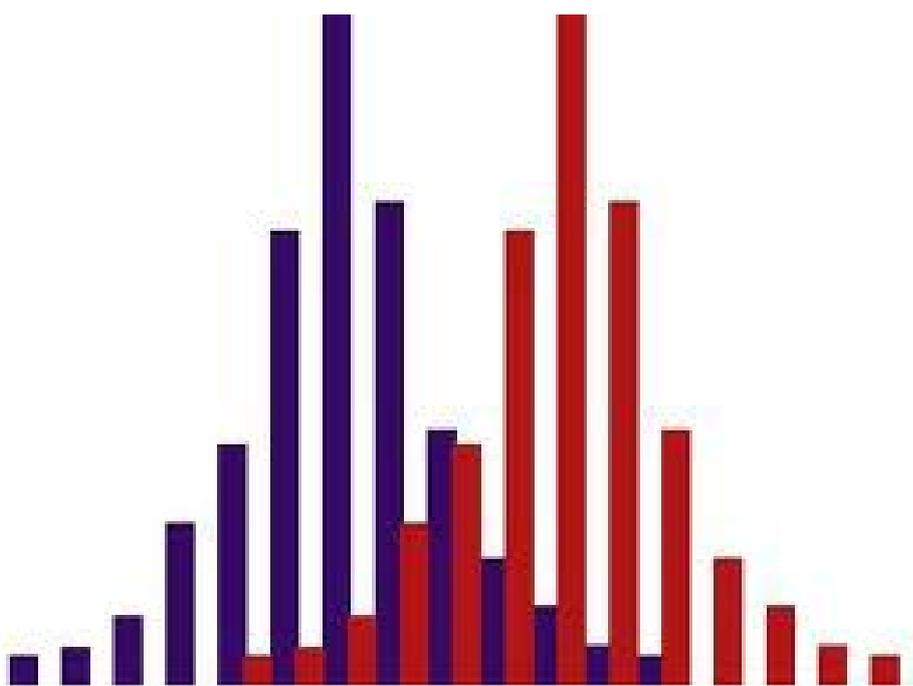


# LA PRATIQUE DE LA MÉDECINE FONDÉE SUR LES PREUVES EN SANTÉ MENTALE DE L'ENFANT ET DE L'ADOLESCENT

**John Hamilton & Füsun Çuhadaroğlu-Çetin**

Edition en français  
Traduction : Cristina Constantin  
Sous la direction de : Priscille Gérardin  
Avec le soutien de la SFPEADA



John Hamilton MD, MSc

Senior Physician, The  
Permanente Medical Group,  
Inc., Sacramento, California,  
USA

Conflict of interest: none  
reported.

Füsun Çuhadaroğlu-Çetin  
MD

Professor of Child and  
Adolescent Psychiatry,  
Hacettepe University Medical  
School, Ankara, Turkey

Conflict of interest: none  
reported.

Acknowledgments. Professor  
Eric Youngstrom, University of  
North Carolina, Professor Peter  
Szatmari, McMaster University,  
and Dr Eric Daleiden,  
PracticeWise LLC, were very  
helpful reviewers in correcting

Cette publication est à destination des professionnels de la santé mentale, qu'ils soient en formation ou en exercice. Elle n'est pas destinée au grand public. Les opinions exprimées sont celles des auteurs et ne représentent pas nécessairement le point de vue de l'Editeur ou de la IACAPAP. Cette publication tente de décrire les meilleurs traitements et pratiques basés sur des preuves scientifiques disponibles au moment de sa rédaction, traitements et pratiques qui pourraient donc évoluer en fonction des recherches à venir. Les lecteurs doivent mettre en perspectives ces connaissances avec les recommandations et les lois en vigueur dans leur pays. Certains traitements pourraient ne pas être disponibles dans certains pays et les lecteurs devraient consulter les informations spécifiques des médicaments car tous les dosages et les effets indésirables ne sont pas mentionnés. Les organisations, les publications et les sites web sont cités ou mis en lien afin d'illustrer les résultats et de pouvoir rechercher davantage d'informations. Cela ne veut pas dire que les auteurs, l'Editeur ou la IACAPAP endossent leurs contenus ou leurs recommandations, lesquelles pourraient être évaluées de façon critique par le lecteur. De même, les sites web peuvent changer ou cesser d'exister.

©IACAPAP 2017. Ceci est une publication en accès libre sous la [Creative Commons Attribution Non-commercial License](#). L'utilisation, la distribution et la reproduction sur tout type de support sont permises sans permission préalable du moment que le travail original est correctement cité et que l'utilisation n'est pas commerciale. Envoyez vos commentaires sur ce livre ou ce chapitre à [jmreyATbigpond.net.au](mailto:jmreyATbigpond.net.au)

Citation suggérée : Hamilton J, Çuhadaroğlu-Çetin F. Evidence-based practice in child and adolescent mental health. In Rey JM (ed), *IACAPAP e-Textbook of Child and Adolescent Mental Health*. (édition en français; Cohen D, ed.) Geneva: International Association for Child and Adolescent Psychiatry and Allied Professions 2017.

Ce chapitre montre comment utiliser les données des études ayant moins de biais et moins d'erreurs aléatoires afin de prendre de bonnes décisions cliniques pour aider les jeunes nécessitant un traitement pour des problèmes de santé mentale. Plus précisément, ce chapitre montre comment :

- Penser les biais et les erreurs aléatoires comme indésirables
- Comprendre les concepts statistiques les plus pertinents pour l'alignement de votre travail clinique avec des preuves solides
- Faire des recherches sur les bases de données électroniques disponibles gratuitement à toute personne ayant une connexion internet
- Utiliser des bonnes techniques d'entretien clinique pour montrer que les jeunes remplissent ou ne remplissent pas les critères diagnostiques d'un syndrome bien défini
- Trouver les traitements les plus susceptibles d'être efficaces pour ces syndromes.

En bref, ce chapitre montre aux cliniciens comment aligner leur pratique avec *ce qui marche* : des choix qui vont entraîner une amélioration clinique plus substantielle et plus rapide chez ces jeunes par rapport à des choix thérapeutiques alternatifs. Mais avant de connaître la vérité sur ce qui fonctionne, il est d'abord nécessaire de faire la guerre au deux grands ennemis de la vérité : *le biais et l'erreur*.

## LE BIAIS ET L'ERREUR ALEATOIRE

### Le problème du biais

Le biais interfère avec la validité – la précision avec laquelle une étude mesure ce qu'elle a l'intention de mesurer – en créant des déviations systématiques (non aléatoires) par rapport à la vérité sous-jacente (Guyatt & Rennie, 2002). Le biais est différent de *l'erreur aléatoire*. L'erreur aléatoire se trouve diminuée par l'augmentation de la taille de l'échantillon alors que le biais peut rester inchangé peu importe la taille de l'échantillon. Considérons les cibles de (a) à (d) de la figure A.6.1. où le centre de la cible marque l'effet réel. Dans l'étude (a), il y a eu peu

and adding to earlier drafts. The final responsibility for the manuscript, including errors, belongs to the authors. You may contact the corresponding author at john.hamilton@kp.org

### Erreur

Tous les essais ou décisions cliniques sont soumis soit à une erreur aléatoire, soit à une erreur systématique, soit aux deux. Les erreurs aléatoires sont dues à des changements inconnus et imprévisibles des instruments de mesure ou des conditions environnementales. Les erreurs systématiques (ou biais), par contre, sont des inexactitudes qui sont constamment dans la même direction. Ils peuvent se produire parce qu'il y a un dysfonctionnement de l'instrument de mesure (par exemple, une échelle de notation, des critères de diagnostic) ou parce que l'instrument est utilisé de façon erronée par l'expérimentateur (par exemple, un clinicien qui surestime –ou sous-estime– systématiquement, le risque suicidaire).

Figure A.6.1 Exemples de biais et d'erreur aléatoire dans des échantillons (Kabai, 2011a).

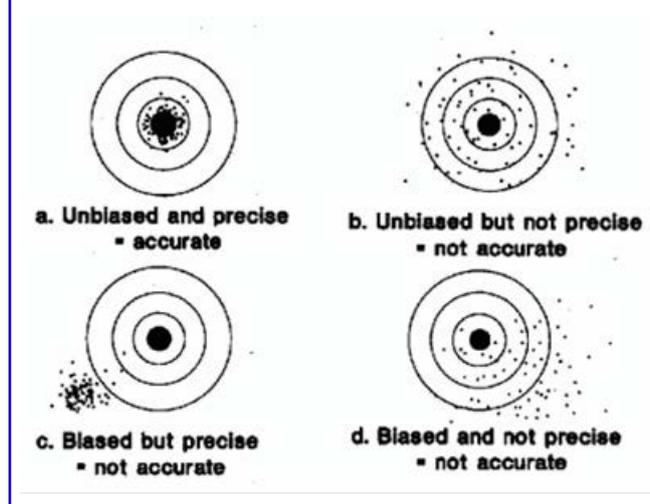
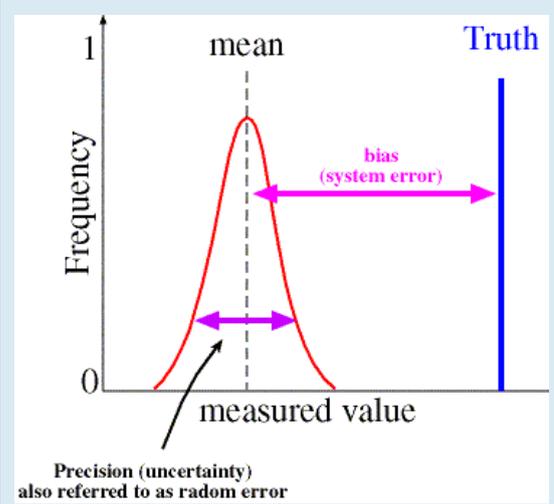


Figure A.6.2 Représentation schématique de biais et d'erreur aléatoire (Kabai, 2011b).



**Tableau A.6.1 Sources courantes de biais et les stratégies pour les réduire.**

| Source de biais   | Stratégies pour réduire la source du biais   |
|---|--|
| <ul style="list-style-type: none"> <li>Le pronostic des patients traités est différent du pronostic des patients témoins</li> </ul> | <ul style="list-style-type: none"> <li>Allocation randomisée dans le groupe traitement ou témoin</li> </ul>  |
| <ul style="list-style-type: none"> <li>Les effets placebo du traitement</li> </ul>  | <ul style="list-style-type: none"> <li>Les patients sont maintenus en aveugle quant à leur groupe traitement ou témoin</li> </ul>  |
| <ul style="list-style-type: none"> <li>La famille rajoute d'autres traitements</li> </ul>   | <ul style="list-style-type: none"> <li>La famille est également maintenue en aveugle</li> </ul>  |
| <ul style="list-style-type: none"> <li>Les évaluateurs du résultat sont biaisés</li> </ul>  | <ul style="list-style-type: none"> <li>Les évaluateurs sont maintenus en aveugle par rapport à l'appartenance au groupe témoin ou traitement du patient qu'ils évaluent</li> </ul> |
| <ul style="list-style-type: none"> <li>Pertes de vue</li> </ul>   | <ul style="list-style-type: none"> <li>Peu ou pas de pertes de vue</li> </ul>  |

### Biais

Le biais est un risque toujours présent dans la recherche, même dans des essais randomisés contrôlés (ECR) bien conçus. Par exemple, les résultats positifs sont rapportés beaucoup plus fréquemment dans les essais financés par l'industrie pharmaceutique (78%), comparés à ceux qui n'ont pas de parrainage de l'industrie (48%) ou ceux financés par un concurrent (28%) (Kelly et al, 2006)..

