

Introduction à la médecine factuelle

An Introduction to Evidence-Based Medicine

Cucherat M (*)

Résumé

La médecine factuelle (« evidence-based medicine ») est une approche fondée sur la prise en compte systématique des données établies et étayées par les essais cliniques lors des décisions thérapeutiques. Les traitements doivent être évalués pour pouvoir fonder les pratiques médicales, non pas sur des arguments théoriques, mais sur des faits vérifiés montrant le bénéfice à attendre des traitements dans la pratique de tous les jours. La preuve de l'efficacité d'un traitement est la démonstration directe, « grandeur nature », qu'il permet bien d'obtenir l'objectif thérapeutique pour lequel il est prescrit. Les preuves sont produites par les essais thérapeutiques contrôlés randomisés correctement conçus et réalisés et utilisant un critère clinique. Les résultats de plusieurs essais cliniques répondant à une même question thérapeutique sont synthétisés par une méta-analyse qui est une synthèse quantifiée, exhaustive et rigoureuse.

Rev Med Ass Maladie 2001;32,1:41-49

Mots clés : médecine factuelle, niveau de preuve, méta-analyse.

Summary

Evidence-based medicine is an approach based on the systematic consideration of well-established data supported by clinical trails for making therapeutic decisions. In order to determine the optimum medical management, decisions concerning treatment should not be based on theoretical grounds but rather on verified facts which clearly demonstrate the benefit patients can hope to receive in normal daily practice. Proof of the efficacy of a given treatment is constituted by the direct, full-scale demonstration that it produces the therapeutic objective for which it is prescribed. This proof is obtained from the results of controlled, randomized clinical therapeutic trials which have been correctly conceived using a clinical criteria. A meta-analysis is a quantified, exhaustive and rigorous synthesis of the results emanating from a number of clinical trials which attempt to answer the same therapeutic question.

Rev Med Ass Maladie 2001;32,1:41-49

Key words: evidence-based medicine, meta-analysis.

(*) Maître de conférences des universités, praticien hospitalier, Service de biostatistiques, Hôpitaux de Lyon.

Adresse pour correspondance : Service de pharmacologie clinique - EA643, Faculté RTH Laënnec, B.P. 8071, 69372 Lyon cedex 08, e-mail : mcu@upcl.univ-lyon1.fr

Les résultats des essais thérapeutiques prennent une place de plus en plus importante dans la médecine moderne. Ils offrent l'avantage de répondre de façon fiable aux questions thérapeutiques se posant dans la prise en charge des malades. Ils permettent de valider les soins proposés aux patients et donnent les garanties de leur efficacité.

Les traitements doivent être évalués pour pouvoir fonder les pratiques médicales, non pas sur des arguments théoriques, mais sur des faits vérifiés montrant le bénéfice à attendre des traitements dans la pratique de tous les jours. La preuve de l'efficacité d'un traitement est la démonstration directe, « grandeur nature », qu'il permet bien d'obtenir l'objectif thérapeutique pour lequel il est prescrit. Les preuves sont produites par les essais thérapeutiques contrôlés randomisés correctement conçus et réalisés et utilisant un critère clinique.

I. OBJECTIF THÉRAPEUTIQUE

Quel est l'objectif de la prescription d'un traitement par un médecin ?

Un traitement est prescrit par un médecin dans le but d'atteindre un certain objectif thérapeutique, dépendant de la maladie et des attentes du patient vis-à-vis de sa prise en charge thérapeutique.

Dans presque tous les cas, l'objectif thérapeutique a trait soit à une amélioration de la **qualité de vie** du patient soit à une augmentation de **sa durée de survie** [1]. Ces objectifs concernant une amélioration soit qualitative, soit quantitative de la survie du patient sont appelés objectifs cliniques.

Dans les pathologies bénignes, les objectifs recherchés pourront être, par exemple, une diminution de l'intensité des douleurs, le raccourcissement de la durée des symptômes, l'obtention d'une guérison accélérée, etc. Dans les pathologies plus graves, l'objectif est de prolonger la survie du patient, de diminuer son risque de complication ou de survenue d'événements cliniques intercurrents. Dans certains cas, ces deux types d'objectifs cohabitent : augmenter la durée et la qualité de la survie.

Avec les pathologies graves, l'objectif idéal serait d'éviter le décès prématuré ou d'obtenir la guérison de la maladie. Malheureusement, ces objectifs sont le plus souvent hors d'atteinte avec les traitements disponibles. Les objectifs deviennent alors *fréquentistes* : **diminuer le risque** de décès prématuré (avec comme espoir de le rendre nul un jour), **diminuer la fréquence des complications** à moyen terme, c'est-à-dire retarder leur survenue.

