

REVUE DE STATISTIQUE APPLIQUÉE

D. SCHWARTZ

J. LELLOUCH

**Attitude de recherche ou attitude pragmatique
dans la formulation d'un problème, avec référence
particulière aux essais thérapeutiques**

Revue de statistique appliquée, tome 13, n° 4 (1965), p. 11-19

http://www.numdam.org/item?id=RSA_1965__13_4_11_0

© Société française de statistique, 1965, tous droits réservés.

L'accès aux archives de la revue « Revue de statistique appliquée » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

ATTITUDE DE RECHERCHE OU ATTITUDE PRAGMATIQUE DANS LA FORMULATION D'UN PROBLÈME, AVEC RÉFÉRENCE PARTICULIÈRE AUX ESSAIS THÉRAPEUTIQUES

D. SCHWARTZ et J. LELLOUCH

Unité de Recherches Statistiques de l'Institut National de la Santé
et de la Recherche Médicale, Institut Gustave Roussy, Villejuif (Seine)

Nous nous proposons de montrer ici -à propos des essais thérapeutiques- comment un problème, tel qu'il est énoncé au stade initial de son élaboration, s'avère presque toujours insuffisamment formulé, et ceci, non pas dans le détail, mais dans ses grandes lignes, en ce sens qu'il recouvre au moins deux problèmes, ou deux types de problèmes reflétant des attitudes entièrement différentes. Nous montrerons en effet que la comparaison de deux traitements conduit à ce choix d'attitude dès qu'il s'agit de définir a) les traitements et b) le mode de comparaison.

Parmi les deux attitudes possibles, il arrive le plus souvent en thérapeutique que l'une soit moins défendable que l'autre, voire indéfendable, pour des raisons d'éthique. Nous éviterons d'abord cet aspect, pour conserver au problème une plus grande généralité. Il fera ensuite l'objet d'une discussion particulière.

A - DEFINITION DES TRAITEMENTS

Exemple 1 :

Soit à comparer, pour le traitement d'un cancer, la radiothérapie seule et la radiothérapie précédée par l'administration d'un agent, sans effet propre, mais susceptible de sensibiliser le malade à l'effet des radiations. Imaginons que la durée de cette administration soit de 30 jours. On peut alors, pour le groupe "radiothérapie seule", choisir entre deux attitudes (figure 1) :

a) faire précéder la radiothérapie d'un temps mort de 30 jours, de manière à la situer au même moment dans les deux groupes.

b) effectuer la radiothérapie d'emblée, la situant ainsi, selon toute vraisemblance, au moment optimum.

Aucune de ces attitudes n'est "meilleure" que l'autre ; la première permet de comparer 2 groupes semblables vis à vis de la radiothérapie et ne différant que par l'administration de l'agent, donc d'apprécier son rôle sensibilisant, apportant ainsi une information précieuse sur le plan de la connaissance biologique. La seconde permet de comparer deux traitements dans les conditions où on les administrerait pratiquement.

Nous les appellerons attitude de recherche et attitude pragmatique.

On notera la raison qui, dans cet exemple, conduit à deux attitudes possibles : dès qu'un essai porte sur deux traitements complexes, formés d'une série d'opérations, dont on veut étudier l'une, la comparaison des deux traitements peut être envisagée, en ce qui concerne les autres opérations, "à conditions égales" ou "en conditions optimum". La première attitude renseigne sur le rôle de l'opération à étudier, l'autre compare deux traitements complexes dans leurs conditions pratiques d'administration.

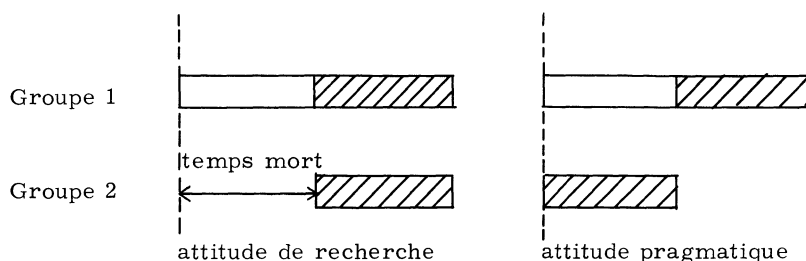


Figure 1 - Attitude de recherche et attitude pragmatique dans l'exemple 1.
(agent à étudier, radiothérapie)

Exemple 2 :

La double attitude n'est pas l'apanage des traitements complexes. Soit maintenant à comparer deux traitements tels que la radiothérapie par photons et par électrons pour certains cancers.

