

4.5 Distinguer les données probantes de grande qualité des données probantes de faible qualité

Toutes les données probantes ne sont pas de grande qualité et fiables pour prendre des décisions. Des outils existent pour de nombreuses formes de données probantes (mais pas toutes) afin d'aider à juger si les données probantes sont fiables (que ce soit pour une seule étude ou un ensemble de données probantes). Comme nous le décrivons ici, ces outils utilisent des scores ou des notes pour aider les utilisateurs à comprendre le degré de confiance qu'ils peuvent avoir dans les données probantes. De nombreux journaux exigent désormais des auteurs qu'ils respectent des normes pour rapporter leurs résultats, telles que CONSORT pour les essais contrôlés randomisés et PRISMA pour les synthèses de données probantes. Toutefois, la plupart des journaux ne demandent pas aux examinateurs d'utiliser des outils spécifiques pour évaluer la qualité des études ou la force des recommandations. Par conséquent, la publication dans un journal à comité de lecture n'est pas un bon indicateur de la qualité.

Enjeu	Réponse
<p>La qualité (ou la fiabilité) des études (et des lignes directrices) varie</p>	<ul style="list-style-type: none"> Des outils d'évaluation de la qualité (ou d'évaluation critique) ont été mis au point pour des modèles d'étude spécifiques (par exemple, l'essai contrôlé randomisé), pour de grandes catégories de modèles d'études (par exemple, l'étude observationnelle, la recherche qualitative et la synthèse des données probantes) et pour les lignes directrices - voir l'annexe à la fin de ce chapitre (section 4.16) pour des exemples (RoB2, ROBINS-I, liste de contrôle JBI, AMSTAR et AGREE II). Les outils peuvent produire un jugement sommaire (par exemple, un faible risque de biais en utilisant RoB2 ou ROBINS-I), un score que certains regroupent en intervalles (par exemple, une qualité élevée en utilisant AMSTAR), un ensemble de scores (par exemple, six domaines en utilisant AGREE II), ou un ensemble de considérations qui peuvent informer un jugement sommaire (par exemple, la liste de contrôle JBI).
<p>Le corpus de données probantes varie dans son degré de certitude (ou la confiance que vous pouvez lui accorder)</p>	<ul style="list-style-type: none"> Des outils d'évaluation de la certitude ont été développés pour un ensemble de données probantes abordant la même question (par exemple, l'effet d'une intervention sur un résultat spécifique ou la signification que les citoyens attachent à un phénomène particulier) - voir section 4.16 pour deux exemples (GRADE et GRADE CERQual) Les outils peuvent produire un jugement sommaire sur la confiance que l'effet réel est similaire à l'effet estimé (par exemple, une certitude élevée avec GRADE) ou que le phénomène d'intérêt est bien représenté par les résultats d'une étude qualitative (avec GRADE CERQual) Un jugement sommaire sur la certitude d'une estimation de l'effet est plus utile qu'un test de signification statistique démontrant qu'une intervention « fonctionne » ou « ne fonctionne pas » (ce qui arrivera par hasard une fois sur 20 si la signification statistique est fixée au niveau 0,05).
<p>Les recommandations varient dans leur force</p>	<ul style="list-style-type: none"> Des outils d'évaluation de la force ont été développés pour les recommandations des lignes directrices (par exemple, GRADE, en plus du classement de la certitude d'un ensemble de données probantes, comme décrit ci-dessus) - voir section 4.16 pour un exemple. Les outils peuvent produire un jugement sommaire sur la question de savoir si la plupart des décideurs choisiraient de procéder à une intervention (par exemple, forte avec GRADE) ou si la plupart d'entre eux auraient besoin de soupeser soigneusement le pour et le contre d'une intervention
<p>Certaines sources de données probantes (ou les approches utilisées pour les générer) peuvent être difficiles à évaluer</p>	<ul style="list-style-type: none"> Il n'existe pas d'outils largement acceptés pour évaluer le degré de confiance que l'on peut accorder à ces données : <ul style="list-style-type: none"> Un expert, bien que des exemples comme The Good Judgement Project existent pour faire des prévisions (nous revenons sur l'opinion d'experts plus loin dans ce chapitre et, dans le cas de l'opinion d'experts sur les paramètres du modèle, dans la section 4.16). Les modèles utilisés pour générer certaines formes de données probantes (que nous abordons dans la section 4.7 en parlant des modèles de changement climatique et dans la section 4.16). Un algorithme d'intelligence artificielle utilisé pour générer certains types de données probantes, bien que des exemples comme TRIPOD commencent à émerger (3)

Distinguer les données probantes de grande qualité des données probantes de faible qualité est particulièrement difficile lorsque les données probantes sont intégrées dans des tableaux de bord, des modèles et d'autres formats, et lorsque des conflits d'intérêts sont en jeu. Nous revenons sur ce dernier point dans les [sections 4.12, 4.14](#) et [4.16](#). Bien que cela ne fasse pas l'objet de ce rapport, distinguer les « données brutes » de haute qualité de celles de faible qualité peut également être un défi, et des organisations comme l'UNICEF ont développé des cadres de qualité des données pour y parvenir (bit.ly/3DQQRrv).