Dans l'hypertension artérielle, le traitement est justifié par le fait que l'élévation de la pression artérielle augmente le risque de complications cardio-vasculaires. Dans ce cas, l'objectif thérapeutique est de réduire le risque de survenue de ces complications

Tableau I
Exemples d'objectifs thérapeutiques

Amélioration de la qualité de vie
– Diminuer la durée d'un épisode de rhume
– Atténuer l'intensité de la fièvre dans la grippe
– Diminuer la fréquence des épisodes douloureux dans l'angor stable
Augmentation de la survie
– Réduire le risque de décès à la phase aiguë de l'infarctus du myocarde
– Diminuer le risque de récidives après un premier infarctus.

Tableau II
Objectifs qui ne sont pas des objectifs thérapeutiques mais seulement des « objectifs opératoires »

Baisse de la pression artérielle dans l'hypertension artérielle
Augmenter la densité osseuse dans l'ostéoporose
Restauration du taux de CD4 dans le SIDA
Restaurer la perméabilité coronarienne à la phase aiguë de l'infarctus du myocarde

potentiellement mortelles : infarctus, accidents vasculaires cérébraux et, sur un autre plan, d'atténuer les symptômes : céphalées, troubles de la vision.

L'élévation de la pression artérielle est avant tout un marqueur de l'état de risque accru. La baisse provoquée de la pression artérielle est – peut être – l'intermédiaire par lequel la réduction de risque sera obtenue.

La baisse de pression artérielle n'est pas la finalité du traitement, mais le moyen par lequel on cherche à obtenir le but recherché : la diminution du risque d'événements cardio-vasculaires.

Le tableau II liste quelques exemples d'objectifs qui ne sont pas des objectifs thérapeutiques mais seulement des « objectifs opératoires ». C'est-à-dire des actions par lesquelles on pense pouvoir atteindre l'objectif thérapeutique.

Ces objectifs sont du domaine de la recherche biomédicale et non pas de la pratique thérapeutique. Le pharmacologue a pour objectif de mettre au point des molécules qui baissent la pression artérielle dans l'espoir de donner au clinicien un traitement qui permet de réduire la fréquence des complications cliniques de l'hypertension.

II. PREUVES D'EFFICACITÉ D'UN TRAITEMENT

L'évolution de la médecine la conduit à chercher dans la science le moteur de son progrès. La médecine est, et doit rester avant tout, une « pratique »,

mais qui va chercher dans l'approche scientifique les fondements de son exercice.

Cette base scientifique permet de rationaliser la pratique thérapeutique pour obtenir une efficacité optimale. De plus, cette approche permet d'éviter des pratiques dont le moteur est autre que la recherche d'un résultat thérapeutique optimum.

Ainsi, en dehors du paradigme scientifique (*), le choix des traitements est potentiellement de l'ordre du rapport de force, de la sociologie, de l'esthétique, de l'idéologie, de la croyance, de l'éso-térisme, etc.

L'inscription de l'activité de soin dans le paradigme scientifique vise à limiter l'influence de ces déterminants arbitraires dans le choix des traitements : c'est la confrontation à la réalité qui corrobore ou infirme les hypothèses thérapeutiques.

Pour s'inscrire dans le paradigme scientifique, l'énoncé qu'un traitement apporte un bénéfice doit donc être basé sur les résultats de la confrontation de cette hypothèse à la réalité par une expérience, et non pas n'être qu'une affirmation issue d'un raisonnement théorique [2].

Les sciences biologiques fondamentales et l'épidémiologie nous permettent de connaître de mieux en mieux les mécanismes des maladies et les mécanismes d'action des traitements. Ces connaissances scientifiques fondamentales permettent d'imaginer qu'un certain traitement apportera un bénéfice clinique dans une certaine maladie. Mais tant que cette hypothèse n'a pas été vérifiée, elle reste théorique. Ce n'est pas parce qu'elle se base sur des faits scientifiques prouvés que la conclusion générique : le traitement apporte un bénéfice clinique, est elle aussi prouvée.

L'évolution actuelle de la médecine vers le paradigme scientifique conduit naturellement à exiger des traitements qu'ils apportent la preuve de leur efficacité clinique. Le principe suivant lequel l'utilisation d'un traitement dans une pathologie doit trouver sa justification au niveau de faits scientifiques est le manifeste de la médecine factuelle (« *evidence based medicine* ») [2].

III. QU'EST-CE QU'UNE PREUVE FIABLE ?

La nécessité de disposer de preuves de l'efficacité clinique d'un traitement étant posée, apparaît alors le problème de définir ce qu'est une preuve fiable du bénéfice clinique apporté par un traitement.

Les arguments avancés comme des preuves doivent être d'un haut niveau de fiabilité car de nombreux facteurs favorisent par essence l'apparition de faux

positifs (*). Les intérêts en jeu font que dès qu'un argument favorable est disponible, même s'il s'agit d'un artefact et donc d'un résultat erroné, cet argument sera retenu et avancé par les promoteurs du traitement. Si aucune précaution n'est prise, des arguments seront disponibles même pour des traitements sans efficacité.

