

Introduction à la lecture critique d'articles médicaux et scientifiques

Marius Fieschi, Roch Giorgi, Joanny Gouvernet, Gilles Bouvenot

1. Introduction

Le médecin a, vis-à-vis de ses patients, une obligation de moyens et ne peut recourir qu'à des traitements éprouvés. Par ailleurs, l'importante et rapide diffusion des connaissances scientifiques et médicales ne permet plus d'admettre que, face à des patients présentant des cas semblables, les pratiques diagnostiques et thérapeutiques des praticiens puissent varier de façon notable.

Considérons deux patients identiques en tous points (de même âge, atteints de la même maladie grave au même stade évolutif...) se présentant dans deux hôpitaux différents et se voyant proposer deux traitements « lourds » différents : il est vraisemblable que l'un des deux traitements est meilleur que l'autre pour le cas présenté. Écartons l'erreur ou la différence d'appréciation de la situation par l'un des praticiens. Sur quels critères ces praticiens se fondent-ils pour proposer des conduites thérapeutiques aussi différentes, après évaluation de leurs conséquences pronostiques ? Peut-on penser que l'évocation de leur expérience personnelle - très probablement différente - suffise à justifier à elle seule deux attitudes différentes avec leurs conséquences prévisibles, par exemple en termes de survie ? Peut-on demander à chacun d'eux d'expliquer son choix et de l'argumenter sur la base de faits aussi indiscutables que possible ? Si une étude, publiée dans un journal scientifique de haut niveau, corrobore l'option thérapeutique de l'un des deux, renforce-t-elle la crédibilité de sa proposition ?

Prendre une décision pour un patient donné est l'aboutissement d'un processus où la rigueur intellectuelle et l'éthique ont des places d'honneur. Or, rigueur intellectuelle et éthique exigent une perpétuelle mise à jour des connaissances établies et des savoirs utiles. La pratique médicale classique est donc confrontée à la difficulté de maîtriser la connaissance produite par la recherche clinique et la masse de plus en plus importante des informations scientifiques diffusées. Le problème de la crédibilité et de la validité de ces informations se pose. Ainsi, du fait de la rapidité de l'évolution des connaissances et de leur accessibilité grâce aux moyens offerts par les technologies de l'information et de la communication, analyser, critiquer et synthétiser les données nouvelles de la littérature scientifique devient,

pour tout médecin, un exercice primordial. Cet exercice vise d'abord à évaluer la crédibilité d'une information : plus forte est la crédibilité d'une information, plus cette information change de statut pour devenir *in fine* un « fait établi » par la science ; mais cet exercice vise aussi à évaluer sa pertinence et son utilité pour la pratique.

Archie Cochrane a porté, dans les années 70, un jugement sur la profession médicale en affirmant : « C'est sûrement une critique importante de notre profession que nous n'ayons pas prévu de faire des résumés critiques, par spécialité ou sous-spécialité, avec une mise à jour périodique de tous les essais pertinents ». C'est sur ce constat que s'est élaboré le « paradigme » de la médecine basée sur le niveau de preuve (Evidence-based medicine ou EBM des auteurs anglosaxons) qui recommande de nouveaux comportements de la part du médecin et définit les règles d'élaboration de la connaissance, impliquant une plus grande utilisation des données de la littérature pour guider la pratique.

Dans la démarche décisionnelle, le médecin doit minimiser l'importance de l'intuition et de l'expérience clinique non systématique et non contrôlée, et prendre en considération les résultats de la littérature considérés comme établis afin d'exercer une « médecine dont les pratiques sont directement inspirées de faits scientifiques validés, d'interprétation indiscutable » [1].

La littérature médicale scientifique, très abondante, rapporte de très nombreuses observations, estime de nombreux paramètres biologiques (dans certaines conditions, de manières différentes, avec une précision différente), établit de nombreux faits, ... Quelle crédibilité accorder à toutes ces informations ? Peuvent-elles toutes être considérées comme des faits scientifiquement établis ? Ces questions renvoient *ipso facto* à d'autres questions : comment lire un article médical ? Sous quel angle aborder la critique scientifique du travail rapporté par ce type de document ? Outre la crédibilité du résultat, quel est son intérêt en clinique, quelles sont son utilité pratique et son applicabilité ? Autrement dit, il s'agit de développer une approche centrée sur le malade et fondée sur le niveau de preuve d'une information médicale. La lecture critique est un processus d'évaluation du résultat publié et de son interprétation en considérant systématiquement sa validité (c'est-à-dire les conditions dans lesquelles il est vrai) et donc sa pertinence pour la pratique.

