regard des questions d'évaluation posées. La rigueur de l'étude se réfère au caractère systématique et transparent de la collecte, de l'analyse et de l'interprétation des données qualitatives. Sa crédibilité se réfère au caractère bien fondé et plausible des arguments relatifs aux données probantes générées.

UTILISER LES REVUES SYSTÉMATIQUES POUR ÉVITER LE « CHERRY PICKING » (LA SÉLECTION DE DONNÉES ALLANT TOUTES DANS LE MÊME SENS EN IGNORANT LES AUTRES)

Les décideurs doivent examiner la qualité de l'ensemble des données probantes disponibles,

et non seulement de certaines données choisies. Ils doivent au contraire utiliser une agrégation des données issues de la recherche. D'ailleurs, une synthèse fidèle de l'information existante est considérée comme « *la contribution la plus importante des chercheurs au processus décisionnel* ». ⁷⁵ L'essentiel de la réflexion sur la qualité entreprise ci-dessus a porté sur la « recherche primaire », autrement dit des études individuelles, qu'ils s'agissent d'expérimentations, de questionnaires ou de séries d'entretiens. Mais il est important de considérer les « synthèses de recherche », et le travail consistant à rassembler et récapituler une série d'études individuelles. L'idée n'est pas nouvelle, et le concept d'une prise de décision sur la base des données probantes accumulées est utilisé depuis des siècles.⁷⁶

Notre discussion précédente sur les biais cognitifs a montré que les gens sont susceptibles de tomber dans le piège inconscient qui consiste à rechercher des données qui confirment nos croyances. Il est donc nécessaire lorsqu'on rassemble la recherche existante de veiller particulièrement à inclure également les études dont les conclusions ne nous plaisent pas. Une façon d'éviter le « cherry picking » est d'utiliser ce qu'on appelle des revues systématiques.⁷⁷ Elles se veulent exhaustives, et visent à recueillir un maximum d'études portant sur la question posée et pertinentes pour la résolution des problèmes à résoudre sur le terrain. Elles utilisent des méthodes précises pour déterminer quelles conclusions fiables peuvent être tirées des études existantes et évaluer les méthodes et modèles utilisées par ses études pour en déterminer la qualité.

Quelques unes des principales caractéristiques⁷⁸ d'une revue systématique sont :

Une question de recherche clairement formulée, des objectifs et des critères pour la sélection des études.

Une méthodologie explicite et reproductible pour minimiser les biais.

Une recherche systématique pour identifier un maximum d'études remplissant les critères de sélection.

Une évaluation formelle de la validité des conclusion des recherches incluses.

Une présentation systématique ainsi qu'une synthèse des caractéristiques et conclusions des études incluses.

ÉTUDE DE CAS



Utilisation de revues systématiques pour développer l'encadrement du personnel.

Le Comité international de secours (International Rescue Committee, IRC) s'appuie énormément sur les preuves issues de la recherche dans le cadre du développement de la documentation et outils d'orientation des programmes à l'intention des équipes de terrain. Ceci, en réalisant des revues de la recherche croisant un grand nombre de sources de grande qualité portant sur des interventions ou approches spécifiques, et en utilisant les connaissances ainsi acquises sur les impacts, contextes et conditions des interventions pour décider de l'opportunité et de la manière d'adapter ces interventions à des contextes spécifiques.

Ce qui n'est pas sans difficultés car la majorité des études disponibles ne contiennent pas certaines informations critiques comme la fidélité de la mise en oeuvre, les mécanismes de causalité ou les spécificités liées au genre.

Cet effort généralisé pour assurer une disponibilité des données à destination de leurs équipes est réalisé par le biais du cadre logique d'intervention de l'IRC (Outcomes and Evidence Framework oef.rescue.org). Sur cette plateforme électronique accessible au public, l'IRC détaille les effets et sous-effets examinés dans le cadre de leur travail, les théories générales du changement (ou parcours) permettant d'atteindre ces objectifs et des indicateurs pour les mesurer.

Pour chaque effet et sous-effet, l'IRC a synthétisé les meilleures données quantitatives portant sur l'efficacité des interventions visant à modifier ou améliorer l'effet ou sous-effet concerné, en se concentrant tout particulièrement sur des données issues de revues systématiques. Pour les sujets sur lesquels une revue systématique n'existe pas encore, l'IRC a identifié et fait la synthèse d'évaluations d'impact individuelles.

Les approches systématiques ont également le mérite d'explicitier la manière dont elles recherchent des études et décident de leur inclusion ou non. Ce qui permet, en théorie,⁷⁹ à d'autres de reproduire cette revue systématique.

En 2018, Christl Donnelly et coll. ont suggéré quatre principes pour aider les chercheurs, décideurs et autres à commander, effectuer, partager, évaluer et utiliser les synthèses de données, notamment les revues systématiques.⁸⁰ Les synthèses de données devraient, selon eux, être inclusives, rigoureuses, transparentes et accessibles, critères qu'ils précisent comme suit :

Inclusive

Associe les décideurs et pertinent pour eux.

Inclut de nombreux différents types et sources de données.

Implique des compétences et personnes variées.

Rigoureuse.

Utilise un ensemble de données le plus complet possible.

Reconnaît les partis pris et veille à les minimiser.

Est examinée de manière indépendante dans le cadre d'un processus d'assurance de la qualité.

Transparente

Décrit clairement la question de recherche, les méthodes, les sources de données et le processus d'assurance de la qualité.

Identifie les difficultés et les points litigieux.

Signale les hypothèses, restrictions et incertitudes, y compris tout type de lacunes dans les données probantes.

Indique les intérêts personnels, politiques et organisationnels et gère les éventuels conflits qui en découlent.

Accessible

Est rédigée dans un langage clair.

Est disponible pour une durée appropriée.

Est librement accessible en ligne.

MÉTA-ANALYSE

La méta-analyse représente une part importante des revues systématiques pour lesquelles tout ou partie des études incluses sont suffisamment similaires statistiquement (« homogènes ») en termes de population, intervention, moyens de comparaison et résultats pour être rassemblées et agrégées. Cela aboutit à une estimation cumulée des effets, dont le risque de parti pris est généralement inférieur à celui des études individuelles prises isolément, et dont la précision est plus grande. Les méta-analyses sont souvent représentées par des graphiques en forêt⁸¹ comme celui de la **Figure C.1**, issue d'une revue systématique récente sur l'impact des programmes de soutien envers les populations affectées par des urgences humanitaires. Les méta-analyses sont peut-être plus connues pour combiner les résultats d'essais randomisés, mais elle peuvent également être utilisées pour des données provenant d'autres types d'études, comme les études cas-témoin, transversales ou de cohorte.

