

LECTURE CRITIQUE DES ARTICLES MÉDICAUX

V. Juger de l'intérêt d'une nouvelle thérapeutique

**J.P. COLLET¹
L.R. SALMI²
ET LE DÉPARTEMENT
D'ÉPIDÉMIOLOGIE
ET DE BIostatistique
DE L'UNIVERSITÉ
McMASTER³**

La lecture d'un compte rendu d'essai thérapeutique doit permettre de différencier les traitements efficaces de ceux qui ne le sont pas, et de repérer les dangers inhérents à leur prescription (fig. 1). Quelques exemples cliniques vont montrer l'intérêt d'une lecture critique de cette littérature pour un meilleur traitement des patients :

- L'examen biologique d'un cadre de 48 ans a trouvé une concentration élevée de cholestérol plasmatique ; il n'a pas d'antécédents particuliers, un examen clinique normal, et un ECG sans anomalies. On lui prescrit un régime sans graisses, et un traitement par clofibrate. Qu'en pensez-vous ?
- On vient de découvrir un ulcère duodéal à un patient dont le père a lui-même été traité pour un ulcère gastrique 15 ans plus tôt (cryothérapie). On lui prescrit de la cimétidine. Quelles réflexions vous inspire cette situation ?
- On pose une indication de cerclage du col chez une femme enceinte qui a eu dans le passé une fausse couche spontanée. Est-ce justifié ?
- Un homme a eu une angine de poitrine sévère avec un important rétrécissement de l'artère coronaire gauche. Il a eu un pontage coronaire au cours duquel l'artère mammaire interne a été ligaturée. Qu'en pensez-vous ?
- Après six mois d'hospitalisation pour schizophrénie, une patiente sort avec de l'imipramine. Est-ce le bon traitement ?
- La méningite tuberculeuse d'un enfant est immédiatement traitée par streptomycine, isoniazide et rifampicine. Est-ce justifié ?

- Deux personnes âgées ont un syndrome identique d'ischémie cérébrale transitoire. L'une a eu une endartériectomie de la carotide interne, alors que l'autre a commencé un traitement prolongé à base d'aspirine. Comment expliquer cette contradiction apparente ?

- La montée de la pression artérielle de 110/70 à 140/90 chez une femme enceinte fait entreprendre un traitement par la méthyl-dopa. Est-ce souhaitable ?

- Une femme a une polyarthrite avec un facteur rhumatoïde positif ; on lui donne de l'indométacine. Auriez-vous fait de même ?

- La découverte d'une perte d'audition de 35 dB au cours d'un examen à l'école conduit un enfant à avoir une paracentèse avec pose de "yoyos". Est-ce une bonne indication ?

Ces patients ont été traités, soit pour prévenir la maladie (exemple du clofibrate), soit

pour la guérir (exemple des antibiotiques), ou pour l'améliorer (cas de l'indométacine). Toutes ces prescriptions se réfèrent à des études biologiques de référence et des suivis de cas similaires qui ont montré une évolution clinique favorable sous traitement (même pour la ligature de l'artère mammaire interne (1)).

Cependant, alors que nous considérons que certains traitements sont réellement utiles (c'est-à-dire qu'ils vont faire plus de bien que de mal aux patients qui les prennent), nous sommes plus circonspects quant à l'efficacité d'autres, et carrément ironiques à propos du consensus qui a préconisé par exemple le traitement de l'ulcère par le froid ou la ligature de l'artère mammaire interne pour améliorer l'angine de poitrine. Cette attitude est due au fait que nous voulons apprendre à partir de l'expérience, en espé-