Les concepts de base de la méthodologie de l'épidémiologie, des essais thérapeutiques contrôlés et des tests diagnostiques définissent des standards de rigueur en recherche clinique. Ils fournissent, avec les biostatistiques, un socle de référence pour apprécier la crédibilité d'un résultat, pour définir des *check-lists* et structurer la démarche de la lecture critique et la rendre plus explicite. La connaissance de ces concepts de base est donc un préalable à l'appréciation

critique de la lecture médicale.

2. La crédibilité du fait établi, la notion de niveau de preuve

L'analyse statistique constitue, habituellement, une part importante d'une étude clinique, mais ne suffit pas à apprécier l'intérêt du résultat obtenu. Un bon jugement clinique et scientifique est également nécessaire, car la crédibilité d'un fait observé est liée aux conditions dans lesquelles l'étude a été réalisée.

La crédibilité du résultat d'une étude clinique dépend à la fois de la précision de sa description, de sa mesure, mais aussi de ses conditions d'observation. Exemple : un article qui se propose d'établir l'efficacité d'un nouveau médicament doit définir le type de patients sur lesquels l'étude a été réalisée, quels ont été très précisément la thérapeutique administrée, le critère d'évaluation de l'effet du médicament, dans quelles conditions ce critère a été mesuré et à quelle thérapeutique existante le nouveau traitement a été éventuellement comparé.

La valeur de l'information médicale obtenue est alors appréciée non seulement par l'adéquation de la méthodologie mise en oeuvre pour produire des connaissances fiables et non biaisées, mais aussi par la connaissance du contexte médical de sa réalisation.

Il est important de juger de la crédibilité du résultat proposé par l'auteur de l'article et de se faire une opinion argumentée de l'intérêt et des limites de l'étude que l'on analyse. Afin d'aider les lecteurs à qualifier la crédibilité des études, on utilise les niveaux de preuves proposés par la médecine basée sur le niveau de preuve, rapportés dans le tableau 1.

Il convient toutefois de noter que le niveau de preuve n'est pas le seul critère de jugement de l'intérêt médical pratique d'un article. Ainsi, si les articles qui produisent un très haut niveau de preuve sont toujours pris en considération, certaines études de niveau de preuve moins élevé peuvent présenter un certain degré d'intérêt. Il est utile, à ce propos, de rappeler les conséquences pratiques de la publication de la lettre à la rédaction de WG. McBride de 1961 dans laquelle il faisait part de son observation concernant l'augmentation des cas de phocomélie chez les nouveaux nés de femmes prenant du thalidomide pendant leur grossesse par rapport aux femmes n'en prenant pas¹. Par ailleurs certains articles ne peuvent, par définition, prétendre au niveau de preuve le plus élevé du fait même de l'objet de leur étude. On conçoit aisément par exemple que, pour des raisons éthiques, il ne soit pas possible de réaliser un essai contrôlé randomisé pour étudier le lien de causalité entre consommation de tabac et survenue du cancer broncho-pulmonaire². On ne trouvera alors que des études de

¹ McBride WG. Thalidomide and congenital abnormalities. *Lancet* 1961;ii:1358.

² Cela nécessiterait de constituer aléatoirement un groupe d'individus consommant obligatoirement du tabac et

niveau de preuve maximal égal à 3.

Tableau 1. Les cinq niveaux de preuve de l'EBM.

Niveau 1.	Information fournie par des essais contrôlés randomisés avec des effectifs suffisants et des résultats méthodologiquement indiscutables
Niveau 2.	Information fournie par des essais contrôlés non randomisés ou avec de petits effectifs ou des résultats méthodologiquement discutables
Niveau 3.	Information fournie par une étude prospective, par des essais ouverts sans témoins (non contrôlés) non randomisés
Niveau 4.	Information fournie par une série de cas, par une étude rétrospective cas-témoin comportant des biais
Niveau 5.	Opinion d'expert, expérience personnelle.