LES GRAPHIQUES EN FORÊT

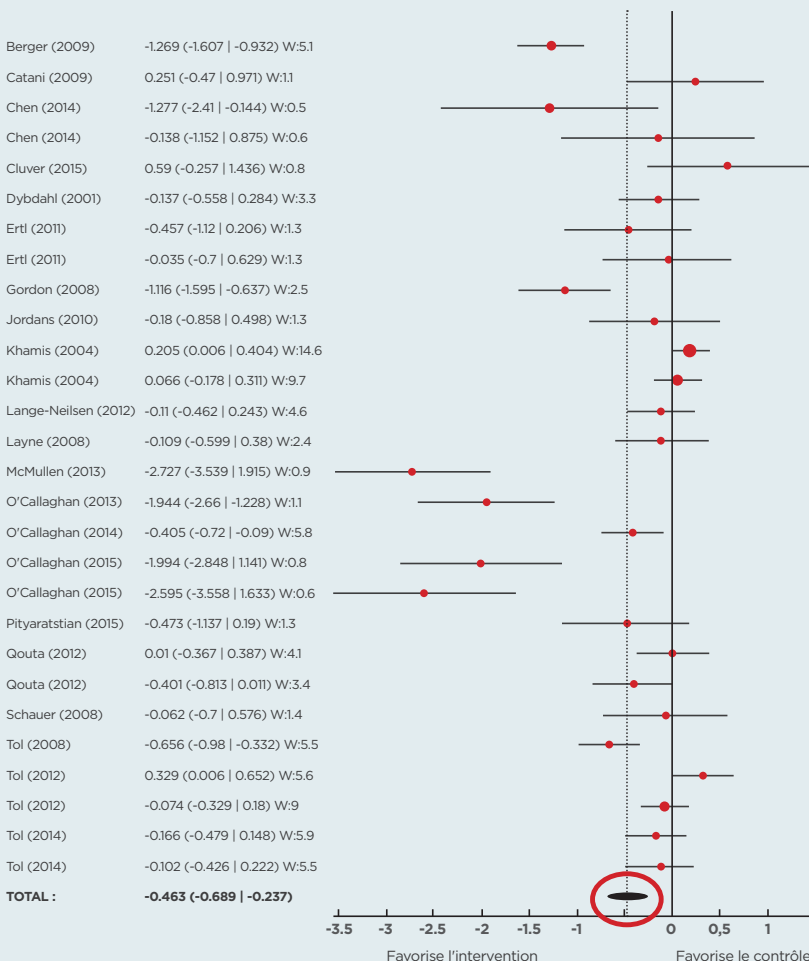
La **Figure C.1** présente les conclusions de 28 évaluations d'impact comparant des personnes en situation d'urgence humanitaire ayant bénéficié de programmes de soutien psychosocial ou de santé mentale, avec d'autres n'ayant pas bénéficié de ce type d'intervention.⁸² Dans ce graphique en forêt, les données présentent les effets des programmes sur les syndromes de stress post-traumatique

Figure C.1: Graphique en forêt sur les effet de programmes de soutien psychosociaux ou de santé mentale sur les syndromes de stress post-traumatique⁸²

Mesure : Continu : d (Hedges g)

Hétérogénéité : Q=206; df=27; p=0; I²=86.9%; tau-carré=0.29

Modèle à effets aléatoires : -0.463 (-0.689, -0.237)



(SSPT) comme une variable continue. Pour chaque étude figurant dans le graphique, le point rouge représente l'effet moyen de l'intervention, et les lignes horizontales traversant les points rouges représentent l'intervalle de confiance pour chaque étude. La ligne pleine verticale partant de zéro sur l'axe horizontal indique une absence d'effet du programme, et les résultats de toutes les études sont rassemblées pour calculer l'effet estimé global des programmes. Cette nouvelle valeur est illustrée par le losange noir (entouré de rouge) dans le bas du graphique. Il représente *l'effet estimé cumulé* issu du regroupement et de l'agrégation des effets moyens et de leur variance pour les 28 évaluations d'impact incluses dans la revue. Il nous permet de conclure qu'en moyenne les programmes de soutien de santé mentale et psychosociale ont un léger effet positif sur le SSPT comparativement à une absence d'intervention.

ÉVALUATION RAPIDE DES DONNÉES

La préparation de revues systématiques et de méta-analyses peut prendre beaucoup de temps. Cela signifie que si une étude systématique à jour n'est pas disponible, et qu'une décision doit être prise rapidement, on peut avoir besoin de rechercher soi-même les données pertinentes puis les évaluer et les synthétiser plus rapidement qu'on ne le ferait normalement pour une revue systématique dans les règles de l'art. Heureusement, si les revues systématiques sont déjà nombreuses, leur nombre continue d'augmenter rapidement, et dans la section D, nous montrons comment les trouver ainsi que d'autres collectes disponibles. Cependant, si vous ne trouvez pas ce que vous cherchez parmi les revues systématiques existantes, ou que vous trouviez des revues caduques, vous serez peut-être amené à commander une revue systématique « réduite », comme une évaluation rapide des données probantes.⁸³ Ces revues rapides prennent habituellement 1 à 3 mois, et sont planifiées de manière à répondre aux besoins des décideurs et praticiens qui n'ont pas le temps d'attendre

une revue systématique complète. Elles utilisent la même structure de base et les mêmes étapes qu'une revue systématique complète, mais ne sont pas aussi intensives, exhaustives, ou complètes. Elles sont amenées à prendre des raccourcis dans la recherche, l'évaluation critique, l'extraction et collecte de données, ainsi que l'analyse statistique des études incluses. Les conclusions sont également présentées de manière plus courte et moins détaillée que pour une revue systématique complète. Elle peuvent être inférieures à 25 pages, avec un résumé analytique de 3 pages, et un document d'information d'une page à l'intention des décideurs. Les limites des évaluations rapides de données par rapport aux revues systématiques sont leur manque d'exhaustivité et de détail, et leur plus grande tendance aux parti pris. Par conséquent, il est nécessaire de faire preuve de plus de prudence lorsqu'on est amené à baser une décision sur une évaluation rapide de données que pour une revue systématique complète.