Les intérêts en jeu conduisent à la mise en avant des arguments faussement en faveur de l'efficacité et à la mise sous silence des arguments qui ne sont pas en faveur de l'efficacité même s'ils sont le reflet de la réalité. Devant cette pression favorisant l'émergence des arguments positifs à tort, il est nécessaire de n'accepter que les arguments qui constituent des preuves quasiment irréfutables du bénéfice apporté par un traitement. Ne peuvent donc être retenues que les preuves les plus fiables, celles qui ont le moins de risque d'être erronées (biaisées).

Un argument positif doit toujours être considéré comme douteux car on sait que si un argument positif à tort apparaît par hasard il sera utilisé et mis en avant par des promoteurs du traitement. Dans ce contexte, seul les arguments d'un haut niveau de fiabilité sont probants. Ceux moins robustes et potentiellement sujets aux erreurs (biais) ne sont pas convaincants car ils ne permettent pas d'exclure qu'ils sont positifs à tort.

Tous les types d'arguments candidats à cette place de preuve ne remplissent pas toutes les garanties de fiabilité nécessaires dans ce domaine. Parmi tous les types d'arguments qui peuvent être avancés comme preuve, seulement certains possèdent toutes les caractéristiques garantissant leur fiabilité. Les autres sont sujets aux biais, c'est-à-dire à différents types d'erreur qui leur font courir le risque de ne pas refléter la réelle efficacité clinique d'un traitement.

IV. COMMENT OBTENIR LES PREUVES NÉCESSAIRES À LA VALIDATION D'UN TRAITEMENT

L'impérative nécessité de disposer de preuves fiables de l'efficacité d'un traitement étant posée, apparaît la question de savoir comment les obtenir. Cette quête est semée d'embûches, et il convient de se protéger contre plusieurs phénomènes pouvant conduire à des conclusions erronées, c'est-à-dire à de fausses preuves. En l'absence, ou avec une prise en compte insuffisante de ces pièges, il existe un risque d'avancer des arguments erronés comme preuve de l'existence d'une efficacité qui en réalité n'existe pas.

Ces pièges sont les suivants :

– l'utilisation d'un critère de jugement non clinique ;

(*) Un paradigme est un cadre conceptuel de raisonnement.

(*) Dans ce cadre, les faux positifs sont des arguments en faveur de l'efficacité d'un traitement alors que celui-ci en est dépourvu.

- les limites du raisonnement théorique basées sur la physiopathologie ;
- la variabilité biologique et les risques statistiques ;
- les facteurs de confusion et les biais ;
- la sélection des arguments en fonction des résultats ;
- le biais de publication.

Sans protection contre ces sources potentielles d'erreur, un argument en faveur de l'efficacité ne présente pas un niveau de fiabilité suffisante pour constituer une preuve. Il est susceptible d'être biaisé par de nombreux phénomènes.

Quelles sont donc les qualités que doivent posséder les arguments en faveur de l'efficacité pour être considérées comme des preuves suffisamment fiables ?

A. Limites des raisonnements théoriques

L'utilisation d'un traitement pour une pathologie donnée peut être justifiée de plusieurs manières.

Une première approche, traditionnelle, consiste à justifier le traitement par son mécanisme d'action. À partir de la connaissance de la physiopathologie de la maladie et de celle des actions pharmacologiques du traitement, il est possible d'élaborer un raisonnement théorique mécaniste qui laisse présager d'un bénéfice thérapeutique.

Ce type de raisonnement n'est évidemment pas strictement théorique. Il se fonde sur des résultats expérimentaux documentant chaque étape du raisonnement. C'est la combinaison des arguments qui est spéculative comme l'est, par voie de conséquence, l'effet déduit. À l'expérience, cette approche apparaît d'une fiabilité imparfaite.

Cependant, malgré l'importance de nos connaissances fondamentales, celles-ci restent encore parcellaires. Toutes ces raisons font que les déductions faites à ce niveau ne peuvent garantir l'exactitude du raisonnement et de ses conclusions (le traitement est efficace). Il existe de nombreux exemples où ce type de raisonnement a été pris en défaut et où les prédictions du modèle mécaniste n'ont pas été confirmées dans un essai clinique. Le plus fréquemment un traitement prédit comme efficace s'est avéré sans effet, mais avec parfois, chose plus gênante, la révélation d'un effet délétère.

L'essai clinique est une confrontation à la réalité des hypothèses du raisonnement théorique. Son résultat, lorsque l'essai est correctement conçu et réalisé, permet de mesurer en conditions réelles le bénéfice apporté par le traitement. Il permet de démontrer que l'utilisation d'un traitement s'accompagne bien d'un bénéfice réel.