3. Le rapport de lecture

La lecture critique vise à établir un rapport de lecture qui doit :

- Identifier la question posée et la méthodologie adoptée pour y répondre ;
- Apprécier l'adéquation du plan expérimental de l'étude au but poursuivi ;
- Apprécier la qualité de la mise en œuvre de l'étude pour apporter le niveau de preuve suffisant ;
- Identifier les biais importants ; apprécier leur impact sur la fiabilité des résultats ;
- Vérifier si la méthode statistique est adaptée et correctement mise en oeuvre
- Apprécier l'intérêt clinique pratique de l'étude, qui doit être analysé et qualifié.

3.1. La forme de l'article (c'est-à-dire le plan de rédaction)

La structuration d'un article scientifique répond à des standards dont la connaissance permet aussi de répondre aux questions qu'il convient de se poser pour le caractériser et l'évaluer. Autrement dit, sachant comment l'auteur a structuré son article, le lecteur peut se reporter à une partie bien précise de cet article pour trouver les réponses aux questions qu'il se pose relativement à la qualité du travail présenté. Deux standards sont aujourd'hui proposés : le standard IMRD (Introduction, Méthodes, Résultats, Discussion) et le standard IHMRD (Introduction, Historique, Méthodes, Résultats, Discussion) proposé par l'AFNOR (Association française de normalisation). La structuration la plus couramment rencontrée avec

un autre groupe d'individus n'en consommant pas puis de comparer le nombre de cancers broncho-pulmonaires apparus au cours du temps dans chacun des deux groupes.

le standard IMRD est détaillée dans le tableau 2.

Tableau 2: Structure IMRD d'un article médical scientifique.

Introduction	Fixe brièvement le contexte général du problème étudié et présente une synthèse de la connaissance utile pour comprendre l'intérêt et l'importance de la question posée par les auteurs. Le dernier paragraphe de l'introduction formule en général l'objectif de l'étude.
Matériels et Méthodes	Décrit précisément ce qui a été fait pour répondre à la question énoncée dans l'introduction : plan expérimental (schéma de l'étude), sélection des patients, répartition des malades dans les groupes étudiés (randomisation, ...), variables observées et relevées, modalités de mesure des variables, ... Permet de savoir si l'étude a été conduite avec une méthodologie appropriée et reproductible.
Résultats	Commence en général par une description des échantillons étudiés, qui permettra de savoir à quels malades on pourra appliquer le résultat obtenu. Donne ensuite les résultats en rapport avec l'objectif principal de l'étude, puis des résultats en rapport avec d'éventuels objectifs secondaires.
Discussion	Permet aux auteurs de proposer une interprétation de leurs résultats, de la validité de ces résultats, de leur caractère général, de leur importance théorique et pratique, en les confrontant éventuellement aux données de la littérature.

3.2. La question posée par les auteurs

Une question bien formulée doit permettre d'identifier précisément le type de sujets qu'elle concerne ainsi que le type d'intervention (ou d'exposition) et les critères de l'évaluation faisant l'objet de l'étude. Par exemple, il peut s'agir de déterminer si l'utilisation de fibrates (*intervention étudiée*) réduit le risque d'accidents cardiaques coronariens (*critère d'évaluation*) chez des hommes de moins de 70 ans ayant un taux de HDL-Cholestérol ≤ 1 mmol/l (*caractérisation des sujets*).

3.3. Qualification des études

Qualifier le type de l'étude permet de mieux apprécier s'il est le mieux adapté pour répondre à la question posée. On distingue deux grands types d'études : les études expérimentales et les études d'observation.

L'information fournie par une étude expérimentale est généralement considérée comme plus fiable que celle fournie par une étude d'observation. Par construction, une étude expérimentale est prospective et basée sur un protocole préétabli qui ajoute à la qualité du travail réalisé (on sait exactement quelles seront les données à recueillir, chez qui, à quel moment, de quelle manière,...). En effet, si les données rétrospectives sont souvent disponibles pour un coût moins élevé, elles sont, en revanche, d'une moins bonne fiabilité :

données non recueillies systématiquement, recueillies suivant un protocole différent selon les patients...

Les chercheurs visent, dans les protocoles des études expérimentales, à minimiser la variabilité des réponses aux facteurs étudiés, ce qui peut les conduire à ne pas inclure certains types de patients. Ainsi, une étude expérimentale restreint généralement, par ses critères d'inclusion et de non inclusion, le champ des situations rencontrées. Il conviendra donc de porter attention à l'intérêt clinique de telles études et à la généralisation de leurs résultats.