Malgré ces limitations, elles sont fréquemment commandées et utilisées par les décideurs et responsables de mise en oeuvre de programmes, en particulier lorsque le temps presse et qu'une évaluation systématique n'est pas disponible.

Un type particulier d'évaluation rapide, appelée une évaluation rapide des besoins en recherche, peut également être utilisé pour identifier rapidement les lacunes dans les données probantes. Au Royaume-Uni, l'Équipe de soutien rapide pour la santé publique comprend un plan en commun avec Evidence Aid pour la mise en oeuvre de ce type d'évaluation, afin d'identifier les doutes importants pouvant être abordés par la recherche dès les premiers stades d'une urgence humanitaire associée à une épidémie.

L'IMPORTANCE DE LA RÉPÉTITION ET DE LA CORROBORATION

Le concept de lacunes dans les données probantes nous amène à un autre facteur à considérer pour évaluer la qualité d'une

synthèse d'études : le nombre d'études à inclure pour être confiant que les données sont suffisamment solides. Est-ce que deux ou trois études pourraient suffire si elles sont vraiment bonnes ? Ou bien est-ce qu'il en faut des douzaines, ou même des centaines ? La vérité est qu'il n'y a pas de nombre magique en la matière. Pourtant, nous ne pouvons ignorer le fait que la quantité de données est un facteur important : la force réside dans le nombre et la répétition et la corroboration sont nécessaires.

Même les études les plus reconnues doivent être répétées. Amgen, une société pharmaceutique californienne, a tenté de reproduire 53 études décisives sur le cancer. Leurs résultats, publiés dans *Nature*,⁸⁴ ont confirmé seulement 11% des conclusions de ces études marquantes. C'est particulièrement préoccupant car ces études ont influencé des sociétés pharmaceutiques et des traitements anticancéreux dans le monde entier et l'incapacité à les reproduire pourrait indiquer que leurs conclusions ne sont pas fiables. À l'inverse, il n'est pas nécessaire que ces études soit répétées encore et encore si les données sont déjà suffisamment robustes.⁸⁵ Les études

inutiles sont un gaspillage des capacités de recherche⁸⁶ voire, dans le contexte d'essais randomisés, contraires à l'éthique si des patients continuent de participer à des interventions sous-optimales.

REVUE DES REVUES : ÉVALUER DES ENSEMBLES DE DONNÉES DE HAUTE QUALITÉ

Il reste encore à évaluer la qualité du modèle et de la mise en oeuvre de l'examen des données, or la qualité d'une synthèse est déterminée par la qualité des études incluses. Même la meilleure revue ne peut pas extraire une réponse de bonne qualité d'une recherche mal effectuée.

Il existe également des outils officiels, comme AMSTAR, pour évaluer la qualité des revues systématiques.⁸⁷ On peut aussi utiliser l'une des premières checklists destinées à l'évaluation de la qualité des revues systématiques,⁷⁸ qui posaient une série de questions :

La question est-elle clairement posée ?

La recherche d'études pertinentes est-elle approfondie ?

Les critères d'inclusion étaient-ils pertinents ?

La validité des études utilisées a-t-elle fait l'objet d'une évaluation adéquate ?

A-t-on récupéré d'éventuelles informations manquantes auprès des chercheurs initiaux ?

Dans quelle mesure les résultats obtenus sont-ils dépendants de la manière dont la revue est conduite ?

Les analyses de sous-groupes sont-elles interprétées avec précaution ?

Les conclusions découlent-elles bien des données examinées ?

Les recommandations sont-elles liées à la solidité des données probantes ?

Les jugements sur les préférences (valeurs) sont-ils valides ?

Un *Comment noter* sur l'évaluation de la force des preuves produites par le ministère britannique du Développement international énumère quatre éléments à prendre en compte lors de la vérification des éléments de preuve :⁴¹

La qualité (technique) des études constituant le corpus de preuves (ou la mesure dans laquelle le risque de biais a été communiqué).

La taille de l'ensemble des preuves.

Le contexte dans lequel la preuve est définie.

La cohérence des résultats produits par les études constituant l'ensemble des preuves.

Est-ce que l'absence avérée d'effet est bien différenciée du résultat non significatif ?

L'arbitrage de tous ces critères sera toujours en partie subjectif et doit prendre en compte le contexte de la question posée et de l'intervention considérée. Il est difficile de donner des conseils d'ordre général sur ce qui peut constituer un ensemble de données suffisants à l'application d'une politique donnée.

On commence également à voir apparaître des regroupements de conclusions provenant de multiples revues dans ce que l'on appelle les « revues de revues ». Bien sûr, elles ne sont possibles que lorsqu'il existe déjà un volume suffisant de revues systématiques – et dans certains domaines il y en a peu. Quoi qu'il en soit, ces revues de revues fournissent une évaluation et une synthèse des preuves probantes issues de multiples revues systématiques concernant soit la même intervention, soit des interventions différentes dans le même domaine.⁸⁸

Messages clés pour la partie 2 de la section C

✓ Pour trouver des données probantes dignes de confiance, cherchez des études examinées par les pairs. Mais soyez conscient que l'examen par les pairs est loin d'être parfait. Il peut y avoir des biais inconscients, comme le mimétisme (ou comportement grégaire), par lequel le comportement des examinateurs est influencé par le comportement de leurs pairs.

✓ Lorsqu'il s'agit d'études d'impact et d'examiner « ce qui fonctionne », utilisez les cadres et standards officiels en matière de données probantes, comme ceux de Nesta ou autres.

✓ Une étude n'est jamais suffisante. Évitez de prendre des décisions basées sur des études isolées et cherchez plutôt des études qui ont été reproduites

plusieurs fois. La force est dans le nombre, la répétition et la corroboration sont nécessaires. Même les études les plus reconnues doivent être répétées, mais il faut éviter de continuer à répéter les études lorsque l'ensemble des données probantes est déjà suffisamment robuste.