Il est particulièrement intéressant pour le biologiste de savoir si ces deux modes de rayonnement diffèrent dans leurs effets. Ceci, bien entendu, à condition de les administrer à dose égale.

Mais supposons qu'un des deux modes de traitement provoque des réactions d'intolérance de la peau, le praticien pourra être amené à le comparer à l'autre avec une dose plus faible. Plus précisément, il existe pour chacun des traitements un optimum, compte tenu des inconvénients d'une dose trop forte. Le problème pratique consiste à comparer l'optimum pour les deux modes.

Ici encore il faut choisir entre une attitude de recherche, qui apporte une information biologique importante, et une attitude pragmatique. La première compare les traitements "à conditions égales", la seconde "en conditions optimum" en ce qui concerne, cette fois, les inconvénients secondaires du traitement.

Sous cette forme, le problème se présente dans sa généralité. Les traitements sont administrés dans un certain contexte : soins auxiliaires, régime, mode de vie, inconvénients secondaires, etc... Ces contextes doivent être définis, et ils peuvent l'être de plusieurs manières, dont deux au moins apparaissent clairement : l'égalisation, si on veut connaître le rôle propre des traitements, et l'optimisation, si on cherche à choisir entre deux thérapeutiques.

Le problème dépasse largement le domaine médical : en agronomie par exemple on peut comparer deux variétés de plantes dans des conditions égales de compacité, d'engrais, etc... (attitude de recherche), ou en optimisant ces conditions, compte tenu du coût au sens large (attitude pragmatique).

Attitude de Recherche et Attitude Pragmatique.

Il convient de discuter les termes adoptés. N'est-elle pas grossière, voire indéfendable, la tendance à séparer recherche et pragmatisme ? Tout progrès d'ordre biologique est susceptible d'applications pratiques, tandis que la comparaison de deux traitements sur un plan pratique apporte un élément de connaissance.

Il reste que l'opposition des attitudes à un sens très clair qui est schématisé sur le tableau 1, à propos de l'exemple 1 (où on supposera, pour simplifier, les effectifs assez grands pour qu'on puisse négliger les fluctuations d'échantillonnage) :

- si on adopte l'attitude de recherche (radiothérapie seule différée) on obtiendra dans tous les cas une réponse sous l'angle de la recherche : l'agent S a, ou n'a pas de rôle sensibilisant. Par contre, on ne pourra tirer de conclusion pratique immédiate que si la radiothérapie est supérieure ou égale au traitement combiné, car alors ce sera vrai a fortiori pour la radiothérapie d'emblée.

- on verra de même que l'attitude pragmatique (radiothérapie seule d'emblée) apportera dans tous les cas une réponse au problème pratique (comparaison des deux traitements en conditions optimum) mais n'apportera d'information sur le rôle sensibilisant de l'agent que si le traitement combiné s'avère surpasser la radiothérapie seule.

En définitive, on aperçoit très bien que :

- *L'attitude de recherche conduit de toute manière à des conclusions "de recherche", non utilisables immédiatement. Elle peut aussi conduire à des applications immédiates, mais cette éventualité dépend du résultat de l'essai.*

- *L'attitude pragmatique conduit de toute manière à une application immédiate ; elle peut aussi apporter une information "de recherche", mais cette éventualité dépend du résultat de l'essai.*

Cette vue justifie bien les noms d'attitude de recherche et d'attitude pragmatique.

domaine des conclusions \ attitude adoptée	Recherche (.R)	Pragmatique (R)
Recherche	oui	si $SR > R$
Pragmatique	si $SR \leq .R$	oui

Tableau 1

Conclusions possibles, dans le domaine de la recherche et de l'application immédiate, selon l'attitude adoptée, dans l'exemple 1.

Légende : Radiothérapie d'emblée = R ; différée = .R
Radiothérapie précédée d'un éventuel sensibilisant = SR

B - MODE DE COMPARAISON

La comparaison de deux traitements A et B ne soulèverait aucune difficulté si on connaissait leur vrai "effet" ; mais cette connaissance exacte nécessiterait un nombre de malades infiniment grand ; avec des séries d'effectif limité, on ne peut conclure qu'à la condition de consentir certains risques d'erreur. On considère classiquement :

- le risque de première espèce α (erreur) : conclure à une différence alors que $A = B$.

- le risque de deuxième espèce β (manque de puissance) : conclure $A = B$ alors que les deux traitements diffèrent.