**Tableau A.6.2 Sources d'erreurs aléatoires et stratégies pour les réduire.**

| Sources d'erreur aléatoire   | Stratégie pour la réduction de celle-ci  |
|--|--|
| <ul style="list-style-type: none"> <li>Instruments de mesure imprécis</li> </ul> | <ul style="list-style-type: none"> <li>Améliorer la précision des instruments</li> </ul> |
| <ul style="list-style-type: none"> <li>Petite taille de l'échantillon</li> </ul> | <ul style="list-style-type: none"> <li>Augmenter la taille de l'échantillon</li> </ul>   |

ou pas de déviation systématique par rapport à l'effet réel (peu ou pas de biais), ainsi que peu de dispersion (erreur aléatoire). Dans (b) il y avait aussi peu de biais mais beaucoup d'erreurs aléatoires alors que en (c) il y a beaucoup de biais mais peu d'erreurs aléatoires. En (d) il y avait beaucoup de biais et beaucoup d'erreurs aléatoires.

### Trouver des preuves avec peu de biais et peu d'erreurs aléatoires

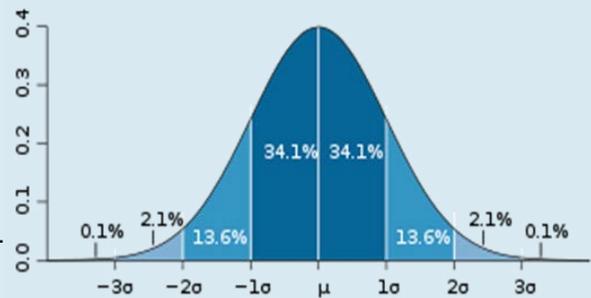
La preuve la plus utile pour guider les décisions cliniques est la preuve avec un *biais minimal et une erreur aléatoire minimale*. La figure A.6.2 est une représentation schématique d'erreur aléatoire et de biais. Le tableau A.6.1 montre les sources courantes de biais et les stratégies pour les réduire alors que le Tableau A.6.2 montre les sources courantes d'erreurs aléatoires et les stratégies pour la réduction de celles-ci.

## CONCEPTS UTILES POUR LA PRATIQUE FONDÉE SUR LES PREUVES

### La moyenne

Quelques concepts statistiques de base sont tout ce qui est nécessaire pour comprendre de nombreuses idées de base utilisées dans les approches fondées sur des preuves. Tout d'abord, il faut rappeler la normale de base ou courbe gaussienne qui décrit la distribution de nombreux aspects de la nature et qui peut être dérivée

**Figure A.6.3** Tracé d'une distribution normale (gaussienne) ayant une valeur moyenne représentée par la lettre grecque mu ( $\mu$ ) et un écart type, delta ( $\sigma$ ), représenté par chaque bande colorée. Notons qu'un peu plus d'un tiers (34,1%) de la population de cette distribution se trouve à un écart type au-dessus de la moyenne; Un peu plus d'un tiers (34,1%) se trouve également à un écart-type en dessous de la moyenne. Seulement 2% environ de la population se trouve à deux ou plus de deux écarts-types au-dessus de la moyenne, et seulement environ 2% se trouve à deux ou plus de deux écarts types en dessous de la moyenne (Wikipedia).



mathématiquement en utilisant la théorie des probabilités. La valeur la plus commune est la même que la valeur moyenne, représentée ici par la lettre grecque mu ( $\mu$ ).

### L'écart-type (ET)

L'écart-type est une mesure de la variabilité: elle reflète la variation ou la «dispersion» par rapport à la valeur moyenne. Un écart-type faible par rapport à la moyenne indique une distribution où les points de données sont étroitement groupés ensemble. Un écart-type large par rapport à la moyenne indique une distribution dans laquelle les points de données sont répartis sur une zone large autour de la moyenne. L'écart-type est parfois décrit comme *la quantité typique qui différencie les cas les uns des autres*. Deux cas sur trois devraient tomber à plus ou moins un écart-type de la moyenne.

### Taille d'effet

La taille d'effet dans une étude de recherche mesure l'importance de l'effet d'un traitement par rapport aux différences entre les cas en général. C'est la différence moyenne normalisée entre les deux groupes dans une étude. Techniquement, ceci est mesurée comme le rapport de la différence due au traitement par rapport à ce que l'on appelle l'écart-type cumulé. Cumulé fait référence à la combinaison des chiffres du groupe témoin et du groupe expérimental pour créer un écart-type cumulé ou regroupé.

$$\text{Taille d'effet} = \frac{[\text{Moyenne groupe expérimental}] - [\text{Moyenne groupe témoin}]}{[\text{Écart type cumulé}]}$$

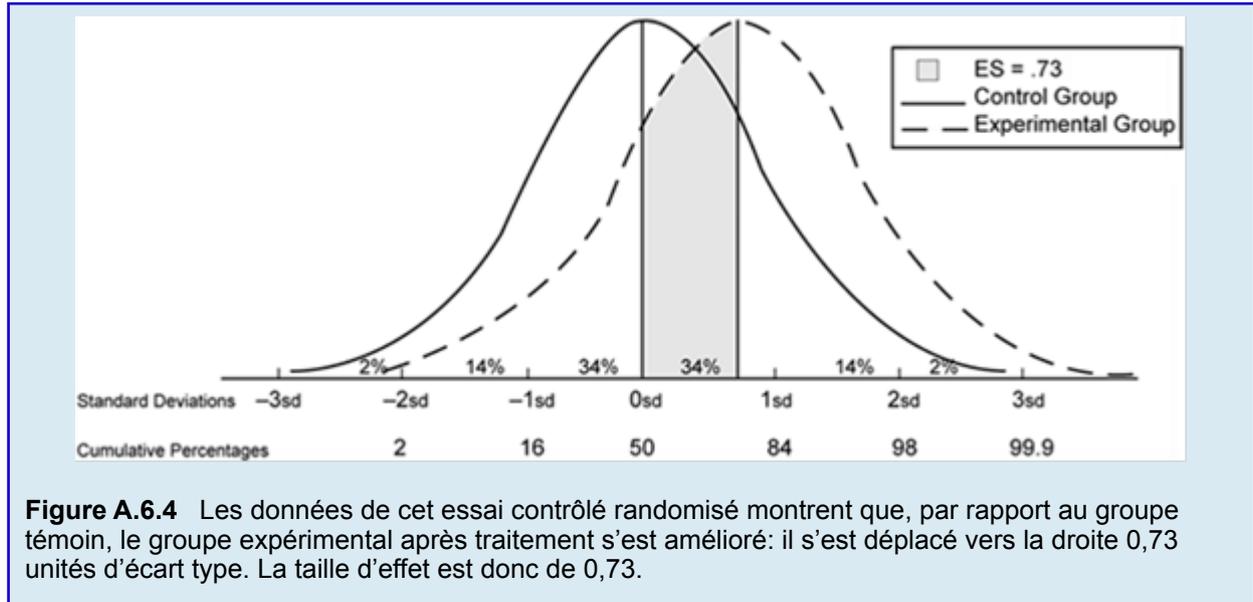
La figure A.6.4 montre une représentation typique des données d'un traitement expérimental en psychiatrie infantile avec une taille d'effet de 0,73.

### Exemple de taille d'effet

Une étude (McCracken et al, 2002) a tenté d'examiner si un antipsychotique atypique (rispéridone) était plus efficace que le placebo pour réduire l'agressivité, les crises de colère ou les conduites auto-agressives chez les enfants atteints d'autisme âgés de 5 à 17 ans. Les scores après 8 semaines sur une mesure de l'irritabilité étaient de 11,3 (ET 7,4) pour le groupe traité et 21,9 (ET 9,5) pour le groupe témoin. Quelle est la taille d'effet?

Taille d'effet =  $(11,3 - 21,9) / 9,5$ , soit 1,1. Il s'agit d'une grande taille d'effet très significative. Les effets de taille supérieurs à 0,5 sont souvent considérés cliniquement significatifs, mais de plus petits effets peuvent encore avoir une pertinence clinique si le traitement peut facilement être appliqué à de grandes populations (par exemple, l'utilisation d'aspirine pour prévenir les crises cardiaques, où la taille d'effet est petite, mais sur de grandes populations, les conséquences sur la santé publique sont importantes).

Comment calculer l'écart-type cumulé peut être trouvé sur des sites statistiques (e.g., <http://www.leeds.ac.uk/educol/documents/00002182.htm>). La taille d'effet met l'accent sur l'aspect le plus important d'une intervention – l'ampleur de l'effet - plutôt que sa signification statistique, qui est déterminée à la fois par la taille de l'effet et la taille de l'échantillon.



Notez que les deux courbes se superposent considérablement.

### Réduction du risque absolu (RRA)

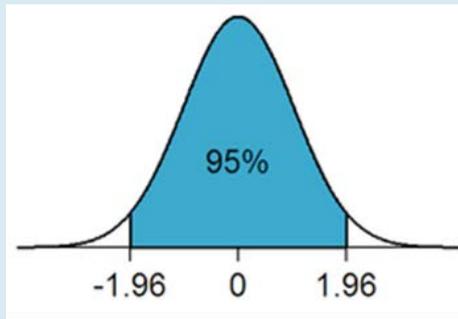
La réduction du risque absolu est une mesure utilisée pour comparer deux alternatives différentes: *de combien un traitement réduit-il le risque d'un «mauvais» résultat prédéfini (comme le fait qu'un jeune déprimé reste déprimé) par rapport à un autre traitement ou à un placebo?* Par exemple, dans l'étude sur le traitement des adolescents avec dépression (TADS), à 12 semaines, 65,2% des adolescents recevant le placebo n'ont pas été classés comme améliorés ou très améliorés comparativement à seulement 39,4% des adolescents recevant la fluoxétine. Par conséquent, la RRA de ne pas s'améliorer pour les jeunes souffrant d'un épisode dépressif majeur qui ont pris de la fluoxétine par rapport au placebo était de  $65,2 - 39,4 = 25,8\%$ . En d'autres termes, la fluoxétine réduit considérablement le risque de ne pas s'améliorer: 25,8%, soit environ 1 sur 4.

La réduction du risque absolu est utile pour deux raisons. Tout d'abord, comme discuté ci-dessous, la RRA peut être utilisée pour calculer le *nombre nécessaire à traiter* (NNT) ou nombre de sujets à traiter. La formule est simplement  $NNT = 1 / RRA$  avec la réponse arrondie au nombre entier suivant. Deuxièmement, la RRA ne surévalue pas l'effet du traitement comme le fait la réduction du risque relatif (RRR) qui est simplement calculée comme le risque dans le groupe de traitement divisé par le risque dans le groupe témoin. Dans le TADS, le RRR lors de la comparaison du résultat chez les jeunes recevant de la fluoxétine par rapport aux jeunes recevant un placebo est le risque que les symptômes dépressifs ne soient pas améliorés de façon significative dans le groupe traité recevant la fluoxétine divisé par le risque que les jeunes ne soient pas significativement améliorés dans le groupe placebo  $39,4 / 65,2 = 0,60$ , généralement exprimé comme 60%. Notons que ce chiffre est très différent de la réduction du risque absolu telle que calculée dans le paragraphe ci-dessus (25,8%). En général, la RRA est une statistique plus utile que le RRR.

### Intervalle de confiance (IC)

Plutôt que de commencer par une définition, un exemple pourrait

**Figure A.6.5** Un facteur important associé au « 1.96 » dans la formule pour le calcul des intervalles de confiance à 95% : 95% de la superficie d'une distribution normale se trouve dans les 1.96 unités d'écart-type par rapport à la moyenne.



être plus utile pour se familiariser avec ce concept important. Supposons que la ministre de la Santé vous dise qu'elle est préoccupée par les adolescents déprimés et qu'elle vous demande de déterminer le niveau moyen de symptômes dépressifs chez les enfants de 13 ans dans la grande ville où vous vivez. Vous acceptez le défi et vous commencez le projet en vous arrangeant pour que tous les 13 ans de la ville complètent un questionnaire d'évaluation des symptômes dépressifs. Le score moyen est de 39,8 sur les 49 questionnaires remplis. Vous calculez l'écart-type et il est de 9,0. La ministre de la Santé veut que vous estimiez, à partir de vos données des 49 questionnaires retournés, à quel point votre valeur moyenne est égale à la valeur *réelle*, également appelée valeur de la *population*. C'est-à-dire, s'il était possible d'évaluer tous les enfants de 13 ans dans la ville avec cette échelle de dépression, ce serait vraisemblablement très proche de la *vraie* valeur. Vous pouvez donc voir que la valeur réelle (ou la valeur de la population) est un concept utile, mais en réalité, il n'est souvent pas pratique de tester tous les enfants. Par hasard, vous avez peut-être fini avec une proportion plus élevée d'enfants déprimés ou d'enfants heureux dans votre échantillon par rapport à ce qui est présent dans la ville en général, même si vous avez réussi à éviter d'être biaisé dans votre échantillonnage (comme le fait de choisir les élèves d'une même école, ou seulement les élèves d'une certaine classe).