✓ Utilisez les revues systématiques, qui visent l'exhaustivité et filtrent les études en fonction de leur qualité – généralement sur la base de leur modèle et leurs méthodes.

✓ Si vous n'avez pas le temps de réaliser une revue systématique et ne pouvez pas en trouver une existante qui réponde à votre question, envisagez de réaliser ou commander une évaluation rapide des données probantes.⁸⁹

Même l'examen le mieux effectué ne peut transformer une recherche de faible qualité en une réponse de haute qualité.

Où devriez-vous chercher des preuves ?

Cette section examine différents types de données probantes et examine comment choisir la plus appropriée à la question qui vous intéresse. Elle traite également de la façon de juger de la qualité des données probantes.

RECHERCHE DE RECHERCHE PREUVES

Avec l'augmentation rapide du nombre de personnes et d'organisations qui travaillent aux examens systématiques, la documentation contient maintenant un nombre important et sans cesse croissant d'examen systématiques, de sorte que vous pourriez être en mesure d'en trouver un qui répond à vos besoins. On estime récemment à plus de 200 000 le nombre d'examen systématiques dans tous les domaines thématiques. Bien sûr, seule une petite proportion de ces études serait pertinente pour le secteur humanitaire, mais (comme nous l'indiquons ci-dessous) des ensembles de ces études sont disponibles gratuitement dans des collections en ligne telles que celles conservées par Evidence Aid. Il existe également de grandes organisations internationales, telles que Cochrane (www.cochrane.org) et Campbell Collaborations (www.campbellcollaboration.org), qui se consacrent à la préparation et à la conservation des examens systématiques et les mettent en ligne.

Cependant, si vous avez besoin de faire vos propres recherches pour étudier les effets de l'action humanitaire, le point de départ est de reconnaître qu'elles ne sont pas toujours faciles à trouver et que vous pourriez avoir besoin de l'aide d'un bibliothécaire ou d'un spécialiste en information qui sera chargé de vous conseiller et, peut-être, d'aider à concevoir et mener la recherche. Les preuves sont éparpillées dans des dizaines de milliers de rapports répartis dans des

milliers de revues, de livres et de sites Web; beaucoup d'entre eux ne sont pas libres d'utilisation ou sont difficiles à consulter. Lorsque vous réfléchissez à l'endroit où effectuer une recherche, il est important de tenir compte de l'endroit où les preuves qui vous intéressent sont susceptibles d'avoir été publiées et indexées, et du degré d'exhaustivité que vous souhaitez atteindre. Par exemple, s'ils étaient publiés dans une revue scientifique, les articles pourraient être disponibles dans l'une des milliers de bases de données bibliographiques électroniques. Il s'agit, par exemple, de PubMed pour les soins de santé, LILACS pour les articles présentant un intérêt particulier pour l'Amérique du Sud ou ERIC pour la littérature éducative. Il peut également être nécessaire d'identifier la littérature grise, comme les rapports des gouvernements ou des ONG, et les recherches présentées lors de conférences, qui nécessiteront des recherches dans des bases de données électroniques telles que Open Grey (www.opengrey.eu) et sur les sites Web des organisations qui pourraient avoir effectué des recherches pertinentes.

Lors de la planification d'une recherche, il peut être utile de diviser les éléments de recherche en types d'intervention ou de politique qui vous intéressent, en population ou milieux pertinents, en mesures des résultats qui seraient les plus utiles à votre prise de décision et en types d'études que vous souhaitez trouver. Vous pouvez ensuite décider lequel ou lesquels de ces domaines est le plus important et le plus utile pour identifier le matériel le plus pertinent pour vous.

Vous devriez considérer les types de sources à rechercher, ce qui pourrait inclure :

Bases de données bibliographiques (p. ex. PubMed, LILACS, ERIC).

Sites Web de revues et de conférences (p. ex. ceux qui se concentrent particulièrement sur les catastrophes ou l'action humanitaire).

Registres de recherche (tels que via le portail de l'OMS pour les essais, www.researchregistry.com et PROSPERO pour les revues systématiques).

Collections en ligne de données de recherche (voir ci-dessous, p. ex. ALNAP, 3ie).

Site Web des organisations concernées (par exemple pour les rapports des gouvernements et des ONG).

Références dans les articles.

Correspondance avec les chercheurs et les évaluateurs.

Lorsque vous choisissez les termes à inclure dans votre recherche, vous devriez considérer :

Synonymes de différents temps et lieux.

D'autres mots et expressions qui sont liés à ce qui vous intéresse.

Une formulation plus large.

Une formulation plus étroite ou plus ciblée.

Index des termes ou mots-clés attribués aux articles par les auteurs originaux ou la base de données bibliographiques.

Enfin, vous devez décider si vous appliquerez des restrictions en fonction de la langue et de la période au cours de laquelle la recherche a été effectuée ou publiée.

SOURCES D'EXAMENS SYSTÉMATIQUES ET D'AUTRES RECHERCHES POUR LE SECTEUR HUMANITAIRE

En raison de ces défis dans la recherche de données probantes, vous pourriez trouver plus efficace et plus facile d'utiliser des dépôts fiables de recherches et d'examen systématiques, plutôt que de vous fier à des recherches aléatoires sur l'Internet. Heureusement, plusieurs organisations s'efforcent maintenant de rassembler des preuves pertinentes dans des ressources en ligne. Elles comprennent les éléments suivants :

Le Réseau d'apprentissage actif pour la responsabilité et la performance (ALNAP) est hébergé par l'Overseas Development Institute (ODI) à Londres, Royaume-Uni, avec un site Web contenant plus de 15 000 ressources, y compris des publications ALNAP relatives aux preuves et des détails sur plusieurs milliers d'évaluations.

Les bases de l'Evidence Aid ont été établies au sein de Cochrane à la suite du tsunami de l'océan Indien en 2004, lorsque le texte intégral de plusieurs dizaines d'études Cochrane relatives aux interventions en cas de catastrophe a été mis en ligne gratuitement. Il a été créé en tant qu'organisation caritative en 2015 et a maintenant rassemblé plusieurs centaines d'études systématiques qui sont toutes disponibles gratuitement sur son site Web (www.evidenceaid.org), y compris, mais sans s'y limiter, des collections spéciales d'études concernant les tempêtes de vent, tremblements de terre, Ebola, la santé des réfugiés et demandeurs d'asile, les troubles post-traumatiques et la prévention de la malnutrition aiguë durant les urgences et crises humanitaires.