Nous envisagerons en outre le risque γ de conclure qu'un des deux traitements est supérieur à l'autre, alors qu'en fait il est inférieur (erreur de signe) et nous l'appellerons risque de troisième espèce.

Nous nous proposons de montrer qu'il n'existe pas une façon unique de comparer deux traitements, mais au moins deux façons, recouvrant des attitudes radicalement différentes.

Première attitude.

Soit à comparer deux antalgiques A et B, de formule très voisine, l'une résultant de l'autre par modification d'un radical. Nous supposons qu'il soit intéressant, sur le plan scientifique, de savoir si cette modification se répercute sur l'efficacité de la drogue. Il est alors grave de conclure à une différence qui n'existerait pas, et ce risque d'erreur α doit être minimisé. La solution est classique : c'est le test statistique, où on ne conclut que si la différence observée dépasse un certain seuil, garantissant la valeur α consentie. La seule raison qui empêche de prendre α trop faible est qu'alors le seuil, et par suite β , augmentent. On peut assigner à β une certaine limite, à condition de prendre un nombre de sujets plus élevé. Quant au risque γ , il n'est pas habituel d'y faire allusion, tout simplement parce que si α et β sont faibles, γ est négligeable (le calcul montre que pour $\alpha = \beta = 5\%$, on a $\gamma < 10^{-7}$).

Ces considérations sont illustrées par la figure 2, où la variable étudiée est supposée quantitative (temps de rémission de la douleur par exemple) ; soient μ_A , μ_B les moyennes vraies avec les traitements A et B, m_A et m_B les moyennes qui seront observées sur les sujets de l'essai, n le nombre de ces sujets, supposé le même dans les deux groupes, σ^2 la variance supposée la même avec les deux traitements :

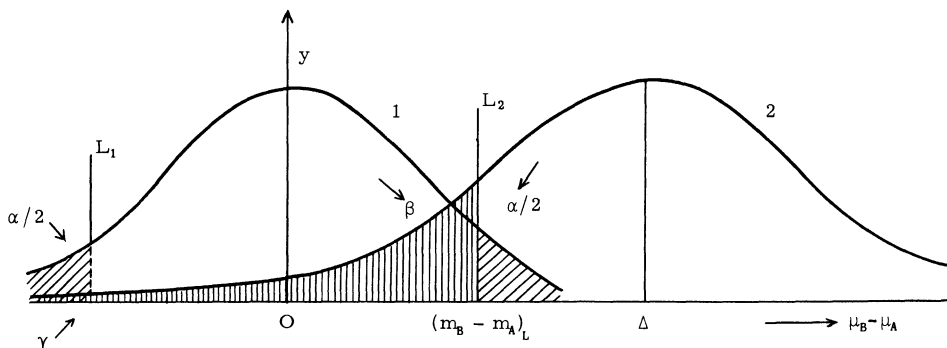
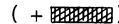
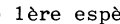
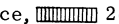


Figure 2 - Fluctuations de $(m_B - m_A)$ dans l'hypothèse nulle $\mu_B - \mu_A = 0$ (courbe 1) et dans l'hypothèse alternative $\mu_B - \mu_A = \Delta$ (courbe 2).

Risques :  (+ ) 1ère espèce,  2ème espèce,  3ème espèce.

- la courbe 1 représente, lorsque $\mu_B = \mu_A$ (hypothèse nulle), les fluctuations de $(m_B - m_A)$; on ne conclut $\mu_B \neq \mu_A$ que si $|m_B - m_A|$ dépasse la valeur seuil

$$|m_B - m_A|_L = \varepsilon_\alpha \sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2}{n}} = \varepsilon_\alpha \sqrt{\frac{2\sigma^2}{n}}$$

où ε_α désigne l'écart réduit correspondant au risque α (pour $\alpha = 5\%$ par exemple, $\varepsilon_\alpha = 1,96$).

- la courbe 2 fait apparaître les risques ainsi encourus, s'il se trouve que $\mu_B - \mu_A = \Delta$ (hypothèse alternative). Pour limiter l'ensemble $(\beta + \gamma)$ il faut resserrer la courbe 2 autour de son axe en prenant un nombre de sujets suffisant, de telle manière que :

$$\Delta - (m_B - m_A)_L \geq \varepsilon_{2(\beta+\gamma)} \sqrt{\frac{2\sigma^2}{n}}$$

où $\varepsilon_{2(\beta+\gamma)}$ désigne l'écart-réduit correspondant au risque $2(\beta + \gamma)$, le risque $(\beta + \gamma)$ étant ici multiplié par 2 parce qu'il n'est envisagé que d'un côté de la courbe 2, alors que les tables classiques de l'écart-réduit correspondent à un intervalle symétrique.