Par exemple les anti-arythmiques de classe I ont été prescrits après infarctus du myocarde en cas d'extra-systolie ventriculaire pour prévenir la mort subite. La valeur péjorative des extra-systoles ventri-

culaires fréquentes est bien documentée. De même, les anti-arythmiques de classe I ont montré qu'ils diminuaient fortement la fréquence des extra-systoles. De ce fait, il semblait logique de penser que l'utilisation de ces traitements devait diminuer la mortalité par mort subite. Ces traitements furent utilisés en pratique pendant de nombreuses années et uniquement justifiés par ce raisonnement théorique. Ce n'est qu'en 1991 que l'essai CAST [3] a évalué quels étaient les résultats réellement produits sur la mortalité (tableau III).

Les résultats de cet essai furent à l'opposé de ce qui était attendu. Au lieu de confirmer la réduction de mortalité pressentie, ils mettaient en évidence un doublement de celle-ci. On a pu calculer que les prescriptions non fondées de ces médicaments auraient entraîné aux États-Unis au moins autant de décès que les guerres de Corée et du Vietnam.

Une explication fut rapidement avancée faisant intervenir les effets pro-arythmogènes et inotropes négatifs des anti-arythmiques de classe I.

Ces effets pour lesquels il a été possible de proposer un mécanisme physiopathologique n'avaient pas été pris en compte dans l'élaboration du modèle thérapeutique théorique alors qu'ils étaient connus. On touche d'ailleurs là à une limitation des modèles thérapeutiques discursifs : la difficulté de prendre en compte des effets contraires à cause de l'absence d'intégration des aspects quantitatifs de la connaissance.

Malgré ses limites, l'étude des mécanismes d'action est indispensable pour imaginer de nouveaux traitements. Cette approche est indispensable pour rationaliser l'évaluation des traitements. En effet, la connaissance des mécanismes fondamentaux de la physiopathologie et de la pharmacologie est une voie sans égale pour la génération de nouvelles hypothèses thérapeutiques.

Sans connaissance fondamentale, comment trouver de nouveaux antiagrégants plaquettaires ? Comment avoir l'idée d'utiliser des fibrinolytiques dans l'infarctus du myocarde ?

Mais l'approche pharmacologique et physiopathologique ne produit que des hypothèses thérapeutiques qui doivent impérativement être confrontées à la réalité pour produire les preuves définitives de l'intérêt de ces hypothèses. Le but de la démarche

Tableau III
Résultats de l'essai CAST [3] de prévention de la mort subite après infarctus par les anti-arythmiques de classe I (AA1).

	DC / n	mortalité
Groupe AA1	39 / 432	9 %
Groupe contrôle	18 / 423	4 %

RR = 2,13 ; p = 0,0004

est de démontrer qu'un traitement apporte en pratique un bénéfice clinique, pas de montrer qu'il existe un cadre formel de pensée pouvant avancer un mécanisme débouchant sur un bénéfice.

Il faut cependant être conscient que dans de nombreux cas, soit les essais n'ont pas confirmé la théorie pourtant séduisante et paraissant robuste, soit après démonstration clinique de l'efficacité du traitement, un autre mécanisme d'action a été trouvé (par exemple les bêta-bloqueurs dans le post-infarctus du myocarde). Ces deux points illustrent le fait que, malgré le haut degré de connaissances que l'on puisse atteindre sur la physiopathologie des maladies et sur les mécanismes d'action des traitements, la complexité biologique peut mettre en échec une approche purement théorique.

UN AUTRE EXEMPLE : AGENTS INOTROPES INHIBITEURS DE LA PHOPHODIESTÉRIASE

Les essais des agents inotropes inhibiteurs de la phosphodiésterase dans le traitement de l'insuffisance cardiaque donnent deux exemples où, de la même manière qu'avec les antiarythmiques de classe Ic (comme la flécaïne ou la quinidine) en post-infarctus, le raisonnement physiopathologie théorique est pris en défaut et n'est pas confirmé par les essais cliniques.

D'après leurs effets pharmacologiques favorables sur des critères physiopathologiques d'ordre hémodynamique, deux produits, l'énoxinome et la vesnarinone, sont apparus comme des traitements qui devaient augmenter la survie des patients insuffisants cardiaques. Ce raisonnement théorique n'a cependant pas été confirmé lors des essais de mortalité où une surmortalité a été observée [4, 5]. De plus dans cet exemple, des critères assez proches de critères cliniques, comme la tolérance à l'exercice ou la qualité de vie, étaient aussi influencés de façon favorable. La seule prise en considération de ces critères qui sont déjà des critères cliniques n'aurait pas permis de mettre en évidence cette surmortalité.

Cet exemple illustre une fois de plus le danger qu'il pourrait y avoir à se contenter de la démonstration d'un effet sur un critère intermédiaire physiologique, comme la fonction ventriculaire, à la place d'un essai sur le critère clinique ultime. Il montre aussi que la mise en évidence d'un bénéfice sur des critères cliniques intermédiaires, comme la tolérance à l'effort ou la durée de l'effort, n'est pas non plus suffisante.