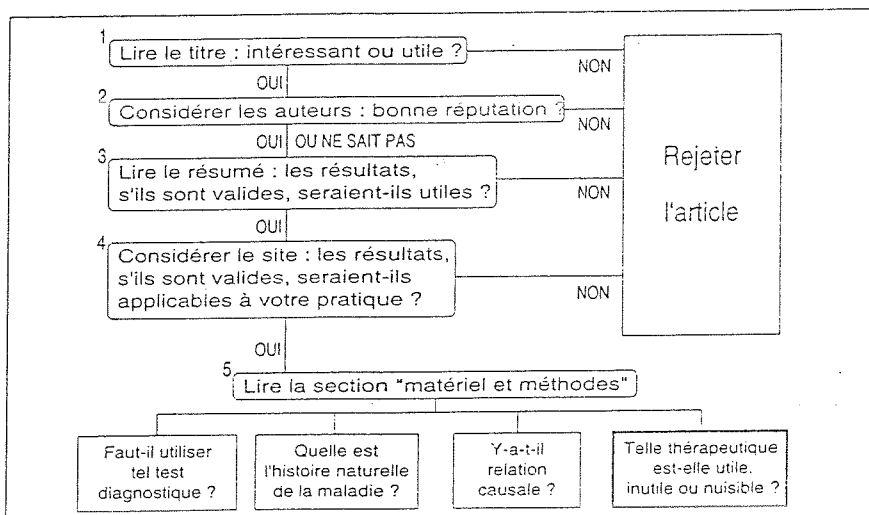


Fig. 1. Les premières étapes de la lecture critique d'un article médical.

1. Unité de pharmacologie clinique, université Claude-Bernard, 162, avenue Lacassagne, 69424 Lyon

2. Département d'épidémiologie et de biostatistique, Pasteur-Mérieux sérums et vaccins, 15-41, avenue Marcel-Mérieux, 69280 Marcy-l'Étoile

3. Cette série d'articles est une traduction et une adaptation de la série *How to read a clinical journal* du Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre, Hamilton, Ontario, Canada, publiée originellement dans le *Journal de l'Association Médicale Canadienne* en 1981

rant qu'un traitement inefficace ou qui fait plus de mal que de bien finira par être repéré. La médecine actuelle exige cependant que l'efficacité d'un produit soit montrée à l'aide de preuves solides obtenues à partir d'essais thérapeutiques avec distribution aléatoire des traitements.

Cet article décrit quelques règles qui permettent de lire de manière critique la littérature sur l'efficacité des médicaments. Les six règles proposées (tableau I) tiennent "du bon sens appliqué"; elles se divisent en deux groupes. Le premier concerne la validité de l'étude (les résultats de l'étude sont-ils vrais ?); le deuxième concerne l'utilité de ses résultats (les résultats de cette étude servent-ils à quelque chose ?).

L'ÉVALUATION DE LA THÉRAPEUTIQUE A-T-ELLE ÉTÉ FAITE DANS DE BONNES CONDITIONS ?

■ La distribution des patients dans les groupes de traitement a-t-elle été faite de manière aléatoire ? Chaque patient inclus dans l'étude doit avoir eu la même probabilité a priori (typiquement 50 %) dans une comparaison de deux traitements) de recevoir chacun des traitements à comparer. Il faut que la répartition des patients entre les deux groupes de traitement soit faite selon une règle analogue à celle qui existe quand on joue à pile ou face. Il est en

général facile de savoir si cette technique de tirage au sort a été utilisée : certaines expressions telles "essai clinique randomisé" ou "allocation aléatoire des patients" apparaissent dans la partie "méthode" de l'article ou parfois même dans le titre. En conséquence, le lecteur qui veut rentabiliser son temps de lecture peut négliger tous les articles sur l'évaluation des thérapeutiques qui ne sont pas des essais cliniques randomisés.

Pourquoi rejeter les résultats des études non randomisées ? Une explication formelle de cette règle serait longue, mais ses conclusions sont simples : l'allocation aléatoire élimine la plupart des biais rencontrés dans les études non randomisées. L'explication pragmatique est également simple : nous avons beaucoup plus de chance d'aider nos patients, et beaucoup moins de chance de leur nuire, si nous utilisons des médicaments dont l'efficacité a été prouvée au cours d'essais cliniques randomisés.