Compte tenu de la valeur de $(m_B - m_A)_L$ indiquée ci-dessus, ces conditions impliquent un nombre de sujets par série :

$$n \geq [\varepsilon_\alpha + \varepsilon_{2(\beta+\gamma)}]^2 \frac{2\sigma^2}{\Delta^2} \quad (1)$$

Si on supprime γ , généralement négligeable, on obtient ainsi la formule classique en test bilatéral :

$$n \geq (\varepsilon_\alpha + \varepsilon_{2\beta})^2 \frac{2\sigma^2}{\Delta^2} \quad (2)$$

Cependant la différence $\mu_B - \mu_A = \Delta$ est bien entendu inconnue. Pour déterminer n , on fixe alors la différence Δ qu'on ne veut pas laisser échapper avec le risque consenti β ; c'est celle-ci qui figure dans la formule 2. Le couple (β, Δ) détermine en quelque sorte le "pouvoir séparateur" de la comparaison (plus β et Δ sont petits, meilleur est le pouvoir séparateur).

Ces considérations font ressortir que la "comparaison" des traitements A et B est une opération mal définie ; il s'agit, en fait, d'une comparaison à des niveaux de risque déterminés" ; la fixation de α ne suffit pas, car elle se répercute sur β ; la fixation de β ne suffit pas, si on n'indique pas à quelle valeur de Δ ce risque est attaché. L'opération n'est donc finalement définie que si on précise le risque d'erreur α et le "pouvoir séparateur" (β, Δ) . Ces trois données se retrouvent dans le nombre de sujets nécessaire.

Deuxième attitude.

Supposons maintenant qu'il s'agisse de comparer deux antalgiques, de formule très différente, et ceci uniquement à des fins pratiques : on veut savoir lequel il vaut mieux adopter.

On s'aperçoit immédiatement que le risque α devient sans gravité : si $A = B$, en quoi peut-il être gênant d'adopter A (ou B) ? Cette op-

tique est si différente de la précédente qu'il convient de s'y attarder un moment. Il est bien clair que, d'un point de vue purement pratique, si les deux traitements sont équivalents, il n'y a aucun inconvénient à adopter l'un ou l'autre. On peut d'ailleurs remarquer que, dans la comparaison pratique de ces deux drogues, l'hypothèse $B - A = 0$ n'est qu'une parmi d'autres, d'ailleurs fort peu probable.

Ainsi le risque d'erreur, qui constituait le centre d'intérêt et dont le contrôle était la motivation essentielle du test statistique, va pouvoir disparaître de nos préoccupations, conférant à la comparaison un aspect radicalement différent.

Etant donné en effet que le risque α nous indiffère, il est logique, non de le minimiser mais, vu ses répercussions sur le nombre de sujets, de lui donner au contraire sa valeur maximum : 100 %. Cela revient, dans la figure 2, à rapprocher les deux limites L_1 et L_2 jusqu'à l'axe Oy. Dès lors :

a) il n'est pas nécessaire de faire un test, car toute différence est significative au risque 100 %. Il suffit d'adopter le traitement qui donne la meilleure moyenne.

b) avec ce procédé, on conclut toujours. C'est bien d'ailleurs ce que montre la figure 2 : si L_1 et L_2 sont confondus avec Oy, le risque β est nul.

c) c'est ici qu'intervient le risque de troisième espèce, qui n'est plus du tout négligeable. Il est par ailleurs grave d'adopter un traitement, si c'est le moins bon. Il faut donc minimiser γ ; mais γ n'a pas de sens isolément, il faut le rapporter à Δ , c'est-à-dire limiter la probabilité γ d'adopter un traitement moins bon que l'autre d'une quantité Δ ; le couple (γ, Δ) détermine en quelque sorte la "limite de sécurité" de la comparaison.

La "comparaison" des traitements A et B reste "une comparaison à des niveaux de risque déterminés", mais où il est intéressant de définir la "limite de sécurité" (γ, Δ) , les autres risques ayant les valeurs optimum $\alpha = 100 \%$, $\beta = 0 \%$.