De plus, ces deux exemples donnent l'occasion de discuter des problèmes d'interprétation de la pertinence clinique d'un effet observé sur une échelle de score et de sa confrontation à un effet sur un critère clinique. La question qui se pose est de savoir si l'amélioration de la qualité de vie ou de la symptomatologie est suffisamment importante pour éventuellement rendre acceptable un surcroît de mortalité. Avant d'envisager le problème éthique d'une

réduction des chances de survie sous prétexte d'une amélioration fonctionnelle, il convient de pouvoir confronter la pertinence clinique des tailles des effets obtenus respectivement sur la mortalité et sur les signes fonctionnels. Ce n'est pas parce qu'il y a détection d'un effet statistiquement significatif sur la qualité de vie(*), que celui-ci est notable et intéressant pour les patients et suffisamment important pour constituer une amélioration substantielle pouvant éventuellement justifier l'acceptation d'une surmortalité. En d'autres termes, la surmortalité observée représente-t-elle un coût acceptable en regard du bénéfice obtenu sur l'amélioration des symptômes ? Il est crucial dans cette situation de pouvoir traduire en terme clinique (évaluer la pertinence clinique) la différence de score de qualité de vie en des termes qui la rendent comparable à la surmortalité.

(*) Surtout qu'avec les critères de jugement continu, des effets de petite taille ne peuvent s'avérer très facilement statistiquement significatifs.

B. La nature probabiliste des phénomènes étudiés comme limite de l'expérience personnelle

L'expérience clinique possède des limites. Dans bon nombre de pathologies, la faible fréquence de survenue des événements et la petitesse des bénéfices attendus font qu'il est impossible de juger à partir de quelques cas de l'efficacité d'un traitement.

Tout médecin a déjà fait la constatation que des patients ayant des caractéristiques similaires peuvent avoir des évolutions très différentes. Par exemple, deux patients similaires avec une hypertension artérielle auront des devenir différents : l'un faisant un accident vasculaire cérébral (AVC) très rapidement, l'autre pas, même après plusieurs années.

Cette variabilité peu réductible des phénomènes biologiques ne permet que de raisonner en probabilité et non plus dans un champ déterministe. Une élévation de la pression artérielle ne détermine pas la survenue d'un AVC, elle en augmente seulement la probabilité (variabilité pour les critères continus).

La nécessité de raisonner en probabilité empêche de pouvoir conclure à partir de l'observation d'un seul individu. Il convient de travailler sur des groupes de patients pour pouvoir mesurer les probabilités et les variations de probabilités avec suffisamment de précision. Chez un hypertendu, la probabilité annuelle d'accidents cardiovasculaires est de 4 %. Il est évident que l'observation d'un seul sujet qui présentera ou pas un événement ne permet pas de savoir si le traitement a modifié ce risque.

PROBABILITÉ NOMBRE ATTENDU

Même sans aucune variation de la probabilité de survenue d'un événement, le nombre d'occurrences de cet événement observé dans différents groupes de même effectif fluctue. Par exemple, avec une fréquence de 10 %, il est possible d'observer, dans un groupe donné, aussi bien zéro événement, que 10 ou 20. Aucune valeur (entre zéro et le nombre de sujets du groupe) n'est impossible, elles sont seulement plus ou moins probables. Ainsi il est du domaine du possible que sur le même nombre de patients ayant la même probabilité d'événement, un médecin n'observe aucun événement et qu'un autre en note 10, cette différence n'étant due qu'au hasard.

La plupart des traitements ne font que modifier la probabilité d'événement sans la supprimer. La mise en évidence de leur efficacité ne peut donc se faire qu'à partir de groupes de patients. L'expérience empirique que peut avoir un praticien sur une série limitée de ses patients ne lui permettra pas d'appréhender des modifications de probabilité.

L'observation de l'évolution satisfaisante de quelques patients ne permet pas de démontrer une efficacité. En l'absence de traitement, ces évolutions favorables sont aussi observées. De plus, l'esprit humain a tendance à oublier les mauvaises expériences et à ne mettre en avant que les bonnes. Tous ces phénomènes concourent à fausser favorablement nos impressions subjectives sur l'efficacité issue de l'expérience personnelle.

Les impressions issues de l'expérience personnelle ne peuvent pas être acceptées comme preuve de l'efficacité, même si elles ont de réelles conséquences psychologiques sur les prescripteurs. La réticence que peut avoir un médecin à utiliser un traitement après avoir fait l'expérience d'un événement indésirable gravissime est bien compréhensible.

La justification d'après l'expérience clinique possède des limites. Dans bon nombre de pathologies, la faible fréquence de survenue des événements et la petitesse des bénéfices attendus font qu'il est impossible de juger à partir de quelques cas de l'efficacité d'un traitement.