Il y a de nombreux exemples d'erreurs dues à l'utilisation de produits dont l'efficacité avait été appréciée au cours d'études non randomisées. Le clofibrate, par exemple, était largement utilisé au moment de la publication de l'essai randomisé (2) qui montrait qu'en fait ce produit augmentait la mortalité ; le médicament fut ensuite interdit dans plusieurs pays. De la même manière, on estime que 2 500 systèmes de cryothérapie pour ulcères gastriques ont permis de traiter des dizaines de milliers de patients avant qu'un essai clinique randomisé ne montre l'inefficacité du traitement (3). Enfin, un

essai randomisé, où des patients souffrant d'angine de poitrine ont eu soit une ligature de l'artère mammaire interne soit un simple abord chirurgical de l'artère sans ligature, a montré que la ligature de l'artère était inutile et que l'amélioration des symptômes pouvait être grande sous placebo (4).

Bywaters résume bien le problème quand il avance que, parmi les adeptes des régimes, il y a deux types d'individus : ceux qui ont de l'enthousiasme, mais pas de groupe témoin, et ceux qui ont un groupe témoin, mais pas d'enthousiasme (5). Sinclair a essayé de quantifier ce phénomène dans les articles sur la prise en charge des détresses respiratoires en pédiatrie (6). Il a montré que les résultats sont plus souvent positifs quand il n'y a pas de groupe témoin (environ 9 études non contrôlées sur 10 rapportaient un succès thérapeutique) que quand il y en a un (seulement la moitié des études contrôlées rapportaient un succès thérapeutique).

En résumé, s'il est vrai que l'essai randomisé peut parfois donner des résultats faux (surtout si l'échantillon est petit), il demeure néanmoins le meilleur instrument pour identifier les thérapeutiques efficaces.

Il existe pourtant une situation exceptionnelle où l'efficacité d'un traitement peut être appréciée sans groupe témoin : lorsqu'une maladie a, de manière constante, un pronostic fatal, comme c'était le cas pour la méningite tuberculeuse avant 1946. Au moment de l'introduction de la streptomycine aux États-Unis, en effet, quelques patients survécurent à la méningite tuberculeuse (7) : l'événement confirmé en Angleterre (8) constitua une preuve suffisante en faveur de l'efficacité de l'antibiotique. En dehors de cette situation très spéciale, il est préférable de ne pas tenir compte des articles d'évaluation quand l'étude n'a pas été réalisée avec allocation aléatoire des patients.

Le problème est différent lorsqu'un médecin doit rechercher dans la littérature le meilleur traitement pour son patient, s'il n'y a pas, comme c'est souvent le cas, d'essais randomisés réalisés dans ce domaine. Il faut alors utiliser les données issues d'études non expérimentales : en général on ne considérera alors que les résultats d'études suggérant une efficacité importante, afin d'être à peu près certain que la sélection des patients n'a pas pu biaiser les résultats. Cette dernière

TABLEAU I

Éléments de l'évaluation critique d'un article sur l'efficacité thérapeutique

A. L'évaluation de la thérapeutique a-t-elle été faite dans de bonnes conditions ?

1. La distribution des patients dans les groupes de traitement a-t-elle été faite de manière aléatoire ?
2. Tous les critères importants ont-ils été mesurés ?
3. Tous les patients randomisés sont-ils inclus dans l'analyse ?
4. La signification clinique et statistique des résultats est-elle discutée ?

B. La thérapeutique peut-elle être utile à vos patients ?

1. Les patients de l'étude sont-ils suffisamment semblables aux vôtres ?
2. La prise en charge thérapeutique est-elle réalisable dans votre pratique ?

règle est évidemment subjective et a permis dans le passé de préconiser le clofibrate, la ligature de l'artère mammaire interne, et la cryothérapie de l'ulcère gastrique. Cette situation est familière pour le clinicien qui doit souvent agir sans avoir une information suffisante.

La meilleure attitude dans ce cas est d'évaluer la probabilité que l'efficacité soit réelle et de mettre cette efficacité en balance avec les conséquences des choix possibles. L'état du patient requiert-il une quelconque action ? Si oui, y a-t-il, parmi les diverses possibilités, une intervention qui ait montré, au cours d'un essai contrôlé, qu'elle faisait plus de bien que de mal ? Si non, quelle est celle qui paraît la plus favorable en termes de bénéfice apporté et de risque encouru ? Dans cette situation, les critères d'évaluation pour juger de la qualité des articles traitant de la causalité proposés dans l'article précédent (9) seront particulièrement utiles.

