

**QUALITE ETHIQUE ET QUALITE METHODOLOGIQUE
DES ESSAIS CONTROLES RANDOMISES :
ANALYSE A PARTIR DES ETUDES PUBLIEES DE 1989 A 1992
DANS QUATRE REVUES INTERNATIONALES**

**(New England Journal of Medicine, Lancet, British Medical Journal,
Journal of the American Medical Association)**

D. BERDEU

DEA 1993

Enfin la dernière critique, et non la moindre, faite aux E.C.R est liée aux questions éthiques qu'ils soulèvent (8, 9). Depuis la fin de la seconde guerre mondiale les problèmes éthiques posés par l'utilisation de sujets humains pour la recherche médicale ont fait l'objet de nombreux débats au sein de la communauté médicale. Le code de Nuremberg (1949) et la Déclaration d'Helsinki de l'Association Médicale Mondiale (1964) fixent à la communauté scientifique internationale des normes éthiques. Parallèlement à ces recommandations internationales, de nombreux pays ont établi des lois, des règlements et des lignes de conduite. Il semble qu'il existe aujourd'hui un consensus à peu près général pour deux questions de principe qui ont reçu une attention toute particulière : celle du consentement libre et éclairé et celle de l'établissement de comités indépendants pour juger de la validité du protocole sur le plan éthique. En fait, de nouveaux problèmes plus spécifiques apparaissent sur le plan de la pratique : comment juger du niveau d'information réellement assimilé par le patient et quelles doivent être les limites des pouvoirs accordés aux institutions de contrôle.

En effet, une nouvelle série d'objections éthiques, développées ces dernières années, repose sur l'idée que le consentement d'un malade, quel que soit son degré d'information, ne peut excuser certains risques inhérents à la technique de l'E.C.R comme celui de se voir attribuer un traitement moins performant (placebo par exemple) ou plus nocif (effets indésirables du traitement à l'étude). D'autre part il semble que les principes éthiques fondamentaux régissant l'expérimentation chez l'homme ne soient pas toujours respectés dans les E.C.R. Le principe éthique inhérent au commencement d'une expérimentation clinique a été clairement exprimé par Sir A. Bradford Hill, qui a introduit l'analyse statistique dans la recherche médicale moderne : "La caractéristique essentielle d'un essai contrôlé (comportant une répartition randomisée) est qu'il doit être possible sur le plan éthique de donner à tous les malades inclus dans l'essai n'importe lequel des traitements en cause. Le médecin accepte, en d'autres termes, de ne vraiment pas savoir si un traitement est meilleur ou pire, plus sûr ou plus dangereux" " Ce principe repose, à son tour, sur un principe encore plus fondamental qui gouverne l'éthique de la recherche sur l'être humain et qui a été énoncé par Claude Bernard : "Le principe de la moralité médicale et chirurgicale consiste à ne jamais effectuer chez l'Homme d'expérience qui lui soit, dans une quelconque mesure, préjudiciable, même si les résultats peuvent être très utiles à la science, c'est-à-dire à la santé des autres ". Il semble que l'E.C.R n'établisse pas de façon très claire la répartition des risques et des bénéfices entre les sujets étudiés et la société et que les impératifs méthodologiques se substituent fréquemment au jugement du médecin, jugement qui ne devrait pas avoir d'autre objectif que le bénéfice individuel de chaque patient.

Contrairement à la qualité méthodologique, la qualité éthique des E.C.R n'a pas encore fait l'objet d'une évaluation faute d'un instrument de mesure adéquat. Or il est évident que d'éventuelles réponses aux problèmes éthiques posés par les E.C.R ne pourront être apportées qu'une fois ces derniers bien caractérisés.

BUTS DE L'ETUDE

Notre travail a été réalisé dans le but d'établir la pertinence des critiques formulées à l'égard des essais contrôlés randomisés en ce qui concerne leurs qualités éthique et méthodologique. Notre objectif principal était de procéder à une évaluation conjointe des qualités éthique et méthodologique des E.C.R et à partir de ces constatations, d'essayer d'analyser les éventuelles relations pouvant exister entre ces deux standards.

Afin d'essayer de répondre à la question "Tout ce qui est scientifique est-il forcément éthique ?", nous avons été amené à proposer un instrument d'évaluation de la qualité éthique des E.C.R sous la forme d'une échelle de mesure. Nous proposons dans ce travail une première validation de cet outil.