Certains « guichets uniques », tels que Social Systems Evidence et le répertoire du réseau COVID-19 Evidence Network to support Decision-making (COVID-END) (décrit à la [section 4.6](#)), utilisent certains de ces outils afin que les décideurs et ceux qui les soutiennent puissent se concentrer sur des synthèses de données probantes de grande qualité ou comprendre qu'ils utilisent les meilleures synthèses de données probantes disponibles (si elles ne sont pas de grande qualité).

La pandémie de COVID-19 a exigé des décideurs qu'ils prennent des décisions difficiles dans des délais très courts, au départ avec peu de données probantes, souvent indirectes, puis, au fil du temps, avec des études, des corpus de données probantes et des recommandations élaborées selon un processus solide. Afin d'appuyer la prise de décision concernant la COVID-19 sur la base d'ensembles de données probantes (plutôt que d'études individuelles), le réseau COVID-END a fait figurer dans son inventaire des « meilleures » synthèses de données probantes celles qui étaient à jour (en fonction de la date de recherche des données probantes), de grande qualité (selon l'outil AMSTAR) et qui fournissaient une évaluation de la certitude des données probantes (selon l'outil GRADE).

Les données probantes ne sont pas toutes de grande qualité, et toutes les données probantes mondiales ne sont pas toutes applicables à un contexte donné. Par exemple, une synthèse des données probantes contenant des études menées uniquement dans des pays à revenu élevé peut avoir une applicabilité limitée à certains pays à faible revenu. Il peut y avoir des différences importantes dans les conditions de base, dans les réalités et les contraintes sur le terrain et dans les caractéristiques structurelles du système local (par exemple, le système de santé national ou le système d'éducation provincial/étatique). L'outil SUPPORT peut également aider les gens à réfléchir à l'applicabilité locale des résultats d'une synthèse des données probantes et à voir comment on peut parfois tirer des conclusions même si les résultats ne sont pas applicables.(4)

Le raisonnement bayésien a suscité une attention croissante en tant que moyen de redessiner délibérément nos « cartes mentales » concernant les défis et les moyens de les relever, non pas en remplaçant tout ce que nous pensions savoir par de nouvelles informations, mais en modifiant notre compréhension à un degré approprié. Ce degré dépend de la confiance que vous aviez dans vos connaissances préexistantes (la probabilité « antérieure » qu'une chose soit vraie) et de la confiance que vous accordez aux nouvelles connaissances. On peut accorder une plus grande confiance aux nouvelles connaissances si elles proviennent d'une synthèse des données probantes de grande qualité qui inclut des études menées dans des contextes similaires au vôtre.



Intermédiaire et producteur de données probantes, Gillian Leng

Cadre expérimentée à la tête d'une agence d'évaluation des technologies et de production de lignes directrices qui soutient la prise de décision en matière de santé et de services sociaux par les gouvernements, les prestataires de services et les patients

Le Royaume-Uni a mené des travaux pendant de nombreuses années pour encourager la synthèse et l'utilisation de données probantes - du premier essai contrôlé randomisé pour prévenir le scorbut chez les marins, aux plus récents centres innovants What Works pour promouvoir l'utilisation de données probantes dans un éventail de domaines politiques. Dans le cadre de ce mouvement fondé sur des données probantes, au cours des 20 dernières années, le National Institute for Health and Care Excellence (NICE) a transformé l'utilisation des données probantes dans la pratique des soins de santé, ainsi que dans des initiatives plus larges de santé publique et de services sociaux.

La pandémie de COVID-19 a considérablement renforcé le besoin de données probantes de haute qualité pour éclairer les politiques et les pratiques, et a également mis en évidence les conséquences négatives des médias sociaux et de la désinformation qui y est associée. Dans ce contexte, les travaux de la Commission mondiale sur les données probantes pour relever les défis sociétaux sont extrêmement importants et doivent être considérés comme une lecture essentielle pour tous les décideurs politiques du monde entier.