Vous dites à la ministre que le score moyen dans votre échantillon est de 39,8, à titre d'estimation approximative. Elle, cependant, veut en savoir plus, et demande: «Quel est votre degré de confiance concernant cette approximation? Quelle est la valeur réelle la plus haute et la plus basse?» Vous répondez: «Je peux vous donner une fourchette ou un intervalle, mais cela dépend de la fréquence à laquelle vous voulez que je sois juste. Voulez-vous que je sois juste autour de 80% du temps quand je vous donne ces estimations ou 95% du temps? Ou même 99%? Plus vous insistez que la réponse que je vous donne contienne la vraie réponse pour la population, plus l'intervalle sera large ». La ministre dit: « OK, je veux que vos réponses sur ce genre d'estimations soient exactes 95% du temps ». La ministre vient de demander l'intervalle de confiance à 95% de votre moyenne calculée.

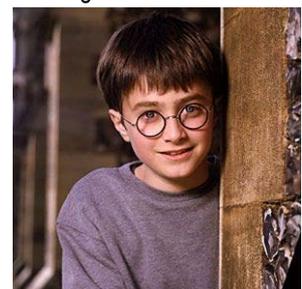
Un statisticien peut maintenant aider. Il nous dit qu'une bonne estimation pour calculer les intervalles de confiance dans les distributions normales est la suivante, où N est le nombre de sujets générant les données (49 dans ce cas) et SQRT signifie la racine carrée:

$$\text{Intervalle de confiance, limite sup} = \text{moyenne calculée} + 1,96 \times ET / \text{SQRT } N$$

$$\text{Intervalle de confiance, limite inférieure} = \text{moyenne calculée} - 1,96 \times ET / \text{SQRT } N$$

Par conséquent, les intervalles de confiance sont:

Les intervalles de confiance peuvent vous transformer en magicien de données.



$$IC \text{ (limite sup)} = 39,8 + 1,96 \times 9,0 / \text{SQRT } 49 = 39,8 + 1,96 \times 9,0 / 7 = 39,8 + 2,52 = 42,32$$

$$CI \text{ (limite inf)} = 39,8 - 1,96 \times 9,0 / \text{SQRT } 49 = 39,8 - 1,96 \times 9,0 / 7 = 39,8 - 2,52 = 38,28$$

Maintenant, vous avez une réponse pour la ministre.

En résumé, les IC sont exprimés avec trois nombres, le premier étant la moyenne, notre meilleure estimation de la vraie valeur. Les deux autres nombres sont bien sûr les nombres indiquant la limite inférieure et la limite supérieure de l'intervalle. Plus nous voulons être précis, plus l'intervalle de confiance sera large: les intervalles de confiance larges permettent que le nombre correct soit là quelque part. *Les intervalles de confiance sont importants parce que l'analyse statistique des études est basée sur l'idée centrale que, malgré les observations faites sur un échantillon limité de sujets, la vérité inférée à tirer de l'étude s'appliquera à la population de tous ces sujets.* Le but principal d'un intervalle de confiance est d'indiquer la précision ou l'imprécision de l'échantillon de l'étude comme estimation de la vraie valeur de la population. Les intervalles de confiance sont donc utiles et même nécessaires chaque fois qu'une inférence est faite à partir des résultats d'une étude sur le reste de la population. Dans le premier exemple, il y avait seulement 49 enfants remplissant le questionnaire sur la dépression, mais la ministre espère généraliser les résultats à tous les enfants de 13 ans dans la ville. Notez également que même si la ministre et vous choisissez votre échantillon judicieusement - par exemple, dans des écoles très différentes dans différentes parties de la ville pour éviter les biais - le problème de la variation aléatoire du niveau des symptômes de dépression reste. Par hasard, vous avez peut-être rencontré des enfants dont le taux de dépression était nettement plus élevé ou plus faible que dans toute la population des enfants de cet âge, même si aucun biais systématique significatif n'a été relevé dans votre méthode d'échantillonnage.

### Nombre de sujets à traiter (NST)

Le nombre de sujets nécessaires à traiter, souvent abrégé NST ou NNT, peut être utilisé pour résumer en un seul nombre l'efficacité d'un traitement spécifique comparé au placebo. Le NST est défini comme *le nombre de personnes que nous devons traiter afin d'éviter un mauvais résultat clinique supplémentaire.* Un mauvais résultat est défini par des critères spécifiques. Par conséquent, les NST faibles (comme 3 ou 4) indiquent un traitement efficace, car le traitement permet fréquemment de «convertir» un patient à un résultat positif, patient qui ne se serait pas amélioré sans le traitement. Les résultats cliniques les plus courants sont



Quelle devrait être la difficulté que les chercheurs exigent pour qu'un jeune soit considéré comme «amélioré» lorsqu'ils calculent le NST?

Des haies plus hautes soulèveront le NST parce que moins de jeunes réussiront à les franchir.

Mann et al (2004) ont mené une méta-analyse pour estimer l'efficacité de l'acamprosate pour arriver à l'abstinence chez les alcooliques. Ils ont rapporté que, parmi les études, 36,1% des participants prenant l'acamprosate avaient atteint l'abstinence à 6 mois comparativement à 23,4% de ceux qui avaient reçu un placebo. Combien est le NST dans ce cas?

Le NST =  $1 / (0,234 - 0,361) = 7,9$ . C'est-à-dire que huit patients (il est habituel d'arrondir au nombre entier suivant) devront être traités avec l'acamprosate pour qu'un patient supplémentaire s'abstienne d'alcool à 6 mois comparés à ceux qui ne reçoivent pas le traitement.

l'amélioration de l'humeur ou l'amélioration du fonctionnement, souvent définies par un nombre seuil sur une échelle.

Notez cependant que le NST dépend beaucoup de la *difficulté* qu'il faut éliminer afin d'être considéré comme «mieux» ou «guéri». Le niveau élevé ou faible de cette difficulté *aura un impact important* sur le NST. Toujours noter, par conséquent, lors de la lecture d'un article contenant un NST, ce que les auteurs ont choisi comme difficulté.

Les NST élevés (comme 25, 30 ou même plus) suggèrent un traitement qui n'est pas très efficace ou un résultat qui est difficile à atteindre. Par exemple, si nous exigeons qu'un résultat positif chez un adolescent déprimé soit un résultat sans aucun symptôme dépressif, il sera difficile d'obtenir ce résultat et le NST correspondant sera plus élevé que si nous avons choisi un résultat plus modeste.

### Nombre nécessaire pour nuire (NSN)

Le nombre nécessaire pour nuire ou nombre de sujets pour nuire est abrégé NNN ou NSN et résume la fréquence avec laquelle un effet secondaire spécifique se produit avec un traitement spécifique. Le NSN est défini comme le nombre de personnes que nous devons traiter afin qu'une seule personne soit lésée par le traitement qui ne l'aurait pas été si elle avait reçu seulement une intervention par placebo. Par exemple, avec des médicaments, un effet indésirable ou un effet secondaire se produit qui n'aurait pas eu lieu si le patient avait reçu un placebo. Par conséquent, les NSN élevés indiquent un traitement avec moins de risques: de nombreuses personnes doivent recevoir le traitement avant qu'une seule personne ne soit lésée par le traitement. Les décisions de recommander un traitement spécifique nécessiteront donc de prendre en compte les bénéfices (NST) et les risques (NSN) du traitement.

## COMPRENDRE UNE ÉTUDE EN EXAMINANT SES STATISTIQUES DE BASE

Maintenant, considérons les résultats d'une étude examinant la réponse des enfants et des adolescents souffrant d'un trouble anxieux en utilisant ces concepts fondamentaux. Une étude appelée Child/Adolescent Anxiety Multimodal Study (CAMS) a recruté des enfants et des adolescents de plusieurs sites souffrant de troubles anxieux - anxiété généralisée, phobie sociale ou anxiété de séparation - et a randomisé ces jeunes pour qu'ils reçoivent l'une des quatre stratégies de traitement possibles: La TCC seule, la sertraline seule, la TCC combinée avec la sertraline, et le placebo.

Le graphique de la figure A.6.7 présente les scores de l'échelle d'évaluation de l'anxiété en pédiatrie (PARS) pour les 4 groupes, chacun recevant un traitement différent. Les scores supérieurs à 13 correspondent à un trouble anxieux. La combinaison de la sertraline et de la TCC, basée sur ces données, est le traitement le plus efficace. Les barres d'erreur sur les chiffres PARS montrent combien il y a de *dispersion* dans ces nombres. Le sommet d'une barre d'erreur est plus facilement compris comme la moyenne plus l'erreur-type de la moyenne (calculée comme l'écart type / racine carrée du nombre de participants), et le bas de la barre d'erreur est la moyenne moins l'erreur-type de la moyenne. Des méthodes plus complexes de calcul d'*erreur-type* représentée par les barres d'erreur existent également.

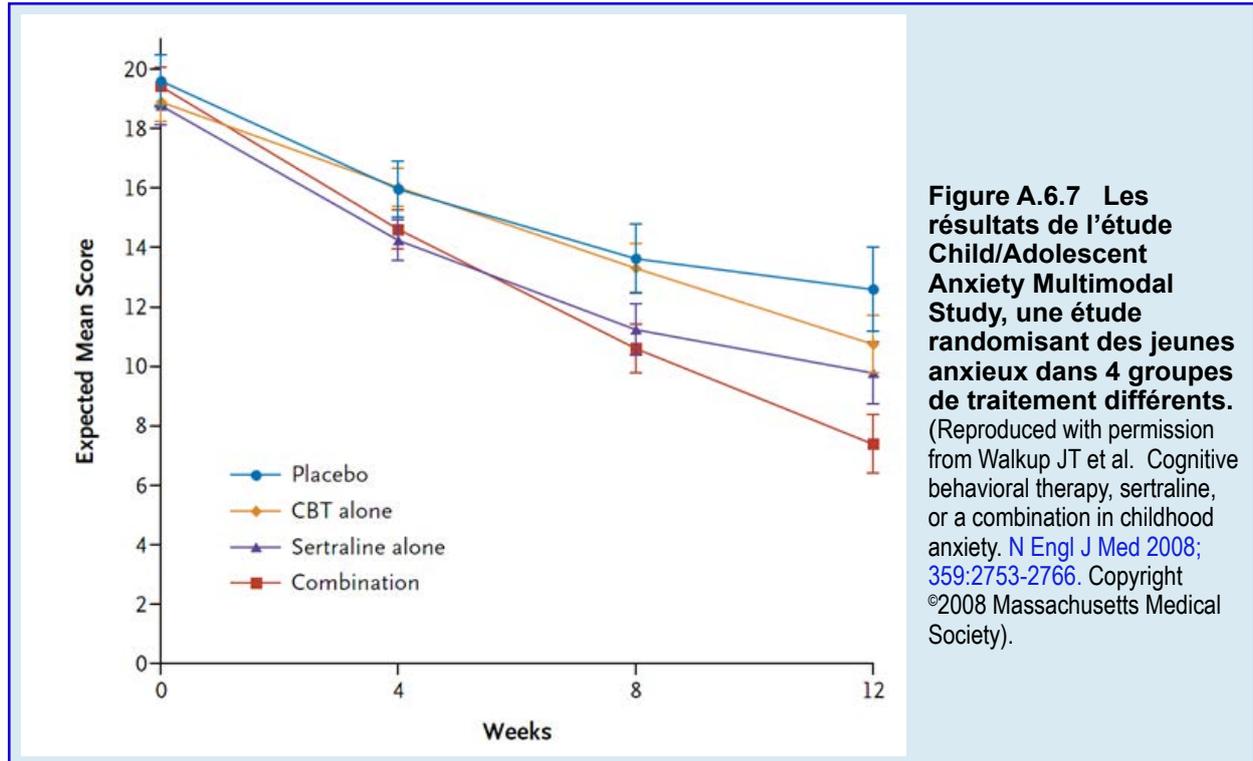
### Nombre nécessaire pour nuire (NSN)

Le NSN est égal à 1 divisé par le taux des non exposés moins le taux chez les individus exposés.