MATERIEL ET METHODES

1. MATERIEL :

Nous avons sélectionné et identifié les essais contrôlés randomisés ("randomized controlled clinical trials") publiés dans 4 périodiques médicaux hebdomadaires internationaux de langue anglaise : le "New England Journal of Medicine," le "Lancet", le "British Medical Journal" et le "Journal of the American Medical Association", entre le 1er janvier 1989 et le 31 décembre 1992. La base de sondage était constituée par la liste de tous les essais contrôlés randomisés publiés entre le 1er janvier 1989 et le 31 décembre 1992 dans les rubriques "original articles" du N.E.J.M. et du LANCET, "original contributions" du J.A.M.A. et "papers" du B.M.J. Ont été exclus de cette étude les essais contrôlés randomisés publiés sous forme de lettre, de résumé ou de "briefs reports" (< 2 pages) ainsi que les essais dont l'objectif n'était pas principalement thérapeutique mais relevant plutôt de la connaissance physiopathologique. Cette première sélection des études a été réalisée par une recherche manuelle compte tenu des réserves émises par plusieurs auteurs sur la fiabilité de la recherche des essais contrôlés randomisés par interrogation de la base de données MEDLINE (10, 11). Une fois la base de sondage décrite de façon exhaustive, un échantillon représentatif de taille $n = 30$ a été constitué par tirage au sort. Nous avons réalisé un sondage élémentaire à un degré par tirage au sort à partir d'une table de nombres au hasard (12). Les 30 études constituant l'échantillon ont été évaluées sur le plan de leurs qualités méthodologique et éthique grâce à deux grilles de lecture : la grille proposée par Chalmers et collaborateurs et une grille d'évaluation éthique originale proposée par les auteurs.

2. METHODES :

2.1 La grille de qualité méthodologique des essais contrôlés randomisés :

Il s'agit d'une grille de 32 items proposée par Chalmers et collaborateurs en 1981 (3) pour évaluer la qualité méthodologique des essais contrôlés randomisés. La cotation de chacun de ces items aboutit à l'obtention d'un score quantitatif, reflet de la qualité globale de chaque ECR. La description des items ainsi que le détail de leur cotation sont donnés dans l'annexe I.

2.2 La grille d'évaluation de la qualité éthique :

Il s'agit d'une grille de 10 items dont le but est d'évaluer la qualité éthique des essais contrôlés randomisés. Comme pour la grille de Chalmers, la somme des scores obtenus à chaque item constitue un score total divisé par le score maximal attribuable. La description des items ainsi que le détail de leur cotation sont donnés dans l'annexe II.

2.3 Mode de passation :

Les articles sélectionnés ont été préparés par la technique de la photocopie différentielle telle qu'elle a été décrite par Chalmers. Ce procédé consiste à photocopier les publications en masquant sur l'original tous les éléments susceptibles d'influencer le jugement du coteur (nom du périodique,

noms des auteurs, ...). Une fois rendues anonymes, les publications étaient ensuite soumises à la lecture de deux cotuteurs indépendants (D.B. et R.A.) avec résolution des discordances en réunion.

2.4 Analyse statistique :

La vérification de la représentativité de l'échantillon a été effectuée par des tests de Chi-deux. Le degré de concordance inter-observateur a été évalué pour chaque grille par la statistique du coefficient de corrélation intra-classe (13). Les réponses à chacun des items ont été analysées de façon individuelle et globale pour chacune des échelles. Les comparaisons entre moyennes observées ont été effectuées au moyen d'analyse de la variance ou du test t de Student en fonction des situations. Les comparaisons entre scores globaux à chaque échelle ("échelle méthodologique" et "échelle éthique") ont fait l'objet d'analyses de corrélation.

RESULTATS

1. DESCRIPTION DE LA POPULATION :

Une double recherche informatisée et manuelle nous a permis d'identifier 528 essais contrôlés randomisés publiés entre le 1er janvier 1989 et le 31 décembre 1992 dans les quatre périodiques sélectionnés (N.E.J.M, Lancet, B.M.J. et J.A.M.A). Sur ces 528 études, 126 (23,8%) ont été rejetées en raison des critères d'exclusion suivants : 72 ECR à but physiopathologique (13,6%), 35 ECR non médicamenteux (6,6%) et 19 ECR publiés sous forme de "briefs reports" (3,6%).