L'Initiative internationale pour l'évaluation de l'impact (3ie) a été créée en 2008 et offre maintenant quatre bases de données consultables en ligne (www.3ieimpact.org). Deux d'entre elles, la 3ie Database of Systematic Reviews et la Database of

Impact Evaluations cataloguent les preuves de l'efficacité des interventions dans le secteur humanitaire. Ces bases de données comprennent également des examens systématiques et des évaluations d'impact sur le cadre plus large du développement international, dont beaucoup sont pertinents pour les interventions dans les situations d'urgence.

Nous tenons à jour une liste plus complète de ces types de ressources en ligne, à : www.evidenceaid.org/online-collections-of-research-for-the-humanitarian-sector

La liste fournit un large éventail de ressources de recherche en ligne, dont beaucoup sont gratuites et faciles d'accès. Celles-ci devraient être utiles à tout décideur politique, ONG ou professionnel de première ligne dans le secteur humanitaire, en fournissant un accès facile à des données fiables et de haute qualité sur l'efficacité des interventions. Si vous souhaitez suggérer des ressources additionnelles pour cette liste, veuillez contacter Evidence Aid : info@evidenceaid.org

CONCLUSION

En conclusion de ce guide sur l'utilisation des preuves dans le secteur humanitaire, nous vous encourageons à tirer parti des résumés de recherche disponibles gratuitement, accessibles et exploitables, tels que les revues systématiques contenues dans les sites internet que nous avons énumérés. Cela vous aidera à passer rapidement aux sources de preuves qui devraient vous aider à éclairer votre politique et votre pratique.

Messages importants pour la section D

✓ Les preuves de la pertinence de la recherche pour le secteur humanitaire sont dispersées dans des dizaines de milliers de rapports répartis dans des milliers de revues, d'ouvrages et de sites internet.

✓ La recherche de ces preuves peut nécessiter l'aide d'un spécialiste de l'information et la conception de la recherche doit tenir compte de ce que l'on cherche, des sources à rechercher,

des termes à utiliser dans la recherche et de toute restriction relative à la langue ou à la période de temps.

✓ Les examens systématiques fournissent des résumés et des synthèses des données de recherche.

✓ Plusieurs répertoires de recherche en ligne sont disponibles, ce qui améliore l'accès aux examens systématiques et à d'autres types de données de recherche.

Vous trouverez peut-être qu'il est plus efficace et plus facile d'utiliser des recueils fiables de recherches et d'examens systématiques.

Notes en fin de document

- ¹ Collins S, Sadler K. (2002) Outpatient care for severely malnourished children in emergency relief programmes : a retrospective cohort study. *Lancet* 360:1824-30.
- ² OMS, Programme alimentaire mondial, Comité permanent de la nutrition du système des Nations Unies et UNICEF. (2007) Community-based management of severe acute malnutrition : a joint statement by the WHO, World Food Programme, UN System Standing Committee on Nutrition and UNICEF. New York : UNICEF.
- ³ Price AI, Djulbegovic B. (2017) What does evidence mean? La plupart des langues traduites « preuve » par « démonstration ». *Journal of Evaluation in Clinical Practice* 23(5):971-3.
- ⁴ Sackett D, et al. (1996) Evidence based medicine: what it is and what it isn't. *British Medical Journal* 312:71. Voir www.bmj.com/content/312/7023/71.
- ⁵ >La définition de la médecine fondée sur des données probantes est venue à un moment où la plupart des décisions médicales étaient fondées sur l'expérience, l'autorité et l'excellence. La pratique médicale n'était pas fondée sur les meilleures preuves scientifiques disponibles. Certains commentateurs et chercheurs ont soutenu que la politique sociale est au même endroit qu'il y a 20 ou 30 ans, c'est-à-dire que c'est l'autorité, plutôt que les données de recherche, qui domine la prise de décision.
- ⁶ Allen C, et al. (2016) Evidence Aid. *Oxford Public Health* August: 51-54; Blanchet K, et al. (2017) Evidence on public health interventions in humanitarian crises. *Lancet* 390:2287-96; and Christoplos I, et al. (2017) Strengthening the quality of evidence in humanitarian evaluations. ALNAP Method Note. Londres : ALNAP/ODI (see www.alnap.org/system/files/content/resource/files/main/alnap-eha-method-note-5-2017.pdf).
- ⁷ Knox Clarke P, Darcy J. (2014) Insufficient evidence? The quality and use of evidence in humanitarian action. ALNAP Study. London : ALNAP/ODI. Voir www.alnap.org/system/files/content/resource/files/main/alnap-study-evidence.pdf.
- ⁸ www.oxforddictionaries.com/definition/english/evidence.
- ⁹ Nutley S, et al. (2013) What Counts as Good Evidence? Londres, Royaume-Uni : Alliance for Useful Evidence.
- ¹⁰ Hallam A, Bonino F. (2013) Using Evaluation for a Change: Insights from humanitarian practitioners. ALNAP Study. London : ALNAP/ODI. Voir www.alnap.org/system/files/content/resource/files/main/alnap-study-using-evaluation-for-a-change.pdf.
- ¹¹ Ariely D. (2009) Predictably Irrational: The Hidden Forces that Shape Our Decisions. London, UK: HarperCollins.
- ¹² HM Treasury. (2011) The Green Book: Appraisal and Evaluation in Central Government. London, UK: HM Treasury.
- ¹³ Christoplos I. (2006) Links between Relief, Rehabilitation and Development in the Tsunami Response: A Synthesis of Initial Findings. Stockholm, Suède : Agence suédoise de coopération dans le développement international.
- ¹⁴ Zsombok CE, Klein G, (editors). (2014) Naturalistic decision making. New York, USA: Psychology Press.
- ¹⁵ Kahneman D, et al. (2009) Conditions for intuitive expertise: a failure to disagree. *American Psychology* 64:515-26. Voir www.ncbi.nlm.nih.gov/pubmed/19739881.
- ¹⁶ These estimates come from Cash Learning Partnership (CaLP). (2018) The State of the World's Cash Report. Oxford: CaLP; Doing Cash Differently: Report of the High Level Panel on Cash Transfers. (2015); et Development Initiatives. (2017) Global Humanitarian Assistance Report.
- ¹⁷ ODI, Development Initiatives. (2016) Counting Cash: Tracking Humanitarian Expenditure on Cash-Based Programming.
- ¹⁸ Oxfam. (2006) Good Practice Review 11: Cash Transfer Programming in Emergencies. Oxford : Oxfam.