Bien entendu, il existe en médecine des situations très fortement déterministes où l'outil probabiliste n'est pas nécessaire. Montrer que le sondage vésical soulage sur le champ le patient en rétention aiguë ne nécessite pas de faire appel au raisonnement en terme de probabilité. Il en est de même pour l'effet narcotique d'un produit anesthésique. Par contre, l'évaluation de la sécurité de ces deux interventions retombe dans le champ du probable et nécessite l'outil statistique.

C. Les études d'observation n'empêchent pas les biais

Plusieurs types d'études de recherche clinique peuvent être envisagés *a priori* pour permettre une confrontation à la réalité des hypothèses thérapeutiques. Il s'avère cependant qu'elles ne sont pas toutes équivalentes et qu'elles ne permettent pas toutes d'obtenir des preuves fiables.

Quels sont donc les types d'études qui permettent une confrontation à la réalité suffisamment fiable des hypothèses thérapeutiques ?

La simple observation n'est pas suffisante car l'observation ne permet pas de prendre en compte les facteurs de confusion et ses résultats sont potentiellement biaisés. Ainsi les études d'observation épidémiologiques ne donnent pas suffisamment de garanties pour fournir les preuves recherchées. Par contre, elles sont idéales pour générer de nouvelles hypothèses. Pour la recherche d'une efficacité thérapeutique les limites des études d'observation sont listées dans le tableau IV.

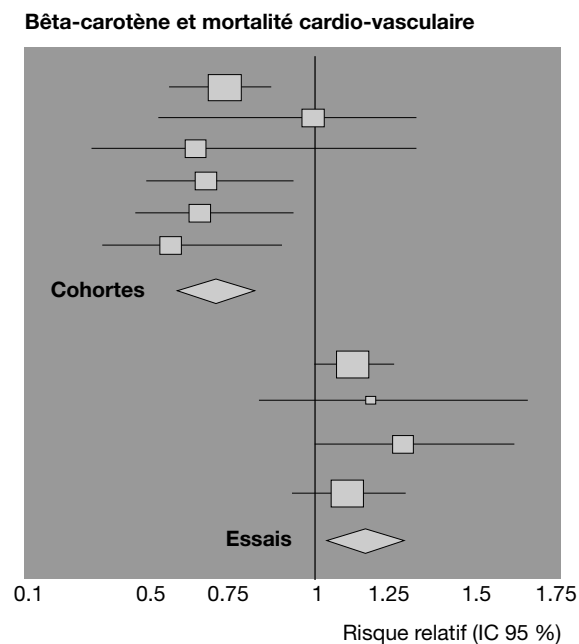
Exemple

Un exemple des limitations des études d'observation pour la recherche des effets d'un traitement est donné par l'évaluation des effets du bêta-carotène en prévention. Les données épidémiologiques [6] et biologiques suggèrent que le bêta-carotène, grâce à ses propriétés antioxydantes, serait protecteur contre les cancers et les maladies cardiovasculaires.

Les études de cohortes aboutissent au résultat que les sujets consommant le plus de bêta-carotène auraient une mortalité cardio-vasculaire (CV) réduite par rapport à ceux qui en consomment le moins.

Tableau IV
Limites des études d'observation
(recherche d'une efficacité thérapeutique)

Type d'étude	Limites
Série de cas	Pas de prise en considération des facteurs de confusion
Étude écologique	Pas de prise en compte des différences géographiques (génomiques, environnementales, etc.)
Étude longitudinale	Pas de prise en compte de l'évolution séculaire de la maladie, de l'évolution de ses déterminants, de sa prise en charge, de sa prévention
Étude cas-témoin	Biais d'indication, biais de mémorisation : les sujets atteints se souviennent plus facilement des traitements qu'ils ont pris
Étude de cohortes	Biais d'indication : les patients reçoivent ou ne reçoivent pas le traitement étudié en fonction de la gravité de leur maladie



L'hypothèse issue des études d'observation que le bêta-carotène pouvait réduire la mortalité CV a été testée dans quatre essais thérapeutiques [7-9]. Leurs résultats vont à l'encontre de ceux des études épidémiologiques et concluent à une augmentation de mortalité.

Plusieurs hypothèses sont disponibles pour expliquer ces discordances : types de sujets différents, doses différentes, mais surtout, et avant tout, le fait que l'effet observé dans les études épidémiologiques peut être dû à un facteur de confusion non identifié.

Par exemple, il est possible que la consommation de forte quantité de bêta-carotène soit simplement un marqueur d'une plus forte appétence pour la santé. Les personnes qui apportent le plus de soins à leur santé font davantage attention à leur alimentation et ont une alimentation plus diversifiée et plus riche en vitamines. De ce fait, les sujets consommant le plus de bêta-carotène pourraient avoir moins de conduite à risque (tabagisme, alcoolisme), se soumettraient plus à des actions de prévention, auraient une plus grande facilité de recours aux soins, etc., et, finalement, auraient donc un meilleur pronostic que ceux consommant moins de bêta-carotène car prêtant moins d'attention à leur santé.