Sur les 392 publications retenues, constituant la base de sondage, un échantillon de 30 études a été constitué par tirage au sort

2. VERIFICATION DE LA REPRESENTATIVITE DE L'ECHANTILLON :

Les tableaux 1 et 2 montrent que les structures de la population et de l'échantillon en fonction de l'origine de la publication ou de l'année de publication sont superposables, ce qui témoigne de la représentativité de l'échantillon et de la validité du sondage effectué.

| Périodique | Population | Echantillon | p |
|------------|------------|-------------|-----|
| N.E.J.M | 173 (44%) | 13 (43%) | N.S |
| Lancet | 114 (29%) | 11 (37%) | N.S |
| B.M.J | 66 (17%) | 4 (13%) | N.S |
| J.A.M.A | 39 (10%) | 2 (07%) | N.S |
| Total | 392 (100%) | 30 (100%) | |

Tableau 1: Répartition des études en fonction du périodique

| Année | Population | Echantillon | p |
|-------|------------|-------------|-----|
| 1989 | 94 (24%) | 6 (23%) | N.S |
| 1990 | 93 (24%) | 6 (20%) | N.S |
| 1991 | 91 (23%) | 10 (27%) | N.S |
| 1992 | 114 (29%) | 8 (30%) | N.S |
| Total | 392 (100%) | 30 (100%) | |

Tableau 2: Répartition des études en fonction de l'année de publication

Si la variable "année de publication" a été retenue en tant que variable de stratification pour l'analyse des résultats (en raison de l'importance de la notion d'évolution de la qualité méthodologique et de la qualité éthique dans le temps), par contre la variable "nom du périodique" n'a pas été retenue en tant que variable de stratification, ni pour le sondage, ni pour l'analyse des résultats, cela afin de prévenir toute discussion de nature polémique.

3. RESULTATS CONCERNANT LA QUALITE METHODOLOGIQUE :

3.1 Qualité méthodologique moyenne des ECR calculée sur l'échantillon :

Le score global moyen de qualité méthodologique obtenu en faisant la moyenne des 30 scores globaux de l'échantillon est égal à $0,65 \pm 0,13$. Le coefficient de corrélation intra-classe mesurant l'accord existant entre les 2 cotateurs est égal à 0,84 ce qui peut être considéré comme une bonne concordance.

3.2 Score moyen obtenu par item :

Les moyennes des scores obtenus à chaque item de la grille de Chalmers sont représentées sous forme d'histogrammes dans la figure 1. On constate un niveau de qualité insuffisant en ce qui concerne l'item 2.2 (description de la population éligible mais non sélectionnée) [$1,20 \pm 1,4$], l'item 2.5 dans ses 2 composantes (description du placebo) [$0,75 \pm 0,78$; $0,43 \pm 0,73$], l'item 2.6 (aveugle de la procédure de randomisation) [$4,4 \pm 6,2$], l'item 2.12 (test de l'efficacité de l'aveugle) [$0,10 \pm 0,39$], l'item 2.13 (vérification de la compliance) [$1,10 \pm 1,44$], l'item 2.14 (équivalent biologique) [$0,38 \pm 1,06$], l'item 3.2 (estimation a posteriori du risque β) [$0,83 \pm 0,93$], l'item 3.5 (prise en compte des sorties d'essai) [$1,33 \pm 1,95$], l'item 3.8 (situation en aveugle du statisticien) [$0,38 \pm 0,81$] et l'item 4.1 (dates de l'étude) [$0,88 \pm 1,03$].

Figure 1: Moyenne des scores obtenus à chacun des items de la grille de qualité méthodologique

3.3 Scores moyens globaux en fonction de l'année de publication :

Les scores moyens globaux de qualité méthodologique en fonction de l'année de publication sont représentés dans le tableau 3. L'analyse statistique ne permet pas de conclure à l'existence d'une différence significative globale entre les 4 années prises en considération dans ce travail : 1989, 1990, 1991 et 1992.

| | 1989 | 1990 | 1991 | 1992 |
|-------------------------------------|------|------|------|------|
| Score moyen à la grille de Chalmers | 0,57 | 0,68 | 0,68 | 0,66 |

Tableau 3: Scores moyens globaux de qualité méthodologique en fonction de l'année de publication

3.4 Scores moyens globaux en fonction du nombre de sujets inclus :

Si l'on classe les ECR en deux groupes, selon que leur nombre de sujets inclus est supérieur ou inférieur à 162 (valeur moyenne du nombre total de sujets inclus dans les ECR faisant partie de l'échantillon évalué), on constate (tableau 4) que les deux scores globaux moyens diffèrent significativement ($p < 0,02$) [$t_{28} \text{ ddl} = 2,71$].