- ¹⁹ Jackson R (Save the Children UK), Kukrety N (Oxfam GB) (2012) Institutionalising cash transfer programming. Voir <https://odihpn.org/magazine/institutionalising-cash-transfer-programming/>.
- ²⁰ Voir www.humanitarianresponse.info/en/programme-cycle/space.
- ²¹ On trouvera dans Carr Z et al. une discussion sur les défis que posent la prise de décisions à la suite des rares circonstances d'une situation d'urgence majeure en matière de rations alimentaires. (2016) Using the GRADE approach to support the development of recommendations for public health interventions in radiation emergencies. *Radiation Protection Dosimetry* 171:144-55; et Ohtsuru A, et al. (2015) Nuclear disasters and health : lessons learned, challenges, and proposals. *Lancet* 386: 489-97.
- ²² Bradley DT, et al. (2014) The effectiveness of disaster risk communication: a systematic review of intervention studies. *PLOS Currents Disasters* August 22; Edition 1.
- ²³ Obrecht A. (2017) Using Evidence to Allocate Humanitarian Resources: Challenges and Opportunities (2017) traite de l'importance de prêter attention à la façon dont les donateurs recueillent, utilisent et partagent les preuves et les informations. Document de travail de l'ALNAP. London : ALNAP/ODI. Voir <https://reliefweb.int/sites/reliefweb.int/files/resources/alnap-ear-resource-allocation-2017.pdf>.
- ²⁴ Nesta. (2013) Understand how innovation works. Vidéo disponible sur www.nesta.org.uk/resources/understand-how-innovation-works.
- ²⁵ Obrecht A, Warner AT. (2016) More than just luck: Innovation in humanitarian action. HIF/ALNAP Study. London : ALNAP/ODI.
- ²⁶ Jones G, et al. (2003) How many child deaths can we prevent this year? *Lancet* 362:65-71. Voir [https://linkinghub.elsevier.com/retrieve/pii/S0140-6736\(03\)13811-1](https://linkinghub.elsevier.com/retrieve/pii/S0140-6736(03)13811-1).
- ²⁷ Nyhan B, Reifler J. (2010) When corrections fail: The persistence of political misperceptions. *Political Behavior* 32:303-30.
- ²⁸ White H. (2009) Theory-based impact evaluation: principles and practice, working paper 3. Delhi, Inde : International Initiative for Impact Evaluation, page 4.
- ²⁹ Waddington H, et al. (2009) Water, Sanitation and hygiene interventions to combat childhood diarrhoea in developing countries: a systematic review. Delhi, Inde : International Initiative for Impact Evaluation.
- ³⁰ Nesta and TSIP. (2014) Guidance for Developing a Theory of Change for Your Programme. Voir www.nesta.org.uk/sites/default/files/theory_of_change_guidance_for_applicants_.pdf.
- ³¹ Stern E. (2015) Impact Evaluation: A Design Guide for Commissioners and Managers of International Development Evaluations in the Voluntary and Community Sector. Londres, Royaume-Uni : Big Lottery Fund, Bond, Comic Relief and the Department for International Development.
- ³² Gerdin M, et al. (2014) Optimal evidence in difficult settings: improving health interventions and decision making in disasters. *PLoS Medicine* 11(4): e1001632.
- ³³ Mulgan G. (2015) The six Ws: a formula for what works. Londres : Nesta. Voir www.nesta.org.uk/blog/six-ws-formula-what-works.
- ³⁴ www.ebola-anthropology.net/about-the-network.
- ³⁵ Faye SL. (2015) L' «exceptionnalité» d'Ebola et les «réticences» populaires en Guinée-Conakry. Réflexions à partir d'une approche d'anthropologie symétrique. *Anthropologie & Santé*. Voir <https://journals.openedition.org/anthropologiesante/1796>.
- ³⁶ <http://pubman.mpdl.mpg.de/pubman/item/escidoc:2096578/component/escidoc:2103624/AAA-Ebola-Report-1.pdf>.
- ³⁷ Fairhead J. (2016) Understanding Social Resistance to the Ebola Response in the Forest Region of the Republic of Guinea: An Anthropological Perspective. *African Studies Review* 59:7-31. doi:10.1017/asr.2016.87.
- ³⁸ Abramowitz S, et al. (2015) Social science intelligence in the global Ebola response. *Lancet* 385:330.
- ³⁹ Petticrew M, et al. (2003) Evidence, hierarchies and typologies: horses for courses. *Journal of Epidemiology and Community Health* 57:527-9.