■ Tous les critères importants ont-ils été mesurés ?

Le jugement sur l'utilité d'un traitement étant directement lié aux critères d'efficacité choisis, il importe de bien vérifier que tous les critères importants ont bien été étudiés.

L'importance de ce point est illustrée par l'exemple du clofibrate (2) dans le traitement de l'hypercholestérolémie (tableau II). Certains résultats sont en faveur d'une efficacité du traitement : la concentration du cholestérol sanguin (facteur de risque de la maladie coronaire) diminue de 10 % sous clofibrate, ce qui est en faveur d'une efficacité biologique du produit. Certains lecteurs cependant auront du mal à se contenter des résultats sur un critère intermédiaire biologique (10), et préféreront savoir si le risque d'accident cardiaque est lui-même réduit sous clofibrate. On constate de fait une diminution du nombre d'infarctus du myocarde (tableau II), ce qui semble démontrer l'efficacité du clofibrate. Pourtant, si l'on considère tous les résultats intéressants (en particulier du point de vue du patient, (11), il faut prendre en compte l'effet du clofibrate sur la qualité de vie d'une part et sur la mortalité totale qui s'avère en fait plus élevée sous clofibrate (tableau II).

Critère d'efficacité	Sous placebo	Sous clofibrate
Changement moyen du cholestérol sérique (%)	+ 1	- 9
Incidence de l'infarctus du myocarde non fatal (/1 000 hommes)	7,2	5,8
Incidence de l'infarctus du myocarde fatal et non fatal (/1 000 hommes)	8,9	7,4
Mortalité (toutes causes) (/1 000 hommes)	5,2	6,2

* Adapté de d'Oliver et al (12)

Incidence d'un événement critique	Traitement chirurgical : nombre (%)	Traitement médical : nombre (%)
Oui	43 (54)	53 (74)
Non	36 (46)	19 (26)
Total	79 (100)	72 (100)

Réduction du risque sous chirurgie :

$$[(53 / 72) - (43 / 79)] / (53 / 72) = 26 \%$$

$$\chi^2 = 5,98 ; p = 0,02$$

* Adapté de Fields et al (12)

Le lecteur attentif portera aussi son attention sur la qualité des critères utilisés qui doivent être explicites, objectifs et mesurés par un observateur qui ne connaît pas le type de traitement reçu par le patient. Ce dernier point est important car il assure que les patients distribués dans chaque groupe (principe actif et traitement de référence ou placebo) sont tous suivis de la même manière. On parle de double insu quand, ni le patient ni le médecin ne sont au courant du traitement administré. Le double insu prévient les biais d'information qui peuvent altérer les résultats.

■ Tous les patients randomisés sont-ils inclus dans l'analyse ?

Le lecteur d'un essai randomisé devrait toujours noter le nombre de patients inclus dans l'étude au moment de l'allocation aléatoire et vérifier que tous ces patients sont inclus dans les analyses qui sont la base des conclusions de l'article. Le tableau III, par exemple, décrit le devenir de patients inclus dans un essai thérapeutique randomisé comparant les prises en charge médicale et chirurgicale des sténoses bilatérales de l'artère carotide (12). Le nombre total de patients qui ont pu être suivis était de 151 en fin d'étude : 79 traités

chirurgicalement et 72 traités médicalement. Une réduction de 26 % du risque d'ischémies transitoires itératives, accidents vasculaires cérébraux ou décès fut observée chez les patients traités chirurgicalement : différence significative tant sur le plan clinique que statistique ($p = 0,02$). Pourtant, 167 patients avaient été initialement inclus dans l'étude, mais 16 d'entre eux avaient eu un accident vasculaire ou étaient morts au cours de l'hospitalisation initiale, et avaient été exclus de l'analyse finale. Il se trouve que 15 d'entre eux faisaient partie du groupe traité par la chirurgie : 5 décès et 10 accidents vasculaires postopératoires. Les résultats de l'analyse après réintroduction de ces patients montrent que la réduction du risque après chirurgie n'est que de 17 %, et n'est plus significative sur le plan statistique ($p = 0,09$) (tableau IV).