$N < 162$ $N > 162$ p

| | | | |
|-------------------------------------|------|------|--------|
| Score moyen à la grille de Chalmers | 0,58 | 0,73 | < 0,02 |
|-------------------------------------|------|------|--------|

Tableau 4: Scores moyens globaux de qualité méthodologique en fonction du nombre de sujets inclus

4. RESULTATS CONCERNANT LA QUALITE ETHIQUE :

4.1 Qualité éthique moyenne des ECR calculée sur l'échantillon :

Le score global moyen de qualité éthique obtenu en faisant la moyenne des 30 scores globaux de l'échantillon est égal à $0,44 \pm 0,15$. Le coefficient de corrélation intra-classe est égal à 0,90, ce qui peut être considéré comme une très bonne concordance.

4.2 Score moyen obtenu par item :

La figure 2 présente les moyennes des scores obtenus à chaque item de la grille d'évaluation de la qualité éthique. On peut constater un niveau de qualité insuffisant en ce qui concerne l'item 3 (taux de refus de consentement) [$0,19 \pm 0,4$], l'item 4 (évaluation préalable du rapport bénéfique/risque) [$0,13 \pm 0,34$], l'item 5 (respect de la clause d'équivalence a priori) [$0,31 \pm 0,48$], l'item 6 (justification éthique de l'emploi ou du non emploi du placebo) [0 ± 0], l'item 9 (règles d'arrêt précoce mises en place pour raison éthique) [$0,38 \pm 0,52$] et l'item 10 (comité de surveillance de l'essai) [$0,25 \pm 0,45$].

Figure 2: Moyenne des scores obtenus à chacun des items de la grille de qualité éthique

4.3 Score moyen obtenu par regroupement d'items :

Les scores moyens obtenus pour chacune des trois classes d'items sont représentés sur la figure 3. Si la catégorie des critères éthiques réglementaires obtient un score très satisfaisant (0,91) et celle des critères accessoires un score assez moyen (0,45), par contre les critères éthiques théoriques fondamentaux sont négligés (0,22).

Figure 3: Représentation des scores en tenant compte du regroupement des critères éthiques en trois catégories

4.4 Scores moyens globaux en fonction de l'année de publication :

Les scores moyens globaux de qualité éthique en fonction de l'année de publication sont présentés dans le tableau 5. Comme pour la qualité méthodologique, il n'existe aucune différence significative entre les quatre années de publication.

| | 1989 | 1990 | 1991 | 1992 |
|--|------|------|------|------|
| Score moyen à la grille de qualité éthique | 0,4 | 0,44 | 0,42 | 0,49 |

Tableau 5: Scores moyens globaux de qualité éthique en fonction de l'année de publication

4.5 Scores moyens globaux en fonction du nombre de sujets inclus :

Contrairement à la qualité méthodologique, il n'existe pas de différence significative entre les scores moyens globaux de qualité éthique si l'on tient compte du nombre de sujets inclus (tableau 6).

N < 162N > 162p

Score moyen

à la grille de 0,48 0,39 NS
qualité éthique

Tableau 6: Scores moyens globaux de qualité éthique
en fonction du nombre de sujets inclus

5. RESULTATS CONCERNANT LES RELATIONS ENTRE LA QUALITE ETHIQUE ET LA QUALITE METHODOLOGIQUE :

Le calcul du coefficient de corrélation entre les 2 séries de scores globaux éthique et méthodologique obtenus sur l'échantillon ($R = -0,15$) montre qu'il n'existe pas de relation entre les deux standards de qualité.

DISCUSSION

Afin de mieux cerner les relations éventuelles entre qualités éthique et méthodologique des essais contrôlés randomisés, nous avons choisi de mesurer celles-ci à partir du seul matériel qui autorise la fiabilité et la validité de ce type d'évaluation, c'est-à-dire à partir de leur expression finalisée sous la forme de publications scientifiques. Plusieurs auteurs ont choisi ce type d'approche pour évaluer la qualité méthodologique des ECR (3, 4, 7, 14) et proposé différents instruments de mesure. Parmi ces derniers, la grille d'évaluation mise au point par Chalmers semble considérée comme la méthode de référence (3, 15). La principale critique que l'on peut faire à ces techniques d'évaluation basées sur l'examen des publications réside dans le fait que l'on est obligé de s'en tenir uniquement à ce qui est rapporté par les auteurs. Afin de souligner les limites de cette méthode, certains auteurs ont fait remarquer, à juste titre, que les consignes rédactionnelles des éditeurs de revues scientifiques imposent parfois certaines limitations de volume, obligeant ainsi les rédacteurs de comptes-rendus d'essais cliniques à alléger leur texte de certaines précisions notamment d'ordre méthodologique. On ne peut que regretter ce type de pratiques si l'on considère que les véritables juges des conclusions d'un ECR sont avant tout les lecteurs de la publication scientifique qui en fait état, c'est-à-dire l'ensemble des membres de la communauté médicale. Comme c'est le cas pour tout travail scientifique faisant l'objet d'une publication, on doit donner au lecteur d'un compte-rendu d'ECR les moyens de vérifier la validité et la pertinence des résultats qui lui sont présentés, en lui fournissant les moyens de se faire une opinion personnelle la plus objective possible.