Quant au nombre de sujets nécessaire, il reste donné par la formule (1), celle-ci devenant :

$$n \geq (\epsilon_{2\gamma})^2 \frac{2\sigma^2}{\Delta^2} \quad (3)$$

Attitude de recherche et attitude pragmatique.

On voit combien les deux attitudes précédentes diffèrent :

- dans la première, on veut savoir s'il existe ou non une différence ; la proclamer à tort est une erreur. Plutôt que de la commettre, on préfère ne pas conclure. La réponse est donnée par un test.

- dans la deuxième, on veut savoir lequel des deux traitements adopter. Adopter l'un, s'il égale l'autre, n'est pas grave. On conclut à tout coup, sans test.

Ces deux démarches peuvent être qualifiées d'attitude de recherche dans le premier cas, d'attitude pragmatique dans le second. Si nous utilisons ici les mêmes noms que dans la partie A, c'est qu'ils recouvrent

une même notion, caractérisant la formulation du problème dans son ensemble : c'est ainsi que le problème des deux antalgiques de formule voisine devrait être abordé dans une attitude de recherche, non seulement pour le mode de comparaison, émis aussi pour la définition des traitements (conditions égales) ; tandis que le problème des deux antalgiques de formule très différente devrait être abordé dans une attitude pragmatique, non seulement pour le mode de comparaison, mais aussi pour la définition des traitements (conditions optimum).

Répercussion sur le nombre de sujets nécessaires.

Reprenons l'exemple des deux antalgiques. Admettons :

- dans une attitude de recherche, $\alpha = 5 \%$ et $\beta = 5 \%$ pour $\Delta = 1/4$ heure (on veut avoir moins de 5 chances sur 100 de laisser échapper une différence de $1/4$ heure dans les durées de rémission). D'après (2),

$$n = (1,960 + 1,645)^2 \frac{2\sigma^2}{(1/4)^2}$$

- dans une attitude pragmatique, $\alpha = 100 \%$ (d'où $\beta = 0 \%$), $\gamma = 5 \%$ pour $\Delta = 1/4$ heure (on veut avoir moins de 5 chances sur 100 d'adopter une drogue donnant $1/4$ heure de rémission de moins que l'autre). D'après (3),

$$n = (1,645)^2 \frac{2\sigma^2}{(1/4)^2}$$

Bien entendu, le deuxième résultat est inférieur au premier du fait de l'absence du terme ϵ_a , et ceci est vrai d'une manière générale : le nombre de sujets est plus faible dans l'attitude pragmatique que dans l'attitude de recherche, à valeur égale de Δ et du risque correspondant (γ dans le premier cas, β dans le second). Mais la conclusion réelle est plus complexe, car il n'y a pas de raison d'adopter pour la "limite de sécurité" (γ, Δ) les mêmes valeurs que pour le "pouvoir séparateur" (β, Δ). La fixation de ces valeurs est affaire de circonstances, et il ne paraît pas possible de donner une règle générale.

Cas de la démarche unilatérale.

Nous avons traité jusqu'ici le cas symétrique ou bilatéral dans lequel on recherche si B diffère de A, cette différence pouvant être aussi bien positive que négative.

Soit maintenant à comparer à un antalgique classique A un antalgique nouveau B. Celui-ci ne nous intéresse que s'il est supérieur à A. S'il est inférieur ou même égal au traitement classique, ce dernier sera conservé. Le problème est alors de type dissymétrique ou unilatéral.

La question qui vient immédiatement à l'esprit est la suivante : à partir de quelle différence le traitement nouveau sera-t-il préféré à l'ancien ? S'il donne une durée de rémission plus élevée d'au moins $1/4$ d'heure, $1/2$ heure, 1 heure ? Supposons que cet "handicap" soit de 1 heure, c'est-à-dire que si $B - A = 1$ heure, il est indifférent d'utiliser A ou B : pour ce "point d'équilibre" la démarche est bilatérale ; à condition de retrancher 1 heure aux durées de rémission procurées par B, on est ramené au problème précédent.

La fixation de cet "handicap" est, bien entendu, malaisée, et on peut être certain que le médecin prié de la chiffrer fera la grimace. En fait, on ne voit pas que la difficulté puisse être esquivée. Procéder d'abord à l'essai, et choisir entre A et B après coup, selon la différence observée, revient seulement à reporter cette difficulté. Dès lors qu'on a admis la fixation à priori de α , Δ , β ou γ , la fixation de l'handicap n'est que la suite logique de cette politique.