Dans l'étude TADS (Treatment for Adolescents with Depression Study) (2004), les auteurs ont rapporté que 11,9% des participants traités par la fluoxétine seule avaient signalé un événement indésirable, comparativement à 5,4% pour le placebo. Le NNH est de combien?

Dans ce cas  $NSN = 1 / (0,119 - 0,054) = 15,4$ .

En d'autres termes, quinze patients déprimés devront être traités avec de la fluoxétine pour qu'un patient supplémentaire subisse un événement indésirable lié au médicament.



Dans ce cas, l'ampleur clinique du traitement sur le pronostic a été évaluée en calculant la taille d'effet pour chacun des 3 groupes d'intervention par rapport au groupe placebo: est-ce que les 3 groupes d'intervention sont vraiment différents par rapport au groupe placebo? Sur la base d'améliorations à l'échelle PARS, la taille d'effet à 12 semaines était de 0,86 pour la thérapie combinée, de 0,45 pour la sertraline seule et de 0,31 pour la TCC seule.

Les auteurs ont également calculé le nombre de sujets à traiter. Rappelons que pour calculer un NST, nous devons choisir un «obstacle» que chaque sujet doit surmonter. Les «obstacles» plus hauts sont plus difficiles à « franchir » et donc généreront des NST plus élevés; des obstacles les plus bas génèrent des NST plus faibles. Les auteurs ont choisi comme difficulté à dépasser que les jeunes soient évalués comme améliorés ou très améliorés sur le PARS à la fin du traitement à 12 semaines. En utilisant ce critère, les auteurs ont calculé le NST pour la sertraline seule à 3,2 et pour la thérapie cognitivo-comportementale seule à 2,8 alors que pour la combinaison des deux traitements, le NST était à 1,7. Ce faible NST est encourageant car il démontre que des traitements efficaces existent pour de nombreux jeunes souffrant de troubles anxieux.

Les auteurs ont également utilisé des statistiques pour transmettre aux lecteurs la certitude qu'ils ont que leurs résultats n'étaient pas le résultat du hasard causés par des événements aléatoires qui ensemble ont entraîné des améliorations chez les jeunes recevant la sertraline et la TCC. Pour évaluer cette possibilité, les auteurs ont calculé les intervalles de confiance pour chaque taille d'effet et pour chaque nombre de sujets à traiter. Les intervalles de confiance à 95% de certitude autour de la taille d'effet calculée de 0,86 étaient de 0,56 à 1,15 pour les jeunes recevant la combinaison de la sertraline et de la TCC comparativement aux jeunes recevant le placebo, ce qui est un bon résultat en pédopsychiatrie. Les tailles d'effet de 0,8 et plus indiquent un effet important alors que une taille d'effet à 0,3 ou

inférieur indiquent un petit effet selon les conventions de Cohen (1988).

## FIABILITÉ ET VALIDITÉ DES MÉTHODES DE DIAGNOSTIC

### LE FONDEMENT D'UNE PRATIQUE FONDÉE SUR LES PREUVES

Harmoniser vos méthodes de diagnostic avec les méthodes de diagnostic utilisées par les chercheurs est important parce que les jeunes dans votre pratique que vous définissez comme ayant, par exemple, un épisode dépressif majeur, auront tendance à avoir des constellations similaires de symptômes dépressifs aux jeunes que les chercheurs ont étudié au moment des tests des traitements pour les troubles dépressifs majeurs. Vous pouvez alors utiliser en toute confiance leurs résultats sur ce qui aide ces jeunes dans votre propre pratique.

Alors, quelle fiabilité pour les procédures que vous utilisez pour identifier les syndromes chez les jeunes? Fiable signifie d'abord que si le test est utilisé à plusieurs reprises il donnera des résultats cohérents avec les résultats qu'il a déjà donné. Par exemple, si un entretien diagnostique bien défini et structuré produit des résultats identiques ou très similaires lorsqu'il est administré à nouveau peu de temps après la première administration, cela suggère que le test est fiable (fiabilité test-retest). Fiable signifie également qu'il donne les mêmes résultats ou très similaires quand il est utilisé par différents observateurs sur le même jeune (fiabilité inter-juge).

#### Fiabilité

La fiabilité est la cohérence obtenue par un instrument de mesure (par exemple, un test diagnostique ou un questionnaire) à travers les individus, le temps et les cliniciens. La fiabilité est mesurée en utilisant la mesure statistique kappa ( $\kappa$ ) pour les variables catégorielles (par exemple, le diagnostic) et le coefficient de corrélation intraclasse (ICC) pour les variables continues (par exemple, les scores d'échelle). Les valeurs pour les deux vont de 0 (pas mieux que le hasard) à 1 (accord parfait). Des fiabilités supérieures à 0,5 sont acceptables et supérieures à 0,7 sont excellentes.

#### Validité

La validité est la mesure dans laquelle un système de diagnostic mesure ce qu'il prétend mesurer. L'évaluation de la validité est souvent indirecte, en s'appuyant sur la capacité d'une mesure à s'accorder avec d'autres mesures qui évaluent le même concept, ou l'accord dans la prédiction du cours, des résultats, de la réponse au traitement, etc. Les coefficients de corrélation et les coefficients de saturation factorielle sont des moyens communs de quantifier au moins quelques types de validité.

**TABLE A.6.3 Exemple d'instruments de mesure fiables (voir Chapitre A.5 également)**

| Disorder  | Instrument   |
|---|--|
| Troubles anxieux y inclus TAG et trouble anxiété de séparation          | <ul style="list-style-type: none"> <li>Screen for Child Anxiety Related Emotional Disorders (SCARED)</li> <li>Voir aussi Chapitres F.1 et F.2</li> </ul>   |
| Troubles thymiques incluant le trouble dépressif majeur et la dysthymie | <ul style="list-style-type: none"> <li>Hamilton Rating Scale for Depression (HAMD-21)</li> <li>Johns Hopkins Depression Checklist for Children (HDCL-C)</li> <li>Voir aussi Chapitre E.1</li> </ul>                                |
| Trouble Déficit de l'Attention/ Hyperactivité                           | <ul style="list-style-type: none"> <li>Swanson Scale for ADHD (SWAN ADHD)</li> <li>The SNAP-IV Teacher and Parent Rating Scale</li> <li>Vanderbilt ADHD Diagnostic Parent Rating Scale</li> <li>Voir aussi Chapitre D.1</li> </ul> |
| Trouble obsessionnel compulsif  | <ul style="list-style-type: none"> <li>Children's Yale-Brown Obsessive Compulsive Scale (CY-BOCS)</li> <li>Voir aussi Chapitre F.3</li> </ul>  |

\*Ces instruments ont été développés et testés en Europe et aux États-Unis. Les syndromes dans des pays éloignés de ces zones géographiques peuvent être sensiblement différents. En outre, il faut se rappeler que les résultats d'un seul instrument ne sont qu'un aspect du processus de conclure voire simplement de tester une hypothèse initiale du diagnostic; De nombreuses autres sources de données telles que l'entretien avec le jeune et avec ses parents, ainsi que des informations sur son passé et la consultation des dossiers médicaux et psychiatriques sont également utiles.

Et quelle est la validité des systèmes de diagnostic que vous utilisez pour identifier les syndromes chez les jeunes? Un système de diagnostic valide est celui qui mesure vraiment ce qu'il prétend mesurer. Par exemple, un système de diagnostic valide pour évaluer la sévérité de la dépression est celui qui mesure réellement la sévérité de la dépression. Les preuves à l'appui de la validité d'un système de diagnostic de la sévérité de la dépression pourraient inclure des données montrant que les auto-rapports de sévérité de la dépression chez les adolescents sont fortement corrélés avec les rapports des soignants formés qui interviewent les mêmes adolescents utilisant ce système de diagnostic. L'accord diagnostique entre les entretiens diagnostiques structurés souvent utilisés dans les études de recherche et les évaluations cliniques, varie fortement selon le trouble et est faible à modéré pour la plupart des troubles (Rettew, 2008).

En résumé, un système de diagnostic fiable et valide est une base essentielle de la pratique clinique. Malheureusement, les entretiens cliniques non structurés ont un mauvais historique de fiabilité et de validité en ce qui concerne le diagnostic des jeunes atteints de troubles psychiatriques. Les cliniciens qui n'ont pas une approche diagnostique structurée tendent à tirer des conclusions hâtivement, omettent souvent d'évaluer des comorbidités importantes et aboutissent à des résultats diagnostiques qui ne sont pas fiables. C'est un problème majeur, mais qui peut heureusement être résolu.

### **Instruments de diagnostic fiables et valides**

La création d'un système de diagnostic fiable et valide est un problème difficile. Les aides les plus utiles pour créer un tel système sont des instruments de diagnostic spécifiques pour des classes spécifiques de troubles. Les aides diagnostiques sont souvent disponibles gratuitement sur Internet, sont très utiles pour définir les troubles et leur gravité et, surtout, servir de base de référence qui peut vous permettre de mesurer le changement des symptômes pendant le traitement. Le tableau A.6.1 montre des exemples de ce qui est disponible pour certains des troubles des plus courants.

### **Un exemple d'utilisation d'outils diagnostiques pour évaluer un adolescent.**

Il est important de traduire correctement les symptômes décrits par les jeunes, leurs parents et leurs enseignants dans un syndrome psychiatrique ou, du moins, dans une partie d'un syndrome (*trouble subsyndromique*). Cette étape prend les plaintes uniques d'un enfant et sa famille, les traduit dans un trouble connu, et essaie d'éviter les nombreuses sources possibles d'erreur dans le processus. Prenons, par exemple, Sam, un garçon de 14 ans dont les parents se plaignent qu'il est constamment irritable et a du mal à dormir. Ses notes baissent à l'école et se replie socialement. L'ensemble de ses symptômes peut s'inscrire dans plusieurs syndromes. Les adolescents peuvent devenir irritables lorsqu'ils se sentent anxieux ou montrent des conduites d'opposition, mais l'irritabilité suggère également la possibilité d'un trouble de l'humeur; son repli social peut également avoir des nombreuses origines possibles. Heureusement, des outils d'évaluation largement disponibles et peu coûteux sont utiles dans ce cas pour trouver le bon diagnostic (voir le tableau A.6.1).

Alors que ces instruments nous aident à évaluer plus en détails et à définir plus précisément les niveaux des symptômes dans les domaines suggérés

par l'histoire clinique, nous avons également besoin de se décentrer du symptôme initial et d'évaluer chez Sam d'autres syndromes moins évidents tels un traumatisme ou un abus, l'utilisation de substances ou un trouble du comportement qu'il a omis de préciser.