- ⁴⁰ Une liste exhaustive et faisant autorité des cadres et méthodes de recherche en sciences sociales est disponible dans Luff R, et al. (2015) Review of the Typology of Research Methods within the Social Sciences. London, UK: ESRC/ National Centre for Research Methods. Voir le lien <http://eprints.ncrm.ac.uk/3721>.
- ⁴¹ Ministère du développement international. (2014) How to Note: Assessing the Strength of Evidence. Voir www.gov.uk/government/uploads/system/uploads/attachment_data/file/291982/HTN-strength-evidence-march2014.pdf.
- ⁴² Adapté de HM Treasury (trésorerie de Sa Majesté), DECC et DEFRA. (2012) Quality in policy impact evaluation: understanding the effects of policy from other influences. London, UK: HM Treasury/DEFRA/DECC; Frost S, et al. (2006) The Evidence Guide; Using Research and Evaluation in Social Care and Allied Professions. London, UK: Barnardo's; Petticrew M, Roberts H. (2003) Evidence, hierarchies and typologies: horses for courses. *Journal of Epidemiology and Community Health*. 57: 527-9; et Stern E. (2015) Impact Evaluation; A Design Guide for Commissioners and Managers of International Development Evaluations in the Voluntary and Community Sector. London, UK: Big Lottery Fund, Bond, Comic Relief and the Department for International Development, tableau 2, page 18.
- ⁴³ Odgaard-Jensen J, et al. (2011) Randomisation to protect against selection bias in health care trials. *Cochrane Database of Systematic Reviews* (4):MR000012.
- ⁴⁴ White H. (2013) An introduction to the use of randomised control trials to evaluate development interventions. *Journal of Development Effectiveness* 5(1):30-49.
- ⁴⁵ Voir par exemple : Schulz KF, et al. (2010) CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *PLoS Medicine* 7(3):e1000251; Higgins JP, et al. (2011) The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 343:d5928.
- ⁴⁶ Puri J, et al. (2015) What methods may be used in impact evaluations of humanitarian assistance? working paper 22, Delhi, Inde: International Initiative for Impact Evaluation. Voir www.3ieimpact.org/media/filer_public/2014/12/08/wp_22_humanitarian_methods_working_paper-top.pdf.
- ⁴⁷ Doocy S, Burnham G. (2006). Point-of-use water treatment and diarrhoea reduction in the emergency context: an effectiveness trial in Liberia. *Tropical Medicine and International Health* 11:1542-52.
- ⁴⁸ White H, et al. (2014) Randomised Controlled Trials (RCTs). Methodological Briefs. Impact Evaluation No. 7. Florence, Italy: Unicef. Voir www.unicef-irc.org/publications/pdf/brief_7_randomised_controlled_trials_eng.pdf.
- ⁴⁹ Gertler P. (2000) Final Report: The Impact of PROGRESA on Health. Washington, DC, USA: International Food Policy Research Institute (IFPRI). Voir www.ifpri.org/publication/impact-progres-health.
- ⁵⁰ Hainmueller J, et al. (2017) Catalyst or crown: does naturalization promote the long-term social integration of immigrants? *American Political Science Review* 111(2):256-76. doi:10.1017/S0003055416000745.
- ⁵¹ Bozorgmehr K, et al. (2015) Effect of restricting access to health care on health expenditures among asylum-seekers and refugees: A quasi-experimental study in Germany, 1994-2013. *PLoS ONE* 10: e0131483.
- ⁵² Rossi R, et al. (2016) Vaccination coverage cluster surveys in Middle Dreib - Akkar, Lebanon: comparison of vaccination coverage in children aged 12-59 months pre- and post-vaccination campaign. *PLoS ONE* 11(12): e0168145.
- ⁵³ Alexander J, Bonino F (2015). A discussion on designs, approaches and examples, ALNAP Discussion Series Improving the quality of EHA evidence, Method Note 4, January 2015. Voir www.alnap.org/system/files/content/resource/files/main/alnap-eha-method-note-addressing-causation-jan2015.pdf.
- ⁵⁴ Waddington H, et al. (2017) Quasi-experimental study designs series paper 6: risk of bias assessment. *Journal of Clinical Epidemiology* 89:43-52; and Bärnighausen T, et al. (2017) Quasi-experimental study designs series—paper 7: assessing the assumptions. *Journal of Clinical Epidemiology* 89:53-66.

- ⁵⁵ Nielsen NS, et al. (2013). The Contribution of Food Assistance to Durable Solutions in Protracted Refugee Situations; its impact and role in Bangladesh: A Mixed Method Impact Evaluation, Volume I-Evaluation Report, Geneva, Switzerland: World Food Program/UNHCR.
- ⁵⁶ Carter R. (2012) Helpdesk Research Report: Theory-based evaluation approach. Université de Birmingham, Royaume-Uni : Birmingham: Governance and Social Development Resource Centre. Voir www.gsdr.org/docs/open/hdq872.pdf.
- ⁵⁷ Westhorp G. (2014) Realist Evaluation: An Introduction. Londres, Royaume-Uni : Methods Lab, Overseas Development Institute. Voir www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/9138.pdf.
- ⁵⁸ Mayne J. (2012) Contribution Analysis: Coming of Age? Evaluation 18(3):270-80.
- ⁵⁹ Pour l'analyse du réseau et la traçabilité des processus, voir : Befani B, et al. (2014) Process tracing and contribution analysis: a combined approach to generative causal inference for impact evaluation. IDS Bulletin 45(6):17-36.
- ⁶⁰ Baptist C, et al. (2015) Coffey How To Guide: Qualitative Comparative Analysis – A Rigorous Qualitative Method for Assessing Impact. Voir www.coffey.com/assets/Ingenuity/Qualitative-Comparative-Analysis-June-2015.pdf.
- ⁶¹ Stern E, et al. (2012) Broadening the range of designs and methods for impact evaluations, London, UK: Department for International Development.
- ⁶² Par exemple, voir le guide de Carol Weiss: Weiss C. (1998) Evaluation: methods for studying programs and policies. New York, USA: Prentice Hall.
- ⁶³ Ioannidis JA. (2005) Contraindications and Initially Stronger Effects in Highly Cited Clinical Research. Journal of the American Medical Association. 294(2):218-28.
- ⁶⁴ Pour un examen systématique des biais de publication, voir Hopewell S, et al. (2009) Publication bias in clinical trials due to statistical significance or direction of trial results. Cochrane Database of Systematic Reviews (1):MR000006. doi: 10.1002/14651858.MR000006.pub3.
- ⁶⁵ Franco A, et al. (2014) Social science. Publication bias in the social sciences: unlocking the file drawer. Science 345:1502-5.
- ⁶⁶ Sense about Science. (2006) I don't know what to believe: Making sense of science stories. Voir www.senseaboutscience.org/resources.php/16/i-dont-know-what-to-believe.
- ⁶⁷ Voir, par exemple, le supplément au Guide du Magenta : HM Treasury, DECC and DEFRA (2012) Quality in policy impact evaluation; understanding the effects of policy from other influences. Les lignes directrices montrent comment « des modèles de recherche de meilleure qualité peuvent aider à relever le défi d'attribuer les résultats mesurés à la politique en question (par opposition à d'autres influences), alors que des modèles de moindre qualité réduisent la confiance dans le fait que c'est la politique qui a obtenu ces résultats » (page 5).
- ⁶⁸ <http://gradeworkinggroup.org>.
- ⁶⁹ Farrington RP, et al. The Maryland Scientific Methods Scale, In: Farrington DP, et al. Evidence-based crime prevention. London, UK: Routledge, 2002, chapitre 2.
- ⁷⁰ Pour un modèle d'approches différentes en matière de qualité qui comprend les quatre dimensions de (1) la qualité méthodologique (2) la qualité des rapports, (3) la pertinence et (4) la pertinence par rapport aux politiques et aux pratiques ; voir Boaz A, Ashby D. (2003) Fit for purpose? Assessing research quality for evidence based policy and practice. Working Paper 11. London, UK: ESRC UK Centre for Evidence Based Policy and Practice.
- ⁷¹ Callaway E. (2011) Report finds massive fraud at Dutch universities. Nature 479(15). Voir www.nature.com/news/2011/111101/full/479015a.html.
- ⁷² Boaz A, et al. (2003) Fit for purpose? Assessing research quality for evidence based policy and practice. Working Paper 11. London, UK: ESRC UK Centre for Evidence Based Policy and Practice.
- ⁷³ La recherche qualitative explore et tente de comprendre les croyances, les expériences, les attitudes, le comportement des personnes et leurs interactions. Elle génère des données non numériques qui peuvent être recueillies par le biais, par exemple, d'entretiens approfondis, de groupes de discussion, d'analyses documentaires et d'observations des participants.