Cette fixation est, bien entendu, de nature pragmatique : c'est que l'essai envisagé ici est de nature pragmatique. Il reste à se demander comment le problème devrait être résolu dans un essai axé sur une attitude de recherche. Notre opinion est que la recherche aboutit plus rarement à des problèmes unilatéraux car si une différence, qu'on attend positive ou nulle, se révèle en fait négative de manière significative, on hésitera le plus souvent à l'attribuer au hasard. Aussi la démarche unilatérale nous paraît-elle être, dans la majorité des cas, la signature d'une attitude pragmatique.

On remarquera que nous avons fait intervenir ici d'autres variables que l'effet principal du traitement, de façon à définir le coût global ; cette préoccupation nous avait échappé dans le problème bilatéral. Nous pensons que dans le problème bilatéral, on devrait également considérer le coût global, du moins si on adopte l'attitude pragmatique ; tandis qu'on examinerait plutôt un effet déterminé du traitement si on adopte l'attitude de recherche.

C - PROBLEME D'ETHIQUE

Si la définition des traitements et celle de leur mode de comparaison offrent le choix entre deux attitudes, l'une d'entre elles peut être éthiquement moins défendable que l'autre, voire indéfendable. Il reste donc, dans le domaine médical, à examiner ce point de vue.

Dans les exemples cités, une attitude apparaît difficilement défendable : retard dans la radiothérapie, radiothérapie par photons ou électrons à doses égales. D'une manière générale, si on peut définir les deux attitudes par "à conditions égales" pour l'une, "en conditions optimum" pour l'autre, seule cette dernière, l'attitude pragmatique, est admissible. Il résulte que seuls sont défendables les essais suivants :

- les essais purement pragmatiques,
- les essais mixtes, lorsque les "conditions égales" se confondent avec les "conditions optimum". Ainsi, si un retard de 30 jours en radiothérapie ne change rien, si les réactions cutanées aux photons et aux électrons sont pratiquement du même ordre, on pourra adopter une attitude de recherche qui sera en même temps l'attitude pragmatique.

Il reste à examiner quel mode de comparaison est éthiquement préférable. Les essais pragmatiques ne posent pas de problème, la comparaison devant y être effectuée sur le mode pragmatique. Mais dans les essais mixtes, il faut bien choisir l'une ou l'autre des deux attitudes, ne fût-ce que pour fixer le nombre de sujets nécessaires. Peut-on indiquer laquelle est la plus conforme à l'éthique ?

Cette décision ne peut résulter que d'un bilan où interviendraient notamment, d'une part les inconvénients d'un nombre de sujets élevé (la moitié des sujets reçoivent le moins bon des deux traitements ; en

outre, pendant la durée de l'essai, les malades extérieurs sont privés de celui des deux traitements qui s'avèrera le meilleur), d'autre part la contribution comparée des deux attitudes au progrès de la thérapeutique. Cette question n'admet pas de réponse générale, et le moins qu'on puisse dire est que, pour un problème donné, elle n'est guère simple. Nous ne pensons pas que ce soit ici le lieu pour en dire davantage.

D - RESUME ET CONCLUSION

L'exposé qui précède n'est qu'un aperçu. Il vise surtout à poser des questions, dont la solution demande sûrement à être discutée. Dans les grandes lignes, il peut être ainsi résumé :

Les termes "comparaison de deux traitements A et B" constituent une formulation incomplète d'un problème, qui recouvre au moins deux problèmes différents.

Un premier choix se présente dans la définition des traitements ; ceux-ci peuvent être administrés à égalité de toutes autres conditions, ou à optimum de ces conditions. La première attitude permet de comprendre le pourquoi d'une différence, la seconde indique le meilleur des deux traitements dans des conditions pratiques.

Un deuxième choix se présente dans le mode de comparaison. La comparaison de deux traitements, dans le domaine des effets variables, ne peut être qu'une "comparaison à des niveaux de risque définis". Le choix de ces niveaux détermine deux attitudes : la première vise à détecter une différence dont la connaissance est scientifiquement importante, la seconde indique celui des deux traitements qu'il faut adopter.

Le premier choix comme le second correspondent, dans l'ensemble, à deux attitudes : l'une vise préférentiellement à une quête de connaissance, l'autre à une décision immédiate.

Ce choix se présente pour tout problème de comparaison de traitements. Dans le domaine médical s'ajoute une implication éthique, qui interdit certaines options. Mais il reste toujours important de distinguer clairement les options possibles, même si certaines ne peuvent être adoptées.