Les auteurs du rapport utilisé dans cet exemple ont donné l'information sur tous les patients, ce qui a permis de construire le tableau IV. Mais que peut faire le lecteur quand le devenir des sujets manquants n'est pas rapporté ? Une approche conservatrice consiste à attribuer arbitrairement un mauvais résultat à toutes les personnes manquantes du groupe qui a le meilleur devenir. Si, malgré cette manœuvre, les résultats restent significatifs, alors le lecteur peut accepter les conclusions de l'étude.

■ La signification clinique et statistique des résultats est-elle discutée ?

La signification clinique d'un résultat fait référence à l'importance de la différence observée

sur le critère clinique d'efficacité ; elle est habituellement décrite en terme d'amplitude d'un résultat. Dans le tableau II, par exemple, on constate que les patients sous clofibrate ont un risque de décès augmenté de $(6,2 - 5,2)/5,2 = 19\%$ par rapport au groupe sous placebo. Une telle différence devient cliniquement significative quand elle conduit à un changement des habitudes médicales. La perception par les praticiens de la signification clinique d'une augmentation de 19 % du taux de mortalité a été reflétée par la réduction considérable de la prescription de clofibrate après la diffusion de l'information. Dans cet exemple, la signification clinique est liée au point de vue médical ; dans d'autres circonstances, le point de vue du malade peut être pris en compte, en particulier quand il faut juger de la signification de l'effet d'une thérapeutique sur la qualité de la vie.

La signification statistique au contraire ne nous donne aucun renseignements sur l'importance ou l'amplitude d'un effet, mais nous informe sur la probabilité que la différence constatée ne soit pas due à la chance mais soit réelle. En d'autres termes, la signification statistique d'une différence n'est rien d'autre qu'une prise de position sur la probabilité que cette différence soit due au seul hasard. Si la probabilité est faible que la différence de 19 % observée entre les patients sous clofibrate et ceux sous placebo soit due à la chance (disons inférieure à 5 % ou 0,05), par convention, cette différence est dite statistiquement significative. Cette probabilité est appelée "valeur p" ou "alpha" ; elle représente la probabilité de se tromper lorsque nous concluons qu'une différence existe, quand, en fait, elle n'existe pas.

L'élément déterminant de la signification clinique est le changement de l'activité clinique ; si les résultats d'une étude vous conduisent à abandonner un ancien traitement pour un nouveau, la différence d'effets entre les traitements est cliniquement significative. Les éléments associés à la signification statistique sont plus complexes. La signification statistique d'un résultat augmente (la valeur p diminue) quand le nombre de patients dans l'étude est plus grand, quand l'effet du traitement est stable d'un moment à l'autre ou d'un patient à l'autre, et quand la mesure du critère d'efficacité est fiable.

Sur les bases de ce qui vient d'être dit, le lecteur pressé peut utiliser deux points de repère. Premièrement, si la différence est statistiquement significative ($p < 0,05$), l'est-elle également sur le plan clinique ? Si c'est le cas, les résultats sont à la fois réels et intéressants. Deuxièmement, si la différence n'est pas statistiquement significative, il faut se demander s'il y avait assez de patients dans l'étude pour mettre en évidence une différence clinique potentiellement significative ? Comme nous l'avons déjà dit, le nombre de sujets dans une étude est un élément déterminant de la signification statistique. Ainsi, plus la population étudiée est grande, plus la différence de résultats aura de chances d'être statistiquement significative (réelle), même si elle est cliniquement sans intérêt (trop petite pour justifier un changement des habitudes thérapeutiques).