Même si en fait le bêta-carotène augmente quelque peu leur mortalité, celle-ci se maintient inférieure à celle des sujets contrôles. Dans les essais randomisés, où tous ces facteurs comportementaux et liés à la catégorie socioprofessionnelle sont également répartis dans les deux groupes, le vrai effet du bêta-carotène peut apparaître. Cette explication n'a pas la prétention d'être « l'Explication » de ces discordances observées entre études d'observation et études randomisées. Elle n'est qu'une possibilité parmi beaucoup d'autres. Elle

a comme seul objectif de montrer comment une étude d'observation peut être faussée par un facteur de confusion. Les conséquences des facteurs de confusion connus peuvent être corrigées par un ajustement, mais pas ceux des facteurs de confusion non connus ou non enregistrés. Les études d'observation ne permettent donc pas de garantir l'absence de biais.

Si les prescriptions avaient suivi les résultats des études d'observation sans attendre les résultats des essais, la pratique qui en aurait découlé aurait eu un résultat inverse à celui recherché, en induisant des décès.

D. Nécessité de confirmation des résultats

Le résultat d'un seul essai répondant aux critères que nous venons de voir ne constitue cependant pas encore une preuve suffisamment fiable. En effet, malgré une méthodologie irréprochable, le résultat d'un essai unique peut encore être inexact du fait de la présence d'une erreur statistique ou d'un artefact. La nature aléatoire implique un risque d'erreur dans la conclusion sur l'existence de l'effet du traitement qui ne peut pas être éliminé, même s'il est contrôlable grâce au test statistique. Un résultat statistiquement significatif au seuil $\alpha = 5\%$ laisse la possibilité d'une conclusion fautive dans 5% des cas.

De plus, le résultat d'un essai peut être dû à un artefact. Même avec une méthodologie irréprochable, il n'est pas possible d'exclure avec une certitude de 100% l'existence d'un biais ou l'existence de circonstances exceptionnelles conduisant à un résultat positif à tort.

Seule la vérification du résultat par au moins un autre essai et sa confirmation permet d'éliminer ces deux possibilités. La nécessité de vérifier les résultats va entraîner dans de nombreux cas une multiplication des résultats à prendre en compte avant de se fixer sur l'efficacité du traitement. Cette multiplicité de l'information entraîne à son tour de nouvelles difficultés : celle de faire une synthèse de résultats qui peuvent être apparemment discordants car soumis aux risques d'erreurs statistiques alpha et bêta (*) et le risque d'une sélection arbitraire en fonction de leurs résultats des essais présentés pour argumenter l'effet du traitement. Ces difficultés sont solutionnées par la méta-analyse [10].

La multiplicité des essais disponibles dans un domaine s'accompagne souvent de résultats apparemment discordants, et tirer une conclusion globale par la seule logique discursive n'est pas toujours possible.

(*) Risque alpha (α) de première espèce : risque de mettre en évidence un effet du traitement qui n'existe pas en réalité. Risque bêta (β) de seconde espèce : risque de ne pas mettre en évidence un effet du traitement qui existe en réalité.

En général, des essais concluants (donnant un résultat statistiquement significatif en faveur de l'efficacité) coexistent avec des résultats non concluants. Ces discordances apparentes donnent la possibilité de soutenir les deux conclusions opposées : celle de l'existence de l'efficacité et celle de son absence.

Les tenants de l'existence de l'efficacité argumenteront à partir des résultats concluants, les résultats non significatifs étant expliqués par un manque de puissance.

Les défenseurs de l'absence de preuves d'efficacité mettront en avant les résultats non significatifs et expliqueront les résultats positifs par le fait du hasard (risque de première espèce alpha).

La solution à ce problème est apportée par la méta-analyse qui est une technique de synthèse quantitative.

E. Sélection arbitraire

Un autre écueil à éviter dans la recherche des preuves est celui de la sélection arbitraire des essais en fonction de leurs résultats.

Assez fréquemment, les synthèses réalisées sous forme de revue générale ne retiennent que les résultats positifs et ne font pas mention de ceux qui ont été négatifs, donnant ainsi une impression plus favorable en faveur de l'efficacité que ce qu'elle aurait dû être après prise en compte de tous les résultats.

Pour éviter ce problème, les preuves d'efficacité doivent être issues d'une synthèse non arbitraire de tous les résultats d'essais, qu'ils soient en faveur ou non de l'efficacité.

Un exemple de sélection arbitraire des résultats couramment avancé pour justifier l'efficacité d'un traitement est donné par le travail de Ranskov réalisé sur les essais d'hypocholestérolémiants dans la prévention du risque cardiovasculaire [11].