L'utilisation d'un outil d'évaluation complet et reconnu peut être utile ici. Le Kiddie-SADS (K-SADS) est une échelle d'entretien diagnostique semi-structuré, largement utilisée et généreusement placée dans le domaine public pour l'utilisation sans but lucratif et la recherche éthique (Kaufman et al, 1997). Notez que le K-SADS n'est pas un entretien structuré - dans lequel le clinicien répète des questions écrites mot à mot - mais plutôt un entretien semi-structuré dans lequel le clinicien a le choix parmi plusieurs questions possibles qui évaluent le même domaine. Les évaluateurs n'ont pas à poser toutes les questions, mais plutôt s'arrêter lorsqu'ils disposent d'informations suffisantes. Voici quelques questions de dépistage du K-SADS conçues pour détecter l'abus possible d'alcool:

- *Quel âge aviez-vous lorsque vous avez consommé de l'alcool pour la première fois?*
- *Quelle est votre boisson préférée?*
- *Avez-vous un groupe d'amis avec lequel vous consommez de l'alcool habituellement ou vous consommez de l'alcool habituellement seul?*
- *Où consommez-vous de l'alcool habituellement? À la maison? Dans des soirées? Dans la maison d'un ami? Dans la rue? Dans des bars?*
- *Y a-t-il des moments particuliers où vous êtes plus susceptible de consommer de l'alcool que d'autres? Soirées étudiantes ou autres soirées?*
- *Quel âge aviez-vous quand vous avez commencé à consommer de l'alcool régulièrement, disons deux verres ou plus par semaine?*
- *Au cours des six derniers mois, y a-t-il eu au moins une semaine pendant laquelle vous avez consommé au moins deux verres?*

Les auteurs de la K-SADS suggèrent un seuil clair pour continuer l'évaluation: le jeune ou le parent indiquent que le jeune consomme deux boissons

### **Quel est un bon moyen pour dépister un traumatisme ?**

Voici un aperçu de la manière dont l'entretien semi-structuré K-SADS interroge sur les traumatismes, mais seulement après une période "de préparation" utilisée pour recueillir des informations de base et mettre les jeunes à l'aise. Notez que l'utilisation correcte de la K-SADS nécessite une formation et la liste ci-dessous ne comprend que quelques-unes des questions sur le traumatisme (Kaufman et al, 1997)

#### **1. Événements traumatiques**

Question: Je vais vous poser des questions sur un certain nombre de mauvaises choses qui arrivent souvent aux enfants de votre âge et je veux que vous me disiez si une de ces choses vous est déjà arrivée. N'oubliez pas de me dire si certaines de ces choses se sont produites, même si elles ne se sont produites qu'une seule fois.

- a. Accident de voiture: Avez-vous déjà été impliqué dans un accident de voiture grave? Qu'est-il arrivé? Avez-vous été blessé? Est-ce que quelqu'un d'autre dans la voiture a été blessé?
- b. Autre accident: Avez-vous déjà été impliqué dans un autre type d'accident grave? Dans un accident de vélo? Dans d'autres accidents? Qu'est-il arrivé? Avez-vous été blessé?
- c. Feu: Avez-vous été impliqué dans un incendie? Votre maison ou école a-t-elle jamais pris feu? Avez-vous déjà commencé un incendie qui a échappé à tout contrôle? Qu'est-il arrivé? Quelqu'un a-t-il subi des blessures physiques graves? Y a-t-il eu beaucoup de dégâts?

Le K-SADS continue avec des questions sur le témoignage d'une catastrophe comme un orage violent ou un crime violent, de la violence domestique, ou des abus physiques ou sexuels.

alcoolisées ou plus par semaine pendant au moins 4 semaines. Les seuils explicites du K-SADS pour considérer un symptôme comme présent ou absent sont utiles, de même que les questions proposées. Elles sont claires et directes, mais offrent au clinicien des multiples façons d'obtenir les informations nécessaires. La spécificité des questions du K-SADS, leurs seuils clairs et la présence de questions de dépistage sont utiles aux cliniciens qui tentent d'améliorer la fiabilité et la validité de leurs entretiens cliniques. L'utilisation de toute l'échelle K-SADS nécessite une formation spécifique et peut être irréaliste dans de nombreux contextes cliniques, mais ses domaines d'évaluation, ses questions proposées et ses seuils sont utiles.

Une fois que l'évaluation diagnostique de l'adolescent est satisfaisante, fiable et valide, la prochaine étape est le choix du traitement.

## CHOISIR UN TRAITEMENT

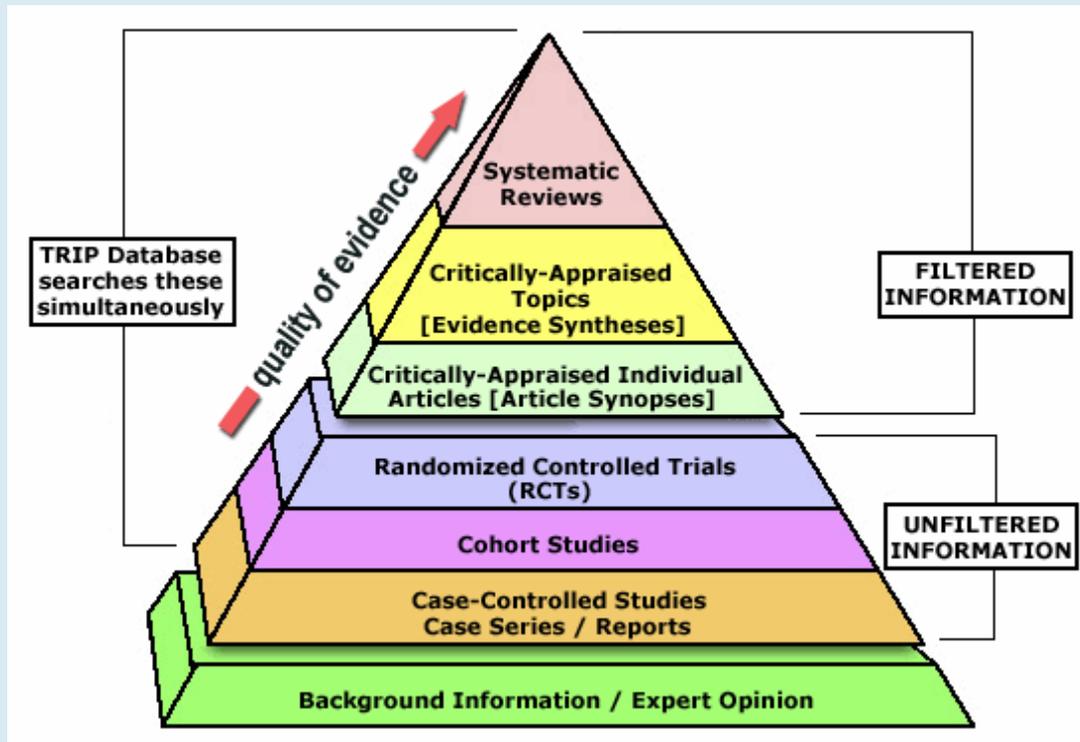
### Le concept d'une pyramide de preuves

Une *pyramide de preuves* est un moyen utile de résumer les meilleures preuves, c'est-à-dire les plus proches d'une description précise de la vérité. Les sources de preuves avec moins de biais et moins d'erreurs aléatoires sont placées vers le haut de la pyramide et sont les plus crédibles. Les pyramides sont une bonne représentation de la transition de centaines voire des milliers de sources de «preuves peu solides» telles que des études de cas, des essais non contrôlés et même des opinions d'experts -placées à juste titre près de la base de la pyramide - vers les peu nombreuses revues systématiques d'essais contrôlés randomisés et des méta-analyses qui synthétisent toutes les preuves les plus crédibles - placées à juste titre au sommet de la pyramide. Comme il existe des nombreux case reports et d'autres études susceptibles de contenir des biais et des erreurs aléatoires et relativement peu d'études qui synthétisent judicieusement toutes les preuves crédibles, la forme d'une pyramide avec la base large et le sommet étroit est particulièrement appropriée. (Figure A.6.8)

Les pyramides sont une bonne forme pour représenter les différents types de preuves, car elles sont larges à la base et étroites au sommet. Notre but est d'arriver au sommet, où la vue est bonne, minimisant ainsi les biais et l'erreur aléatoire.



**Figure A.6.8** Une pyramide des preuves où plus haut sur la pyramide signifie moins de biais et d'erreur aléatoire.



Source: University of California, San Diego, library system and its [TRIP database](#) which uses 'filters' (software programs) to select published evidence most free from bias and random error.

### *Information générale*

En bas, les informations de base sur le sujet, comme la neurophysiologie de la dépression ou l'autisme, par exemple, et l'opinion des experts ne fournissent que des preuves faibles. Les informations de base ne font que "préparer la scène" pour ce qui pourrait arriver: de nombreuses pièces de théâtre peuvent se jouer en utilisant la même scène. Si, par exemple, les études d'imagerie montrent des changements dans le flux sanguin des régions corticales frontales dans la dépression, cela ne nous aide pas nécessairement à choisir le traitement le plus efficace.

### *Avis des experts*

De même, l'*avis des experts* peut être complètement faux. L'expert peut être employé par une compagnie pharmaceutique et ressentir la pression de privilégier un certain médicament, ou peut ressentir de la honte au sujet des résultats négatifs, ou de l'orgueil au sujet des résultats positifs et incliner, inconsciemment, les résultats dans une direction positive, ou simplement être fidèle à un certain traitement parce il lui est familier. Le point principal est que l'opinion d'un seul individu non étayée par des données systématiques recueillies selon une méthodologie reconnue constitue une preuve faible. L'humour familier formule cette conclusion comme: «A Dieu, nous faisons confiance. Le reste du monde, apportez des données.»

### *Séries de cas*

Les séries de cas et les case reports se trouvent également au bas de la pyramide des preuves. Un praticien isolé évaluant un seul cas, effectuant le

traitement, puis évaluant les résultats est très vulnérable aux biais. Les séries de cas produisent des données de niveau faible parce qu'il n'y a pas de groupe de comparaison, bien que leurs données puissent être utiles aux chercheurs qui génèrent des hypothèses pour des futures études.

### ***Études cas-témoin***

Les études cas-témoin peuvent être utiles pour montrer une association entre l'exposition à un facteur de risque et un devenir. Dans les études cas-témoins, le devenir est connu, par exemple le suicide chez l'adolescent. L'investigateur crée alors un groupe témoin d'individus sans ce devenir mais présentant des caractéristiques similaires aux individus ayant le devenir connu et détermine le taux d'exposition à un facteur de risque chez les témoins par rapport à ceux dont le devenir est connu. En utilisant cette méthodologie, on a montré que l'accessibilité et la disponibilité d'une arme à feu dans la maison étaient associées à un risque accru de suicide chez les adolescents (Brent et al., 1993). C'est une preuve utile lorsqu'on conseille les parents d'un adolescent potentiellement suicidaire.

### ***Études de cohorte***

Les études de cohorte peuvent être utiles pour démontrer de façon prospective une association entre l'exposition à un facteur de risque spécifique et un devenir. Ici l'échantillonnage est par l'exposition, et non pas par le devenir. Les enfants exposés à certains types d'abus, par exemple, sont connus pour être plus à risque de développer une dépression. Certaines études de cohortes utilisent des très larges bases de données, incluant même tous les enfants d'un pays entier, générant donc une grande puissance statistique (la capacité de détecter même un petit effet)

### ***Essais contrôlés randomisés***

Les essais randomisés contrôlés (ECR) sont les principaux éléments constitutifs des niveaux supérieurs de la pyramide. Les ECR ont des avantages uniques: un groupe de contrôle choisi au hasard et, souvent, le maintien en aveugle des sujets, des cliniciens et des évaluateurs quant à l'affectation à l'intervention active ou au groupe témoin. Les niveaux plus élevés de la pyramide sont conçus pour éliminer les limites des données ECR, comme l'élimination des données provenant d'ECR mal effectués, tout en combinant des données provenant de plusieurs ECR bien conçus pour générer un plus grand nombre de sujets étudiés afin de minimiser les erreurs aléatoires. Méfiez-vous des ECR non publiés.

### ***Évaluation critique***

L'évaluation critique d'articles individuels élimine les données d'essais avec des défauts significatifs. Peut-être qu'un tel essai n'a pas suffisamment maintenu l'aveugle des évaluateurs, par exemple, ou a perdu une grande proportion de l'échantillon au suivi. Si tel est le cas, les résultats peuvent comporter un biais significatif et les études sont exclues. Les résultats d'articles individuels évalués de manière critique sur le même sujet, maintenant connus pour être exempts de sources évidentes de biais, peuvent alors être regroupés pour augmenter la taille de l'échantillon afin de réduire l'erreur aléatoire. Envisagez d'utiliser des listes spécifiques de critères lors de l'évaluation d'une étude, comme la liste de vérification des normes uniformisées pour le rapport des essais (CONSORT), si vous en avez le temps.