3.4. Objectifs des études et choix des plans expérimentaux (schémas)

Certains schémas d'études sont considérés comme meilleurs que d'autres pour répondre à certaines questions. Ainsi les essais cliniques contrôlés sont généralement considérés comme le schéma d'étude obligatoire quand il s'agit d'évaluer l'efficacité d'une thérapeutique. Mais d'autres schémas d'études sont appropriés pour répondre à d'autres objectifs de recherche : des recherches sur l'étiologie ou sur les facteurs de risque peuvent être conduites dans le cadre d'études « cas-témoins » ou d'études de cohorte. Ces différents schémas d'étude sont définis dans d'autres chapitres de cet ouvrage ainsi que dans le glossaire.

3.5. La construction du fait scientifique : impératifs méthodologiques

3.5.1. L'identification des biais

De nombreux biais ont été individualisés. Nous n'en citerons que quelques uns qu'il convient impérativement de débusquer. La Figure 1 présente le schéma général d'une étude de type « essai thérapeutique ». A chaque stade de l'étude, des biais peuvent s'introduire si des précautions méthodologiques ne sont pas prises. Ces biais, ainsi que les mesures qui doivent être prises pour les limiter, sont décrits dans le tableau 3.

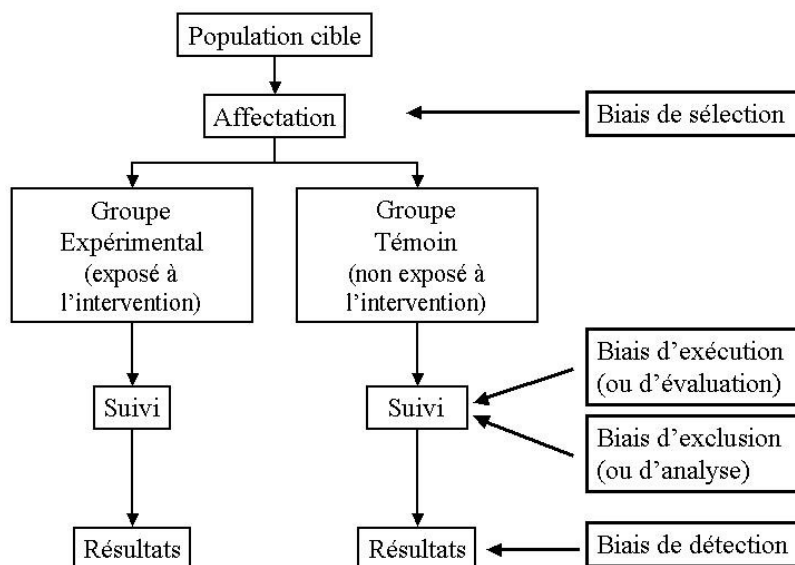


Figure 1. Schéma général d'une étude comparative d'intervention. Exemple : essai thérapeutique

Tableau 3. Biais et procédure permettant de limiter leur impact.

Type de Biais	Définition	Procédure limitant ce biais
Sélection	Introduction d'une différence systématique entre les groupes. Les groupes de sujets à comparer initialement seront différents en termes de facteurs pronostiques connus mais aussi inconnus.	Randomisation des interventions étudiées pour répartir de la même façon dans les deux groupes en début d'étude les facteurs mesurables et non mesurables. Vérifier quels sont les sujets inclus dans l'étude, quels sont les sujets non inclus.
Exécution (évaluation)	Introduction d'une différence systématique dans les soins et le suivi (intervention testée mise à part).	La notion d'aveugle permet de minimiser la part de subjectivité.
Exclusion (analyse)	Résulte notamment de l'existence et de l'importance du nombre des sujets perdus de vue pendant l'étude et qui seront éventuellement exclus de l'analyse des résultats.	Limiter, par la qualité de la réalisation de l'étude, le nombre de sujets perdus de vue. Les prendre en compte dans l'analyse, même si on ne dispose pas de leurs résultats.
Détection	Biais dans la mesure du critère choisi pour l'évaluation du résultat.	Critère de jugement mesurable, plutôt que déclaratif, fiable, validé.

3.5.2. Notions de métrologie. Qualité des mesures. Qualité des indicateurs

Afin d'éliminer des erreurs dues à l'instrument de mesure, deux qualités doivent être recherchées pour assurer la validité des variables faisant l'objet de l'étude :

a) L'exactitude, car l'information doit refléter la réalité qu'elle est censée décrire. Par

exemple la mesure de la pression artérielle systolique sanguine à l'aide d'un appareil automatique doit être le reflet de la PAS réelle.