En 1992, les résultats de 24 essais étaient disponibles, 14 étaient favorables à l'efficacité en prévention des hypocholestérolémiants et 10 ne l'étaient pas (résultats non statistiquement significatifs ou effet délétère). Ranskov a étudié la fréquence de citations de ces essais dans la littérature au travers des articles de synthèse, des éditoriaux, des articles de recommandations pour la pratique.

	Nombre moyen de citations par an
Résultats favorables (n = 14)	40
Résultats non favorables (n = 10)	7,4
Résultats favorables publiés dans une revue à diffusion large (n = 8)	61

Fréquence de citations dans les années suivant la publication de deux essais publiés dans le JAMA :

Essai	1 ^{re} année	2 ^e	3 ^e	4 ^e
LRC (favorable)	109	121	202	180
Miettinen (non favorable)	6	5	3	0

La fréquence de citation dépend grandement du résultat. Les essais négatifs disparaissent presque totalement de la mémoire collective.

Pour constituer une preuve fiable, une série de résultats d'essais doit être issue d'un processus de collecte non arbitraire, sélectionnant les essais sur leur caractère non biaisé et non sur leurs résultats. Les résultats en faveur du traitement ainsi que ceux en défaveur doivent être pris en considération.

F. Biais de publication

Les essais thérapeutiques ont d'autant plus de chance d'être publiés que leurs résultats s'avèrent statistiquement significatifs [12]. De ce fait la littérature biomédicale ne reflète pas la réalité, mais en donne un aperçu plutôt optimiste en ne publiant pas les essais en faveur de l'absence d'effet du traitement. Cette situation est susceptible de biaiser le résultat d'une synthèse de l'information disponible (biais de publication). Toute synthèse doit donc se baser sur une recherche exhaustive de tous les essais réalisés, publiés et non publiés.

V. CONCLUSION

Les différents points que nous venons de voir conduisent à conclure que l'obtention de preuves fiables de l'efficacité d'un traitement nécessite :

- une vérification directe, sur **critères cliniques**, que le traitement permet d'atteindre l'objectif thérapeutique pour lequel il est pressenti,
- que cette vérification s'effectue sans biais par le moyen **d'essais thérapeutiques randomisés correctement conçus et réalisés**,
- que **plusieurs résultats concordants** soient disponibles pour éliminer un artefact dû, soit au risque d'erreur statistique alpha, soit à une étude biaisée,
- que l'ensemble des essais conduits, quels que soient leurs résultats, soit disponible afin de pouvoir en faire une **synthèse** loyale pesant les résultats positifs et négatifs,
- que la synthèse des essais concernant un même traitement soit effectuée en utilisant la technique de la **méta-analyse**, capable de prendre en compte, d'une part, l'inflation du risque alpha lié au nombre des études et, d'autre part, le risque bêta.

RÉFÉRENCES

1. Boissel JP. *L'information thérapeutique*. Paris : Masson; 2000.
2. Evidence-based medicine working group. *Evidence-based medicine. A new approach to teaching the practice of medicine*. JAMA 1992;268:2420-5.
3. Echt DS, Liebson PR, Mitchell LB, Peters RW. *Mortality and morbidity in patients receiving encainide, flecainide or placebo. The Cardiac Arrhythmia Suppression Trial*. N Engl J Med 1991;324:781-8.
4. Cohn JN, Goldstein SO, Greenberg BH et al. *A dose-dependant increase in mortality with vesnarinone among patient with severe heart failure. Vesnarinone Trial Investigators*. N Engl J Med 1998;339:1810-6.
5. Cowley AJ, Skene AM. *Treatment of severe heart failure: quantity or quality of life? A trial of enoximone. Enoximone Investigators*. Br Heart J 1994;72:226-30.
6. Jha P, Flather M, Lonn E, Farkouh M, Yusuf S. *The antioxidant vitamins and cardiovascular disease*. Ann Intern Med 1995;123:860-72.
7. Alpha-Tocopherol Beta Carotene Cancer Prevention Study Group. *The effect of vitamin E and beta carotene on the incidence of lung cancer and other cancers in male smokers*. N Engl J Med 1994;330:1029-35.
8. Hennekens CH, Buring JE, Manson JC. et al. *Lack of effect of long-term supplementation with beta carotene on the incidence of malignant neoplasms and cardiovascular disease*. N Engl J Med 1996;334:1145-9.
9. Omenn GS, Goodman GE, Thornquist MD et al. *Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease*. N Engl J Med 1996;334:1150-5.
10. Cucherat M, Boissel JP, Leizorovicz A. *La méta-analyse des essais thérapeutiques*. Paris : Masson; 1997.
11. Ravnskov U. *Cholesterol lowering trials in coronary heart disease: frequency of citation and outcome*. BMJ 1992;305:15-9.
12. Dickersin K. *The existence of publication bias and risk factors for its occurrence*. JAMA 1990;263:1385-9.