CONSORT est synonyme de normes uniformisées pour rapporter les essais.

[Cliquez ici](#) pour accéder à la liste de vérification CONSORT des informations à inclure lors de la déclaration d'un essai randomisé.

Le diagramme de la figure A.6.9 (page 20) montre comment certaines étapes critiques de l'exécution d'un ECR méritent d'être prises en compte lorsque vous lisez un article. Parfois, par exemple, un auteur peut décrire une étude comme un «essai randomisé» même s'il est clair dans la section des méthodes que l'assignation n'était pas vraiment aléatoire. Il existe de nombreuses sources potentielles de biais dans la réalisation d'un ECR, mais la plupart des praticiens n'ont tout simplement pas la volonté, le temps ou l'expertise statistique et de recherche pour lire attentivement un ECR publié. L'évaluation critique implique un examen critique de chaque étape d'une étude pour s'assurer qu'aucun biais n'influence les résultats finaux. Participer à un cours de médecine basée sur des preuves et participer à un club de lecture de revues pour les personnes intéressées peut aider à améliorer ses compétences d'évaluation critique.

### ***Synthèse des preuves et revues systématiques***

La synthèse des preuves regroupe les données provenant d'ECR individuels qui ont fait l'objet d'une évaluation critique pour déterminer les sources possibles de biais. La mise en commun des données tend à diminuer l'erreur aléatoire pour produire un résultat plus précis.

Enfin, les revues systématiques sont une méthode particulièrement organisée de synthèse des preuves reposant sur des normes de synthèse convenues. Avec cette méthodologie, peut-être le mieux illustrée par la bibliothèque électronique des revues systématiques de la Bibliothèque Cochrane, la méthode utilisée pour combiner les données et tirer des conclusions est bien définie avant le début de la revue. Cette méthodologie est également transparente, dans le sens que le lecteur peut examiner la façon dont les résultats ont été obtenus. Une méta-analyse est un type de revue systématique qui prend en compte les données de plusieurs études pour arriver à une seule estimation de la taille d'effet ou du NST.

## **UTILISER L'INTERNET ET LA PYRAMIDE POUR TROUVER DES TRAITEMENTS EFFICACES**

### **Répondre aux questions pertinentes de votre pratique en utilisant l'Internet**

Essayez de rester curieux quand vous faites un travail clinique : posez-vous des questions pertinentes pour le soin des patients que vous voyez ce jour-là. Cela signifie poser des questions pertinentes pour ce cas spécifique et essayer de répondre en utilisant l'Internet pour regarder la littérature publiée, en se concentrant sur les preuves qui n'ont pas de biais ni d'erreur aléatoire.

La première étape consiste à poser une question. Pour y parvenir, commencez la question par «Quelle est la preuve que ...?» Cela permet de répondre à la question: vous constaterez que soit il n'y a pas de preuve, dans quel cas la réponse est «aucune» ou vous trouverez au moins quelques preuves.

L'objectif d'une recherche est un résultat pratique, assez bon pour apporter de l'information pour les soins sans prétendre être définitif, dans le temps et les ressources disponibles. La technologie permet de rechercher des bases de données pour des résultats valables et pertinents. Le logiciel rend facile les recherches sensibles et spécifiques sur [PubMed](http://www.pubmed.gov), le site Internet de la National Library of Medicine aux États-Unis. PubMed est maintenu par le gouvernement des États-Unis sans frais pour les utilisateurs et aucun abonnement n'est nécessaire. [PsycINFO](http://www.psycinfo.org), le site



Le logo de PubMed, une source gratuite de preuves. Il dispose de filtres pour trouver des revues systématiques, utiles comme résumés de preuves de haut niveau.

Internet pour les études psychologiques, est maintenu par l'American Psychological Association et offre des résumés sans frais pour les utilisateurs. **CENTRAL**, la base de données de la Cochrane Library d'essais contrôlés, est basée au Royaume Uni. CENTRAL est l'abréviation utilisée pour le Cochrane Central Register of Controlled Trials; Il contient des résumés de près de 700 000 essais contrôlés à la fois de PubMed et **EMBASE**, une base de données très présente dans les revues européennes. EMBASE est accessible uniquement aux abonnés et est géré par Elsevier, une société basée à Amsterdam. PubMed est un bon point de départ pour les questions liées aux médicaments et inclut des études qui ne sont pas des ECR. PsycINFO est utile pour les questions liées aux problèmes psychosociaux. CENTRAL est utile pour chercher exclusivement des résultats des ECR.

**Terry: une recherche de 10 minutes pour une solution pour un garçon avec TDAH, type trouble de l'attention.**

Un pédopsychiatre essaie d'aider un garçon de 10 ans avec TDAH type trouble de l'attention. Sur la base des résultats SWAN de son professeur (voir le tableau A.6.3), il a besoin de méthylphénidate 15 mg trois fois par jour pour obtenir une amélioration substantielle (une note de moins de 9 sur les neuf premières questions SWAN, les items de l'inattention). Mais il a tellement perdu du poids sur ce traitement qu'il a dû arrêter. Maintenant, ses parents sont de plus en plus frustrés et cherchent une solution: «Y a-t-il rien d'autre?» Son pédopsychiatre veut un médicament autre qu'un stimulant ou l'atomoxetine - tous deux baissent l'appétit. Notre médecin se rappelle combien la guanfacine a été utile dans le passé pour le traitement des garçons TDAH avec Trouble oppositionnel avec Provocation et des symptômes d'hyperactivité / impulsivité. Mais est-ce que la guanfacine est efficace dans le TDAH type inattention? Est-ce qu'il y a des preuves que la guanfacine est efficace chez les garçons de 10 ans atteints de TDAH type inattention? Si elle est efficace, quelle est la taille d'effet?

1. Un lieu raisonnable pour commencer est [PubMed](#). Cliquez sur «Advanced», sélectionnez dans le menu déroulant «MeSH term», tapez «attention deficit hyperactivity disorder» dans le champ de recherche et cliquez sur «Search». Le résultat devrait ressembler à la figure ci-après. Il y a trop de résultats pour pouvoir les lire tous vous-même - plus de 16 000! Nous avons besoin d'un moyen de limiter les résultats à un nombre gérable en choisissant seulement les résultats du haut de la pyramide afin de voir les meilleures preuves. Voici comment.
2. Commencez au sommet de la pyramide: recherche des méta-analyses publiées. Revenez à la page de recherche et, après avoir répété l'étape précédente, tapez «guanfacine» dans le champ de recherche, les cases «AND» et «All Fields» et cliquez sur «Search». Il en résulte 94 résultats. Maintenant cliquez sur «Limits» et sélectionnez «Meta-Analysis» dans le nouveau menu et cliquez «Search»; 2 résultats, mais ce garçon n'appartient pas à la population étudiée par les deux articles: l'un concerne les adultes, l'autre, les enfants atteints de TDAH et les tics, et notre patient a un TDAH sans tics (voir la figure ci-après). Nous sommes devenus trop sélectifs et nous avons besoin d'élargir la recherche.
3. Descendez d'une marche dans la pyramide vers les ECR en changeant la case «Limits» de «Meta-analysis» en «Randomized controlled trials» et 9

**Recherche PsycINFO**

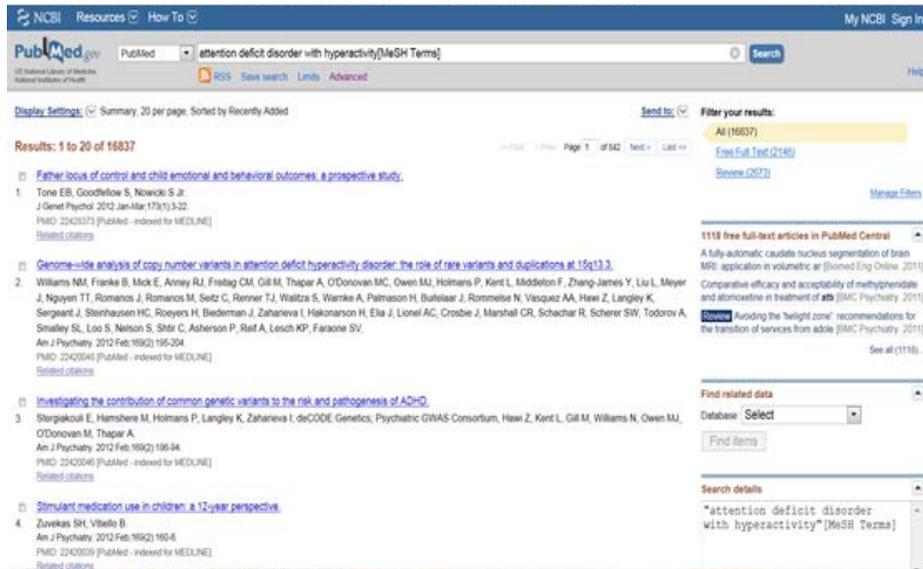
PsycINFO est un excellent choix de base de données pour les recherches de constructions comme «bullying» ou «identity», car il utilise un thésaurus pour les mots clés. En recherchant le terme «bullying», par exemple, les mots clés suivants sont énumérés dans une colonne à gauche des résumés, ainsi que le nombre de résultats pour chaque mot clé.

Mot clé:

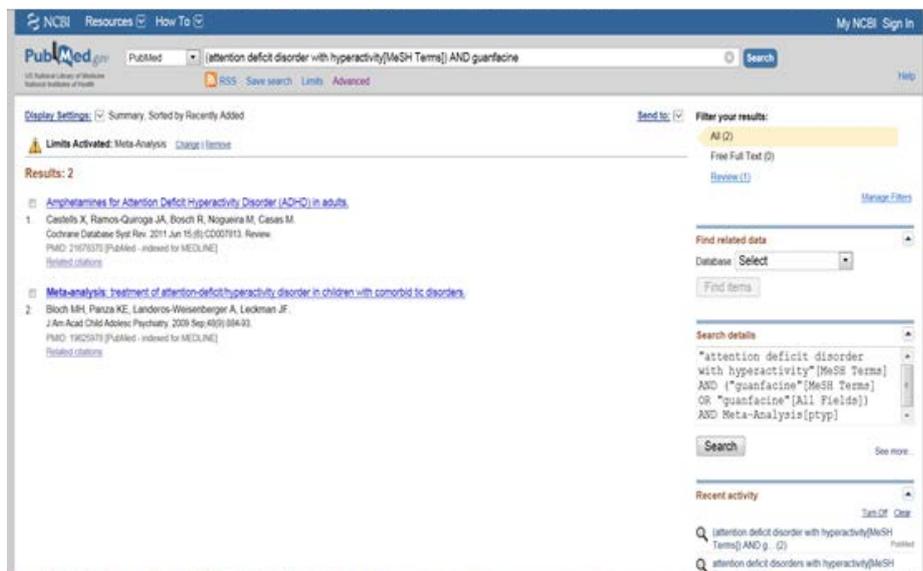
- Bullying (2,808)
- Victimization (1,175)
- Aggressive Behavior (845)
- Peer Relations (543)
- Schools (437)
- 5 more...

En cliquant sur l'hypertexte du site Web, vous avez accès aux résumés disponibles gratuitement. Utilisez l'onglet «Advanced Search» pour rechercher «bullying» et choisissez «Only Show Content Where...» pour inclure le menu déroulant «Treatment Outcome/ Randomized Controlled Trial». Cela réduit 2808 résultats à seulement 10 (Octobre 28, 2011).