- b) Ce reflet n'est jamais parfait et dépend de la précision de la mesure : chez un même individu et dans les mêmes conditions, les résultats de plusieurs mesures de la PAS par un appareil automatique doivent être du même ordre.

Ces deux propriétés participent à l'appréciation de la fiabilité des informations, c'est-à-dire au degré de confiance qu'on leur accorde. Ainsi, le choix de la variable est très important.

Exemple : si l'on veut connaître, dans une enquête, la consommation en graisses de sujets, on peut procéder de plusieurs façons. La plus précise, mais aussi la plus coûteuse, consiste à enregistrer quotidiennement la liste et la quantité des aliments consommés. Parfois, le schéma (étude rétrospective) et le financement de l'étude ne permettent pas ce recueil d'information. On procédera alors en utilisant une information moins précise et moins fiable comme l'évaluation de cette consommation par le seul interrogatoire des sujets. Dans ce cas, la fiabilité du résultat de l'analyse est mise en question et doit être discutée.

Les qualités recherchées d'une mesure sont : fiabilité, reproductibilité, disponibilité en pratique clinique, sensibilité et spécificité. Ces caractéristiques sont appréciées à l'aide d'indicateurs dont un résumé est donné dans le tableau 4.

Le choix d'un indicateur se fait parfois sur des critères de coût ou d'une disponibilité plus aisée. Il est souvent plus pratique d'utiliser ce que l'on appelle une « variable intermédiaire » pour mesurer indirectement un effet clinique, que de mesurer directement cet effet clinique. Il faut cependant être vigilant sur cette manière de procéder, car il se peut que la relation entre la variable intermédiaire et la variable clinique soit trompeuse.

Exemple : l'évaluation de l'efficacité d'une radiothérapie pour un cancer du sein peut reposer sur la diminution de la masse tumorale au scanner. Mais, bien que la réduction de la masse tumorale soit un facteur de bon pronostic à court terme, elle ne s'accompagne pas pour autant d'une amélioration de la survie des patientes à plus long terme. Pour qu'une variable intermédiaire soit reconnue comme une variable dite de substitution elle doit prédire de manière fiable l'effet clinique étudié.

Tableau 4. Indicateurs appréciant la qualité d'une mesure.

Caractéristique	Autre dénomination	Élément de réponse
-----------------	--------------------	--------------------

Sensibilité	Taux de vrais positifs	Qualité de ce test pour identifier les personnes qui sont malades?
Spécificité	Taux de vrais négatifs	Qualité de ce test pour exclure à bon escient les personnes qui ne présentent pas la maladie?
Valeur prédictive positive	Probabilité post-test d'un test positif	Si un sujet est testé positif, quelle est la probabilité pour qu'il présente la maladie?
Valeur prédictive négative	Probabilité post-test d'un test négatif	Si un sujet est testé négatif, quelle est la probabilité pour qu'il n'ait pas la maladie?
Précision		Quelle proportion de tous les tests pratiqués a donné un résultat correct?
Rapport de vraisemblance positif		Dans quelle proportion trouve-t-on un test positif chez un sujet malade par rapport à un sujet sain?
Rapport de vraisemblance négatif		Dans quelle proportion trouve-t-on un test négatif chez un sujet sain par rapport à un sujet malade?

3.5.3. Quelques impératifs méthodologiques liés à l'utilisation de la statistique

L'utilité des tests et des analyses statistiques tient à ce qu'ils permettent de résumer les données en un petit nombre de valeurs, de quantifier l'incertitude inhérente au jugement sur échantillons et de réaliser l'inférence statistique, c'est-à-dire tirer des conclusions fiables sur une population, à partir de l'étude d'un échantillon de cette population.

Dans le jugement sur échantillon - base des études cliniques et épidémiologiques - il faut souligner un autre point méthodologique important : la puissance de l'étude. La puissance d'une étude est sa capacité de détecter la plus petite différence scientifiquement ou cliniquement intéressante fixée par les investigateurs de l'étude.