- ⁷⁴ Spencer L, et al. (2002) Quality in Qualitative Evaluation: A framework for assessing research evidence. London, UK: Cabinet Office; CASP. (2018) CASP Checklist for Qualitative Research. Oxford, UK: Critical Appraisal Skills Programme; National Institute for Health and Clinical Excellence. (2012) The guidelines manual: Appendix H: Methodology checklist: qualitative studies. London, UK: National Institute for Health and Clinical Excellence; et O'Brien BC, et al. (2014) Standards for reporting qualitative research: a synthesis of recommendations. *Academic Medicine* 89: 1245-51.
-
- ⁷⁵ Whitty CJM. (2015) What makes an academic paper useful for health policy? *BMC Medicine* 13:301.
-
- ⁷⁶ Voir, par exemple, la discussion sur l'importance d'utiliser les données existantes lors de la conception de nouvelles études, dans Clarke M. (2004), Doing new research? Don't forget the old: nobody should do a trial without reviewing what is known. *PLoS Medicine* 1: 100-2; et l'historique des examens systématiques dans Chalmers I, et al. (2002) A brief history of research synthesis. *Evaluation and the Health Professions* 25: 12-37; et Clarke M. (2016) History of evidence synthesis to assess treatment effects: personal reflections on something that is very much alive. *JLL Bulletin: Commentaries on the history of treatment evaluation. Journal of the Royal Society of Medicine* 109: 154-63.
-
- ⁷⁷ Allen C. (2014) A resource for those preparing for and responding to natural disasters, humanitarian crises, and major health care emergencies. *Journal of Evidence-Based Medicine* 7: 234-7; and Gurevitch J, et al. (2018) Meta-analysis and the science of research synthesis. *Nature* 555: 175-82.
-
- ⁷⁸ Oxman AD. (1994) Checklists for review articles. *BMJ* 309: 648-51.
-
- ⁷⁹ Pour une critique de l'application des synthèses de recherche aux politiques, voir Pawson R. (2001) Evidence Based Policy: In Search of a Method. Working Paper 3. London, UK: ESRC UK Centre for Evidence Based Policy and Practice.
-
- ⁸⁰ Donnelly CA, et al. (2018) Four principles for synthesizing evidence. *Nature* 558: 361-4.
-
- ⁸¹ Lewis S, Clarke M. (2001) Forest plots: trying to see the wood and the trees. *BMJ* 322:1479-80.
-
- ⁸² Bangpan M, et al. (2017) The impact of mental health and psychosocial support programmes for populations affected by humanitarian emergencies. Oxford, UK: Oxfam.
-
- ⁸³ Ganann R, et al. (2010) Expediting systematic reviews: methods and implications of rapid reviews. *Implementation science* 5:56; et Tricco AC, et al. (2015) A scoping review of rapid review methods. *BMC Medicine* 13: 224.
-
- ⁸⁴ Begley CG, et al. (2012) Drug development: Raise standards for preclinical cancer research. *Nature* 483: 531-3.
-
- ⁸⁵ Un examen de la recherche comportant une ou plusieurs méta-analyses cumulatives a révélé de nombreux exemples montrant que des résultats stables (bénéfiques, nocifs et neutres) auraient été obtenus si une méta-analyse de la recherche existante avait été effectuée avant le début d'un nouvel essai randomisé, ce qui aurait permis de mettre en place des interventions efficaces plus tôt : Clarke M, et al. (2014) Accumulating research: a systematic account of how cumulative meta-analyses would have provided knowledge, improved health, reduced harm and saved resources. *PLoS ONE* 9(7): e102670.
-
- ⁸⁶ Chalmers I, et al. (2014) How to increase value and reduce waste when research priorities are set. *Lancet* 383: 156-65.
-
- ⁸⁷ Shea BJ, et al. (2007) Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology* 7:10; et Shea BJ, et al. (2017) AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of health care interventions, or both. *BMJ* 358:j4008.
-
- ⁸⁸ Smith V, et al. (2011) Methodology in conducting a systematic review of systematic reviews of health care interventions. *BMC Medical Research Methodology* 11:15.
-
- ⁸⁹ Voir, par exemple : Brennan RJ, et al. (2005) Rapid health assessment in Aceh Jaya district, Indonesia, following the December 26 tsunami. *Emergency Medicine Australasia* 17: 341e350 ; et Beebe J. (2014) Rapid Qualitative Inquiry: a Field Guide to Team-based Assessment, second edition. Lanham, Maryland, USA: Rowman & Littlefield.
-

Karl Blanchet, Directeur du centre de santé en situation de crise humanitaire, École d'hygiène et de médecine tropicale de Londres
Karl.Blanchet@lshtm.ac.uk | crises.lshtm.ac.uk

Mike Clarke, Directeur de recherche de l'aide en matière de preuves, Centre de santé publique, Queen's University Belfast
mclarke@qub.ac.uk | www.qub.ac.uk



La traduction de ces informations a été rendue possible grâce aux généreux efforts des traducteurs travaillant avec Translators without Borders (TWB), une organisation bénévole œuvrant pour un monde où le savoir ne connaît pas de barrières de langue.