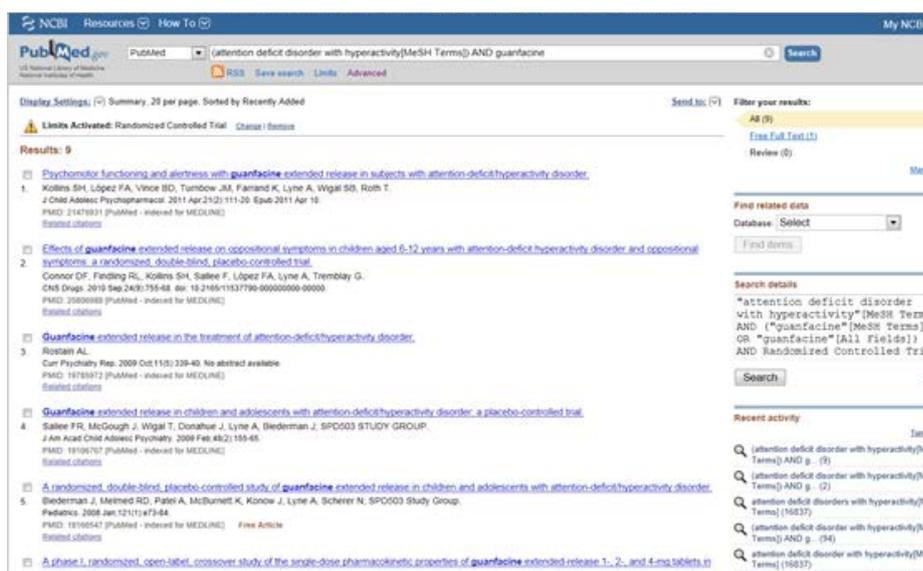
Comme exercice, essayez de trouver tous les essais randomisés qui étudient l'effet de l'apprentissage de la relaxation chez les adolescents.



étape 1 :  
trop de résultats!



étape 2 :  
trop sélectif!



étape 3 :  
9 résultats

articles apparaissent (voir la figure ci-avant).

Les articles 4 et 5 sont les seuls pertinents pour un garçon de 10 ans, et seul l'article 5 est disponible en texte intégral gratuitement (Biederman et al., 2008). Cet article a été financé par des sociétés pharmaceutiques, une source potentielle de biais, et concerne la guanfacine à libération prolongée, qui est coûteuse et n'est pas disponible pour ce garçon. Une recherche rapide sur Internet révèle que la demi-vie de la guanfacine chez les enfants et les adolescents est de 13 à 14 heures, suggérant qu'il est possible d'imiter la préparation à libération prolongée avec une préparation abordable - la guanfacine elle-même - prise au moins deux fois par jour (Strange, 2008).

Quelle est l'importance de l'effet de la guanfacine sur les symptômes d'inattention chez les jeunes atteints de TDAH de type mixte parmi les enfants étudiés? Les auteurs notent que la taille d'effet dans les groupes de traitement guanfacine calculés post hoc étaient respectivement de 0,58, 1,19 et 1,34 pour les groupes de 0,05 à 0,08 mg / kg, de 0,09 à 0,12 mg / kg et de 0,13 à 0,17 mg / kg. Cependant, ces tailles d'effet étaient pour les changements dans les scores totaux du TDAH (symptômes d'inattention et d'hyperactivité / impulsivité). Heureusement, les auteurs précisent que «tous les groupes d'enfants prenant la guanfacine ont montré une amélioration significative sur les sous-échelles d'hyperactivité / impulsivité et d'inattention de l'échelle ADHDRS-IV comparativement au placebo. Les changements de moyenne par rapport à l'état de base d'hyperactivité / impulsivité pour le placebo et les groupes guanfacine de 2 mg, de 3 mg et de 4 mg étaient respectivement de -4,06, -6,94, -7,09 et -9,46. Les changements de moyenne pour l'inattention étaient de -4,78, -8,46, -8,71 et -9,51 respectivement. Même si aucune taille d'effet n'est proposée concernant les changements de l'inattention seule, ces données suggèrent que la taille d'effet sera au moins égale à la taille d'effet pour les symptômes inattentifs et hyperactifs / impulsifs combinés.

Il est raisonnable de conclure que le garçon de 10 ans avec un TDAH de type inattention peut répondre par une réduction significative des symptômes à la guanfacine titrée à une dose de 1,0 à 1,5 ou 2,0 mg deux fois par jour.

### Stratégies Internet pour trouver de l'information concernant un vaste domaine

Plutôt que de chercher des preuves sur une question clinique spécifique concernant un jeune, imaginez que vous écrivez des directives pour votre pays sur la façon de traiter les enfants et les adolescents atteints de schizophrénie. Une question pertinente pourrait être: «Quelle est la preuve qu'il existe des traitements efficaces pour les jeunes atteints de schizophrénie?» Commencez par trouver les données les plus récentes et les plus pertinentes provenant d'essais randomisés. Commencez par chercher sur PubMed des études concernant le traitement chez les enfants et les adolescents atteints de schizophrénie. Voici comment le faire (notez que le nombre de résultats changera avec le temps avec l'ajout des nouvelles études).

1. *Trouvez les mots clés du sujet médical correspondant à votre domaine de recherche.* Entrez «Schizophrenia» dans la barre de recherche PubMed après avoir utilisé le menu déroulant pour changer le paramètre à «MeSH». Cela signifie que vous cherchez maintenant parmi les mots du vocabulaire de termes descripteurs utilisés par la National Library of Medicine aux États-



Cliquez sur l'image pour afficher un didacticiel vidéo sur la façon de chercher sur PubMed (2:27)

### Faire une recherche dans la base de données CENTRAL de la bibliothèque Cochrane

Chercher dans CENTRAL le mot clé «bullying» donne 36 résultats, tous des ECR étudiant l'intimidation. Un ECR fait en Australie, par exemple, en 2009, montre des résultats prometteurs chez les garçons adolescents, et n'est pas apparu dans la recherche PsycINFO pour les ECR de «bullying». Berry K and Hunt CJ (2009), Evaluation of an intervention program for anxious adolescent boys bullied at school. *Journal of Adolescent Health* 45:376-382). CENTRAL est utile pour récupérer des ECR de partout dans le monde et rendre les résumés disponibles gratuitement pour tout le monde.

Unis. Ces termes médicaux sont utilisés pour catégoriser chaque article publié dans la littérature médicale, ce qui permet de retrouver tous les articles sur le même sujet. Dans ce cas, plusieurs (12) termes MeSH sont récupérés, communément appelés 12 résultats. Chacun contient le mot «schizophrenia» soit dans le titre soit dans la description.

2. Envoyez les termes pertinents à la zone Search Builder, en les combinant avec «OR». Seuls 6 apparaissent comme pertinents puisque les 6 derniers concernent une protéine atypique. Après avoir sélectionné ces 6, cliquez sur «Add to Search Builder».
3. Ajouter à la recherche une version tronquée (tw) du concept central en tant que mot textuel avec un «OR». Dans ce cas, ajoutez «Schizophren \*» (tw) dans la barre de recherche avec l'opérateur «OR». Cette étape récupérera tout article contenant un mot commençant par les lettres «Schizophren» dans le titre ou dans l'abrégié. Ce terme supplémentaire peut aider à trouver des articles qui n'ont peut-être pas été correctement classés dans l'un des termes MeSH comme contenant de l'information sur la schizophrénie, ainsi que des articles utilisant uniquement des termes tels que «schizophrénique» ou «schizophréniques»
4. Chercher sur PubMed avec cette chaîne de termes «OR». 125889 résultats sont retrouvés. Il est temps maintenant de devenir plus spécifique et commencer à réduire ce grand ensemble de références au seulement les plus pertinentes.
5. Cliquez sur «Limits» et cochez la case «Children 0 to 18» 22545 résultats restent.
6. Cliquez sur «Limits» et cochez les cases «Meta-analysis» et «Randomized controlled trial». Il reste 945 résultats. Il peut s'agir d'un nombre raisonnable de documents à examiner pour un article important qui donnera les directives pour un pays entier. Deux ou trois personnes pourraient analyser les titres et les résumés et choisir les plus pertinents. Vous pouvez également revenir à «Limits» et choisir la case «Published in last 5 years», ce qui réduit encore le nombre de références. Notez également que de nombreuses références concernent uniquement les adultes, pas les enfants ou les jeunes, et peuvent être abandonnées immédiatement.

Voici une alternative si vous n'avez pas le temps de regarder à travers tant de résultats:

1. Revenez à l'étape 1 et notez que l'un des termes MeSH renvoyés est «Schizophrenia, childhood». Envoyez uniquement le terme «Schizophrenia, childhood» [MeSH] dans le champ de recherche, en demandant au moteur de recherche de ne chercher que ce terme MeSH unique. 1434 résultats apparaissent.
2. Appliquez maintenant les limites «Meta-analysis» ou «Randomized controlled trial». 8 résultats restent. Ces résultats doivent être lus, mais les limites semblent avoir éliminé trop de résultats.
3. Revenez à l'étape 2 avec 1434 résultats et appliquez les limites «Last 10 years». 59 résultats restent. Explorer ces résultats révèle une série d'articles très

**EvidenceUPDATES**  
FROM THE BMJ EVIDENCE CENTRE

Home My Profile My Alerts Search

Home About This Site About BMJ Group

### Se tenir à jour avec la médecine basée sur les preuves

Plus de 5000 articles biomédicaux et les résultats d'environ 100 essais de traitement sont publiés dans le monde chaque jour et le nombre continue de croître. Se tenir à jour sur cette marée d'information est intimidant. Pour faciliter cette tâche, vous pouvez vous abonner sans frais au site [Evidence Updates](#) géré par le Groupe BMJ et l'Unité de recherche sur l'information concernant la santé de l'Université McMaster. Ces établissements collaborent pour que les cliniciens aient accès aux meilleures données actuelles issues de la recherche, adaptées aux intérêts propre de chaque clinicien, pour appuyer les décisions cliniques fondées sur des preuves. Ce service scanne plus de 120 des meilleures revues médicales et plus de 50 000 articles chaque année, les classant en fonction de la validité et de la pertinence, et envoie régulièrement aux abonnés un courriel avec les articles clés de leur spécialité

pertinents sur des sujets tels que «Les effets de la clozapine et de l'olanzapine sur l'épaisseur corticale dans la schizophrénie débutée durant l'enfance» et «La perte d'asymétrie est locale plutôt que globale dans la schizophrénie débutée durant l'enfance». Notez la réduction dramatique du nombre de résultats, ce qui suggère que - comme le site PubMed nous informe - «Schizophrenia, childhood» comme terme MeSH est avant tout un terme historique; Pourtant les 59 résultats tirés des 10 dernières années semblent très pertinents et intéressants pour un document de prise de position.

Gardez un œil sur le nombre de résultats qu'une stratégie produit dans le cadre des objectifs de votre recherche : les résultats sont-ils gérables et adaptés à vos besoins? Si vous sentez qu'ils ne sont pas gérables (trop de résultats et trop peu de temps), continuez à tailler de nouveaux résultats avec d'autres limites. Si vous sentez qu'ils ne sont pas suffisants, changez votre stratégie de recherche initiale pour être plus sensible. Dans cet exemple, ajouter le terme «Schizophren\*» donne plus de 20.000 résultats. Une autre alternative pour augmenter la sensibilité est de cliquer sur l'hypertexte «Citations liées» placé sous les citations, qui semblent particulièrement utiles ; le logiciel rapporte une série de citations de sujet similaire à la citation que vous trouvez utile.

## LA MEDECINE BASEE SUR LES PREUVES (EBM) CLASSIQUE ET LA 'QUESTION SUSCEPTIBLE DE REPONSE'

«EBM classique» est notre expression pour l'approche développée à l'origine par David Sackett et ses collègues de l'Université McMaster au Canada et plus tard dans de nombreux autres endroits. Leur travail a été fondamental pour l'approche discutée ici. Nous distinguons notre approche, qui est adaptée à la réalité du travail en santé mentale des enfants, de «EBM classique», qui a été initialement développée dans le contexte de la médecine interne. Il est important de connaître la «EBM classique» comme précurseur de ce qui est discuté ici.