Or, les études portant sur de petits échantillons manquent en général de puissance. C'est-à-dire que, par exemple, dans une comparaison de deux traitements sur des échantillons de petites tailles, des différences cliniquement importantes peuvent ne pas apparaître au sens statistique du terme, alors que ces mêmes différences auraient pu être statistiquement significatives dans une étude portant sur des échantillons de plus grande taille. Aussi, doit-on retrouver dans la partie « Méthodes » d'un bon article, la mention du calcul du nombre de sujets nécessaire à l'étude, qui explique les tailles des échantillons étudiés et justifie la puissance.

A l'inverse, de petites différences risquent d'être statistiquement significatives si l'étude porte sur des échantillons de grande taille alors même qu'elles sont insignifiantes du point de vue clinique. Prenons l'exemple d'un essai thérapeutique randomisé en double aveugle contre placebo, réalisé pour évaluer l'efficacité de la prise orale de prostacycline sur la distance maximale de marche chez des sujets présentant une artériopathie chronique oblitérante des membres inférieurs (stade II de la classification de Leriche et Fontaine). Le

groupe traité (n = 702) et le groupe placebo (n = 700) étaient comparables au début de l'étude (en particulier pour la distance de marche, l'ancienneté de la claudication intermittente, les traitements antérieurs). Par rapport à l'état basal, la distance maximale de marche a respectivement augmenté de 68 % et de 53 % dans le groupe traité et dans le groupe placebo (p = 0,01). Cet accroissement relatif correspond en fait à un périmètre moyen de marche de 218 mètres dans le groupe traité et de 204 mètres dans le groupe placebo, soit un gain moyen absolu de 15 mètres. Cette considération doit amener tout lecteur à examiner, outre la signification statistique des résultats, l'amplitude concrète des différences observées.

Le nombre d'hypothèses testées dans un article mérite également attention. En effet, si l'on travaille au seuil de signification de 5%, en moyenne une comparaison sur vingt sera statistiquement significative par le seul fait du hasard. Il faut donc se méfier des études qui présentent des résultats de dizaines de comparaisons, car ces comparaisons multiples sont susceptibles d'induire des résultats faussement significatifs.

Les résultats de significativité sont habituellement fournis avec une P-valeur habituellement appelée le « petit p ». Cette P-valeur est la probabilité d'obtenir une différence au moins aussi grande que la différence observée, si l'hypothèse nulle du test statistique est vraie. En d'autres termes, la P-valeur correspond à la probabilité que les différences observées dans une étude soient dues au seul hasard. C'est la probabilité de conclure à l'existence d'une différence si l'on répétait la comparaison à plusieurs reprises. De ce fait, on remarquera que pour rendre compte du résultat d'un test statistique, la valeur exacte de p est bien plus informative que le résultat « test significatif » ou « test non significatif », établi à partir d'une valeur de p choisie arbitrairement (généralement ce seuil de signification est de 5 %).

Outre le « p » il est utile, dans une étude, de quantifier la différence observée. Comme nous le verrons par la suite, p ne permet pas de quantifier un effet observé. D'autres indicateurs tels que la différence arithmétique de moyennes, le risque relatif, ..., permettent de quantifier cet effet. D'autre part, il est toujours nécessaire et pertinent de calculer l'intervalle de confiance de l'effet estimé : cet intervalle, généralement construit à 95 %, signifie qu'il y a 95 % de chances que la vraie valeur de l'effet observé soit comprise entre sa borne minimale et sa borne maximale.

4. Quelques checklists

Il existe des règles communes à la lecture critique de tout article médical : article traitant d'un essai thérapeutique, de l'évaluation d'un test diagnostique, d'une étude portant sur des facteurs pronostiques ou encore de la causalité d'un facteur de risque. Il est donc utile

de disposer de grilles de lectures, de *checklists*, pour aider à la réalisation d'une lecture critique systématisée et rapide. Le tableau 5 précise certaines questions à se poser systématiquement lors de la lecture critique d'un article scientifique médical.