Au contraire, si une population étudiée est trop petite, même une différence clinique importante avec un grand intérêt clinique potentiel peut ne pas être statistiquement significative. Cette situation renvoie aux vocables "faible puissance", "problème d'erreur bêta", ou "risque d'erreur de deuxième espèce" qui sont rencontrés quand la conclusion est qu'il n'y a pas de différence alors qu'en fait il y en a une. Lorsqu'une différence n'est pas statistiquement significative, le lecteur doit donc regarder avec attention la différence des résultats cliniques pour juger d'un éventuel intérêt thérapeutique. Cet argument se trouve conforté par une étude qui montre que la plupart des essais thérapeutiques dont les résultats n'étaient pas significatifs manquaient en fait de patients pour montrer une réduction du risque de 25 ou même 50 % (13).

TABEAU IV

Résultats de l'essai randomisé comparant les prises en charge médicale et chirurgicale de la sténose carotidienne bilatérale : tous patients randomisés*.

Incidence d'un événement critique	Traitement chirurgical : nombre (%)	Traitement médical : nombre (%)
Oui	58 (62)	54 (74)
Non	36 (38)	19 (26)
Total	94 (100)	73 (100)

Réduction du risque sous chirurgie :
 $[(54 / 73) - (58 / 94)] / (54 / 73) = 17\%$
 $X^2 = 2,80 ; p = 0,09$

* Adapté de Fields et al (12)

LA THÉRAPEUTIQUE PEUT-ELLE ÊTRE UTILE À VOS PATIENTS ?

Les patients de l'étude sont-ils suffisamment semblables aux vôtres ?

Pour répondre à cette question, il faut d'abord pouvoir identifier les patients inclus dans l'étude ; leur état clinique et leurs caractéristiques socio-démographiques doivent être clairement décrits. Il faut en effet que vous puissiez juger si les patients de l'étude ressemblent suffisamment aux vôtres pour que vous puissiez appliquer les résultats à votre pratique. Si c'est le cas, le lecteur aura une bonne idée des résultats attendus lorsqu'il utilise la thérapeutique chez ses patients.

La prise en charge thérapeutique est-elle réalisable dans votre pratique ?

D'autres éléments doivent apparaître dans l'article. Le premier concerne la description de la prise en charge qui doit être suffisamment précise pour être reproduite. Qui a fait quoi ? à qui ? avec quel produit ? quelle dose ? dans quelle circonstance ? avec quelle adaptation posologique ? avec quels effets inattendus ou toxiques ? pour quelle durée ? et sur quels critères cliniques juger que le traitement doit être augmenté, réduit ou arrêté ?

La seconde exigence concerne la signification clinique et biologique de la prise en charge. Par exemple, la posologie, la voie d'administration et la durée de traitement doivent concorder avec les données de pharmacocinétique et pharmacodynamique.

La troisième exigence est la disponibilité de la thérapeutique.

La quatrième exigence enfin concerne la description de la prise en charge : le lecteur doit rechercher deux biais potentiels liés à son application, et que l'auteur a dû éviter. Il s'agit des biais de contamination et de co-intervention. Le biais de contamination survient lorsque des patients dans le groupe contrôle reçoivent accidentellement le traitement étudié, ce qui réduit artificiellement la différence clinique entre les

deux groupes étudiés. Le biais de co-intervention survient lorsqu'un complément diagnostique et thérapeutique est appliqué au groupe recevant le produit étudié, mais pas au groupe de référence ; dans ce cas, on assiste à une augmentation erronée et artificielle de la différence entre les groupes. Ce biais de co-intervention peut être prévenu par le double insu (14).

APPLICATION DE CES CRITÈRES À LA LECTURE

L'application de ces critères devrait vous permettre de décider si les résultats d'un essai thérapeutique sont utiles et transposables à votre pratique clinique.