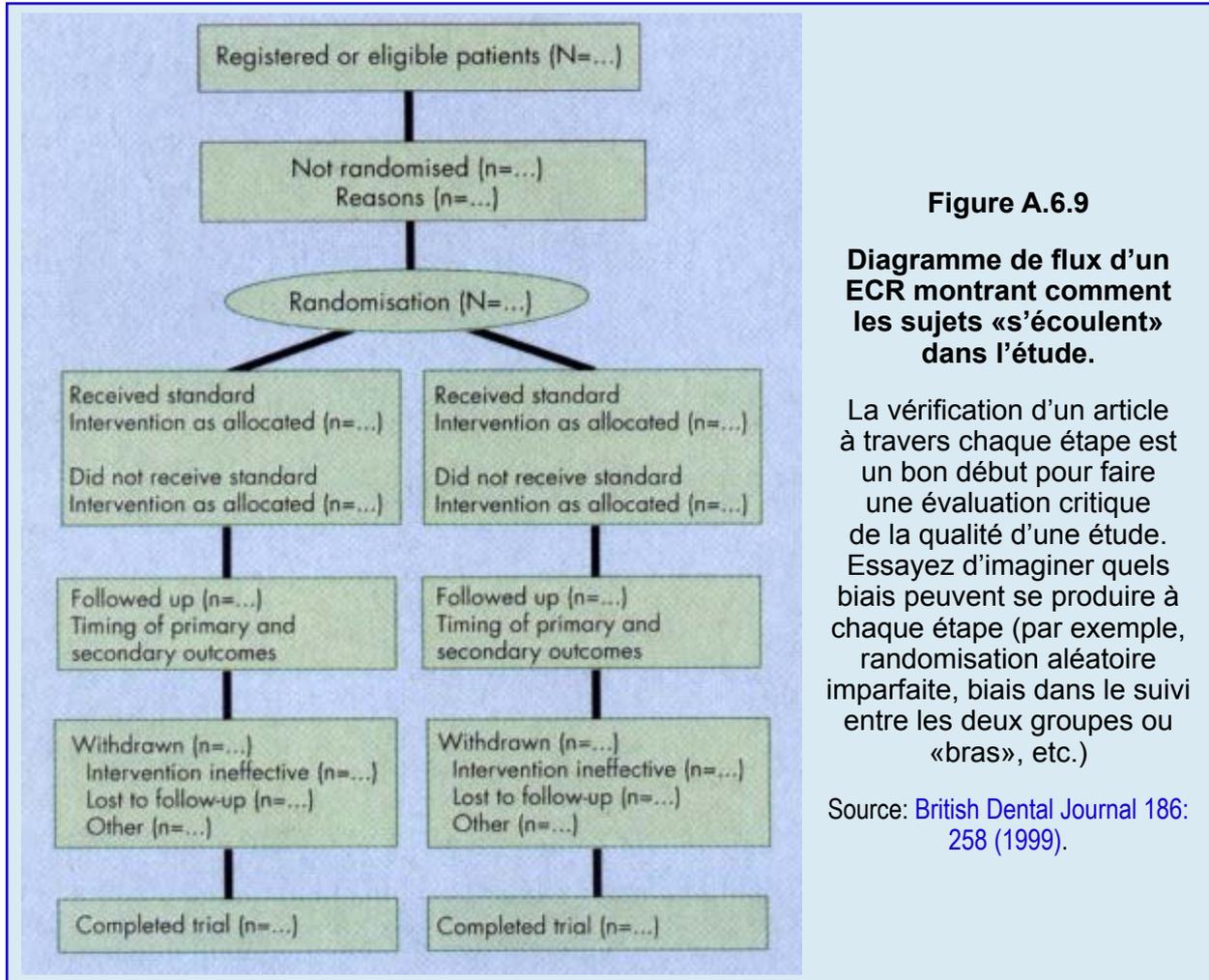
«EBM classique» peut être résumé comme un modèle en cinq étapes:

1. Poser une question «susceptible d'avoir une réponse» sur une situation clinique, en commençant généralement «Quelle est la preuve que ...» un patient spécifique ou un problème, traité avec une intervention spécifique par rapport à une autre intervention ou un placebo donne quels résultats? Ces questions sont parfois appelées PICO questions (Patient / Intervention / Comparaison / Résultat ou Outcome en anglais).
2. Recherche de preuves pour répondre à la question PICO.
3. Évaluation critique de la preuve en posant des questions sur la façon dont les études ont été réalisées.
4. Intégrer l'expertise clinique et les valeurs du patient avec la preuve pour prendre une décision concernant le traitement.
5. Évaluer le résultat.

«EBM Classique» prend également en compte des questions aussi importantes que la sensibilité et la spécificité des tests diagnostiques et la manière d'appliquer la preuve (Sackett et al, 2000).



Cliquez sur l'image pour accéder au site Centre for Evidence Based Medicine, qui contient des informations et de la formation sur l'EBM



**Figure A.6.9**

**Diagramme de flux d'un ECR montrant comment les sujets «s'écoulent» dans l'étude.**

La vérification d'un article à travers chaque étape est un bon début pour faire une évaluation critique de la qualité d'une étude. Essayez d'imaginer quels biais peuvent se produire à chaque étape (par exemple, randomisation aléatoire imparfaite, biais dans le suivi entre les deux groupes ou «bras», etc.)

Source: [British Dental Journal 186: 258 \(1999\).](#)

## PRATIQUES BASEES SUR LES PREUVES ET LA CULTURE LOCALE

La culture - laissée non définie dans ce chapitre - est un facteur important qui colorie la manière dont les idées discutées dans ce chapitre sont reçues par les cliniciens et les patients. Les éléments culturels qui influent sur nos pratiques comprennent les croyances et les valeurs, les pratiques d'éducation des enfants, le rôle de la famille élargie et l'impact de l'Internet.

### Croyances et valeurs

Les croyances concernant le traitement des maladies mentales sont importantes et il est crucial pour les cliniciens de comprendre la culture dans laquelle ils travaillent afin de permettre l'acceptation d'interventions efficaces. Dans les sociétés où la stigmatisation des malades mentaux est plus sévère, il peut être très difficile de convaincre un enfant et sa famille de participer même à un premier entretien psychiatrique. Même s'ils viennent, il peut être très difficile pour eux d'accepter le diagnostic ainsi que le traitement suggéré. Dans certaines cultures, les maladies mentales sont considérées comme étant spirituelles et donc les traitements spirituels sont le premier choix. Les familles peuvent se conformer au traitement seulement si le guérisseur spirituel approuve. Dans un cas de ce type, un guérisseur spirituel turc avait préparé un muska - un collier magique de

papier et de tissu enveloppé pour se débarrasser de la maladie - pour un adolescent psychotique que le guérisseur avait envoyé pour une hospitalisation. Au cours de son traitement, l'adolescent ouvrit le muska. Il ne trouva qu'un mot griffonné: *Haldol*.

Pourtant, les patients recherchent souvent des remèdes magiques à effet immédiat. Il est très difficile parfois d'expliquer, et de les faire accepter, le temps nécessaire pour un bon traitement, ou le conseil d'utiliser un médicament malgré ses effets secondaires potentiels. Dans les cultures où les gens demandent d'abord à leurs voisins ou à leurs proches de conseils concernant le problème et s'appuient sur leurs expériences personnelles ou de tierces parties, le traitement basé sur les preuves suggéré par un clinicien peut ne pas toujours correspondre aux suggestions des voisins et acquiert souvent un statut «à ne pas utiliser» (l'enfant du voisin a bien réagi à un autre médicament, ce qui signifie que celui conseillé peut ne pas être le bon).

### **Croyances concernant les pratiques d'éducation des enfants**

Des croyances comme «les garçons montrent normalement une activité motrice élevée» peuvent être une raison pour une famille de refuser les médicaments pour traiter le TDAH, parce que la famille pense que le médicament ralentit l'enfant et le rend «anormal». Une autre croyance fréquente est que «les enfants ne doivent pas prendre des traitements sur des longues périodes de temps» ce qui entraîne souvent l'interruption du traitement par les parents. Ou, «Grand-mère dit qu'il est comme son père quand il était enfant, et son père se porte bien. Vous êtes sûr qu'il a besoin d'un traitement?»

### **Rôle de la famille élargie**

Dans les cultures traditionnelles, les grands-parents ou les frères ou sœurs aînés des parents peuvent être en charge des décisions de traitement. Ils peuvent refuser d'autoriser les parents à administrer le traitement suggéré. Dans certains cas, les parents qui croient en la médecine basée sur les preuves sont confrontés au conflit de devoir administrer le traitement secrètement. Les membres de la famille peuvent également interférer avec la posologie prescrite, en fonction de leurs opinions personnelles.

### **L'influence de l'Internet**

Les parents ou les adolescents qui vérifient sur Internet peuvent refuser le traitement. Puisque l'Internet peut de la même façon mal informer mais aussi informer, les patients et leurs familles trouvent souvent plus de désinformation que de guidance. Par exemple, ils peuvent chercher un mot sur Google et lire tout ce qui est affiché alors que souvent les informations trouvées mettent l'accent sur les effets secondaires, ignorent les avantages, et s'opposent aux médicaments pour la maladie mentale. Fournir aux parents et aux adolescents des informations et des adresses des sites Internet avec moins de biais et des renseignements plus précis et complets peut améliorer l'adhésion au traitement.

En résumé, un rôle clé du clinicien est de comprendre à la fois le système culturel des croyances et des valeurs de ses patients ainsi que les idées fondamentales de l'épidémiologie et de la recherche clinique qui constituent la première partie de ce chapitre. La combinaison réussie des deux, telle qu'appliquée au patient et à la famille dans le bureau médical, est l'essence de la pratique fondée sur les preuves.

## REFERENCES

- Biederman J, Melmed R, Patel A et al (2008). A randomized, double-blind, placebo-controlled study of guanfacine extended-release in children and adolescents with attention-deficit/hyperactivity disorder. *Pediatrics* 121: e73-84
- Brent D, Perper J, Moritz G et al (1993). Firearms and adolescent suicide. A community case-control study. *American Journal of Diseases of Children*, 147:1066-1071.
- Guyatt GH, Rennie D (2002). *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice* Chicago: AMA Press
- Kabai P (2011b). *Bias and Random Error*. [http://www.behav.org/ecol/wildlife/w\\_05\\_design.htm](http://www.behav.org/ecol/wildlife/w_05_design.htm)
- Kabai P (2011a). *More on Bias (systematic) and Random Error*. [http://apollo.lsc.vsc.edu/classes/remote/lecture\\_notes/measurements/bias\\_random\\_errors.html](http://apollo.lsc.vsc.edu/classes/remote/lecture_notes/measurements/bias_random_errors.html)
- Kaufman J, Birmaher B, Brent D et al (1997). *Kiddie-SADS-Present and Lifetime Version (K-SADS-PL)*. Pittsburgh: Department of Psychiatry, University of Pittsburgh School of Medicine.
- Kelly R, Cohen L, Semple R (2006). Relationship between drug company funding and outcomes of clinical psychiatric research. *Psychological Medicine*, 36:1647-1656.
- McCracken JT, McGough J, Shah B (2002). Risperidone in children with autism and serious behavioral problems. *New England Journal of Medicine*, 347:314-321.
- Sackett DL, Straus SE, Richardson WS et al (2000). *Evidence-Based Medicine: How to Practice and Teach EBM*, 2<sup>nd</sup> edition. New York: Churchill Livingstone.
- Strange BC (2008). Once-daily treatment of ADHD with guanfacine: patient implications. *Journal of Neuropsychiatric Disease and Treatment*, 4:499-506.
- Walkup JT, Albano AM, Piacentini J et al (2008). Cognitive behavioral therapy, sertraline, or a combination in childhood anxiety. *New England Journal of Medicine*, 359:2753-2766
- Wikipedia (2011). *Standard Deviation*. [http://en.wikipedia.org/wiki/Standard\\_deviation](http://en.wikipedia.org/wiki/Standard_deviation)
-