Tableau 5. Première synthèse de la lecture critique

Pouvez-vous trouver cette information dans l'article ?	Problème de méthodologie ?	Impact de ce problème sur la validité de l'étude ?
Quelle est la question posée par les auteurs ?	S'agit-il de l'évaluation de l'impact d'une intervention? De l'importance d'un problème de santé?...	Impact de ce problème sur la validité de l'étude?
Quel est le schéma de l'étude?	Est-il approprié à la recherche?	Si non, les résultats sont-ils utiles?
Quelles sont les variables étudiées et comment sont-elles mesurées?	Les variables sont-elles pertinentes?	Biais et erreurs de mesures?
Y a-t-il des facteurs de confusion potentiels?	Erreurs de mesures? Ces facteurs ont-ils été contrôlés?	Impact de l'importance du problème sur les résultats?
Quelle est la méthode d'échantillonnage?	Biais de sélection?	Impact de l'importance du biais sur les résultats?
Des tests statistiques ont-ils été utilisés?	Les tests sont-ils appropriés aux données? Les intervalles de confiance sont-ils fournis?	Les test sont-ils appropriés aux données? Les intervalles de confiance sont ils donnés? Puissance?
Les résultats ont-ils un intérêt clinique ou « social »?	La taille de l'échantillon était-elle adaptée pour détecter cet intérêt (puissance)?	Le résultat est-il utile ou non concluant?
Quelles sont les conclusions des auteurs?	Ces résultats sont-ils intéressants pour la population qui vous intéresse?	

5. Discussion

L'objectif de la lecture critique d'un article scientifique médical est de savoir s'il y a lieu d'intégrer (et comment) une information (ou connaissance) nouvelle dans la pratique quotidienne. Les qualités méthodologiques évoquées ci-dessus servent à donner une idée de la crédibilité que l'on peut accorder à cette information. Cependant, les qualités méthodologiques ne suffisent pas à faire du résultat d'une étude un résultat utile en pratique courante. Rappelons que cette approche doit être centrée sur le patient.

Il convient donc de savoir si les résultats présentés permettent d'envisager ou non un bénéfice clinique intéressant pour la pratique et sont susceptibles de s'appliquer aux patients « tout venant ». Ces interrogations doivent naturellement amener le lecteur à analyser :

- Les caractéristiques des patients inclus : les patients sur lesquels a porté l'étude sont-ils semblables à ceux de la pratique médicale courante ? On s'interrogera alors sur les

conséquences prévisibles d'une sélection excessive des patients sur le sexe, l'âge, les comorbidités associées, ... La définition de la maladie correspond-elle aux critères diagnostiques consensuels ? Les résultats d'une étude dans laquelle la définition de la maladie repose sur des explorations non réalisables en pratique courante présenteront peu d'intérêts dans ce cadre d'exercice.

- La pertinence des critères de jugement : la mesure de certains indicateurs ne pose pas de problèmes particuliers, comme par exemple le décès³; cette pertinence des critères de jugement a une dimension méthodologique (qualité intrinsèque des indicateurs utilisés) et une dimension clinique (les indicateurs utilisés apprécient-ils au mieux le problème clinique étudié ?).
- La taille de l'effet observé : l'effet est-il intéressant (suffisamment important) du point de vue clinique par rapport aux critères statistiques retenus ? Il existe différentes manières pour présenter des résultats. Les règles de la communication font que tout auteur communique les résultats de ses travaux en utilisant un discours positif. Une étude sur le risque d'hémorragie gastro-intestinale suite à un traitement par AINS en est un exemple. Dans une étude, le traitement AINS étudié multipliait le risque d'accident hémorragique par 1,5 (l'intervalle de confiance à 95 % allant de 1,2 à 2,0) ce qui, exprimé de la sorte, n'est pas négligeable cliniquement. Pourtant, cette même information peut également se traduire sous une autre forme : il faudrait suivre 100 patients pendant 20 ans pour observer, en moyenne, une hémorragie supplémentaire, présentation qui fait alors perdre à l'information toute sa notion de gravité, au sens clinique du terme.

La lecture critique d'un article médical scientifique fait donc appel à des connaissances méthodologiques et à des connaissances cliniques. Toute critique doit reposer sur des arguments objectifs et faire appel au bon sens du lecteur.

Références

- [1] Bouvenot G. La médecine factuelle. *Thérapie* 1996 ; 51 :209-211.
- [2] Carson JL, Strom BL, Soper KA, West SL, Morse ML. The association of nonsteroidal anti-inflammatory drugs with upper gastrointestinal tract bleeding. *Arch Intern Med* 1987 ; 147(1):85-8.

³ En fait, la détermination de la cause exacte du décès peut s'avérer difficile dans certaines situations. Si l'on s'intéresse à la mortalité par infarctus du myocarde et au fait qu'une personne soit décédée dans un accident de la route, le décès est-il attribuable à l'accident lui-même ou bien est-ce un infarctus qui est à l'origine du décès entraînant un accident de la route ?