Le critère le plus important pour l'évaluation de ces articles sera toujours l'existence d'un groupe témoin constitué par randomisation. En fonction du contexte dans lequel vous évaluerez l'article, l'un des cinq autres critères sera ensuite le plus important, et vous pourrez aller directement à ce critère. Si l'article a rempli ce critère, vous pourrez vérifier les autres critères ; au contraire, si le critère qui vous paraît le plus important n'est pas rempli, vous pourrez rejeter l'article. Cette méthode permet ainsi d'améliorer le rendement du temps que vous consacrez à la lecture. Seuls les articles qui respectent tous les critères apporteront une information valide et fiable que vous pourrez utiliser dans votre pratique.

Remerciements : Les auteurs remercient H. Maisonneuve et C. Dupuy pour leurs critiques et commentaires, et M. Portier pour la dactylographie.

Article reçu en juin 1991.

LECTURE CRITIQUE

DES ARTICLES MÉDICAUX

Articles déjà parus :

- Éditorial n° 25 du 1^{er}/12/91 - page 2597.
- I. Introduction, n° 25 du 1^{er}/12/91 - page 2605.
- II. Juger de l'intérêt d'un test diagnostique - n° 26 du 15/12/91 page 2734.
- III. Comprendre l'histoire naturelle et le pronostic d'une maladie - n° 1 du 1/1/92 page 76.
- IV. Déterminer la causalité - n° 2 du 15/1/92 page 213.

RÉFÉRENCES

1. Glover RP, Davila JC, Kyle RH, Beard JC, Trout RG, Kitchell JR. Ligation of the internal mammary arteries as a means of increasing blood supply to the myocardium. *J Thorac Cardiovasc Surg* 1957 ; 34 : 661-78.
2. Oliver MF, Heady JA, Morris JN, Cooper J. A co-operative trial in the primary prevention of ischaemic heart disease using clofibrate. Report from the Committee of Principal Investigators. *Br Heart J* 1978 ; 40 : 1069-118.
3. Miao LL. Gastric freezing : an example of the evaluation of medical therapy by randomized clinical trials. In : Bunker JP, Barnes BA, Mosteller F, eds. *Costs, risks and benefits of therapy*. New York : Oxford University Press ed. 1977 : 198-211.
4. Cobb LA, Thomas GI, Dillard DH, Merendino KA, Bruce RA. An evaluation of internal-mammary-artery ligation by a double-blind technic. *N Engl J Med* 1959 ; 260 : 1115-8.
5. Bywaters EG. Treatment of rheumatic fever. *Circulation* 1956 ; 14 : 1153-8.
6. Sinclair JC. Prevention and treatment of the respiratory distress syndrome. *Pediatr Clin North Am* 1966 ; 13 : 711-30.
7. Hinshaw HC, Feldman WH, Pfuetez KH. Treatment of tuberculosis with streptomycin : summary of observations on 100 cases. *JAMA* 1946 ; 132 : 778-82.
8. Medical Research Council. Streptomycin treatment of tuberculous meningitis. *Lancet* 1948 ; 1 : 582-96.
9. Salmi LR, Collet JP. Lecture critique des articles médicaux : IV. Déterminer la causalité. *Rev. Prat. (Paris)* 1991 ; 41 (2) : 213-8.
10. Yerushalmy J. On inferring causality from observed associations. In : Ingelfinger FJ, Rellman AS, Finland M, eds. *Controversy in internal medicine*. Philadelphia : Saunders ed. 1966 : 659-68.
11. Sackett DL, Gent M. Controversy in counting and attributing events in clinical trials. *N Engl J Med* 1979 ; 301 : 1410-2.
12. Fields WS, Maslenikov V, Meyer JS, Hass WK, Remington RD, MacDonald M. Joint study of extra-cranial arterial occlusion. V. Progress report of prognosis following surgery or non-surgical treatment for transient ischemic attacks and cervical carotid artery lesions. *JAMA* 1970 ; 211 : 1993-2003.
13. Freiman JA, Chalmers TC, Smith H, Kuebler RR. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial. Survey of 71 "negative" trials. *N Engl J Med* 1978 ; 229 : 690-4.
14. Sackett DL. On some prerequisites for a successful clinical trial. In : Shapiro DH, Louis TA, eds. *Clinical trials. Issues and approaches*. New York : Marcel Dekker ed. 1983 ; 65-79.