Chalmers a essayé d'évaluer le taux d'items non rapportés par les auteurs dans leurs publications mais ayant cependant été observés dans la pratique. Il a procédé à une évaluation de la qualité méthodologique d'un échantillon d'ECR publiés et a adressé aux auteurs des publications sélectionnées un questionnaire dont le but était de vérifier si les critères omis dans la publication originale avaient été ou non effectivement remplis. Sur les deux tiers de questionnaires retournés, Chalmers a pu estimer ce taux à environ 58%, ce qui revient à dire que l'on a environ 1 chance sur 2 de considérer à tort qu'un critère n'a pas été rempli lorsqu'il n'est pas rapporté dans la publication.

Ce biais qui peut entraîner une "sous-estimation" du score de qualité méthodologique est contrebalancé par un autre biais de publication qui lui, entraîne au contraire un risque de surestimation de la qualité globale moyenne toutes publications scientifiques médicales confondues. En effet, notre choix délibéré de ne considérer que les quatre meilleurs périodiques internationaux généralistes est à l'origine d'un biais de sélection non négligeable. Ce biais de sélection est certainement responsable de la différence constatée entre nos résultats (score moyen global de qualité égal à $0,65 \pm 0,13$) et ceux, beaucoup plus bas, relevés dans la littérature comme en atteste la figure 5.

Figure 5: Scores moyens de qualité méthodologique des ECR relevés dans la littérature

En effet, ce score moyen global est de $0,42 \pm 0,16$ pour Reitman (7), $0,32 \pm 0,14$ pour Adams (16) et $0,54 \pm 0,14$ pour Nicolucci (17). Tous ces auteurs ont utilisé la grille de Chalmers pour juger de la qualité méthodologique sur des échantillons d'ECR, sans limiter leur recherche à un type de périodique donné. Nicolucci et collaborateurs apportent une très intéressante contribution à l'évaluation méthodologique des essais thérapeutiques contrôlés en étudiant l'ensemble des ECR publiés de 1970 à 1987 dans une seule affection, le cancer du poumon. Quelques résultats de cette étude méritent d'être rappelés : les essais présentent de nombreuses failles méthodologiques (le critère de jugement n'est pas précisé pour 27% des essais, le calcul du nombre de sujets nécessaires figure dans 6% des publications, une analyse statistique correcte ne serait réalisée que pour 17% des essais, ...) et aucune amélioration méthodologique n'est observée au cours du temps. Cette dernière constatation est corroborée par nos résultats, même si notre échantillon était trop réduit pour permettre d'objectiver une éventuelle différence significative dans l'évolution de la qualité méthodologique des ECR sur une période aussi courte (4 ans). Cette absence d'évolution du niveau de qualité global des ECR est assez accablante et témoigne de la nécessité que s'opère une appropriation par la communauté médicale de la méthodologie de l'essai thérapeutique. Le fait que cette qualité soit proportionnelle aux sujets inclus est également préoccupant. En effet, cela dénote qu'il existe deux niveaux de qualité en fonction de la taille de l'essai et donc un certain manque de rigueur pour les essais à effectifs modérés.

Les résultats obtenus dans le registre de la qualité éthique prouvent que l'outil de mesure (grille de qualité) que nous proposons est valide comme en témoigne le coefficient de corrélation intra-classe (validité externe). Cette bonne concordance entre cotateurs était prévisible en raison du caractère relativement objectif des critères et du système de cotation retenu. Le score global moyen calculé sur l'échantillon ($0,44 \pm 0,15$) est assez décevant compte tenu du fait que ces essais sont publiés dans ce que l'on considère comme étant les meilleurs périodiques internationaux. Comme on pouvait s'y attendre, on constate que si la totalité des publications d'ECR font état de l'obtention préalable du consentement du patient, il y a cependant près d'une publication sur cinq (19%) qui ne mentionne pas l'accord préalable d'un comité d'éthique. L'analyse des scores obtenus en classant les critères en trois catégories est plus démonstrative que l'analyse par item. En effet, elle permet d'objectiver de façon plus nette le peu d'intérêt manifesté par les auteurs vis-à-vis des principes éthiques théoriques fondamentaux sur lesquels reposent toute expérimentation portant sur l'être humain. Ces principes sont peut être considérés comme implicites par certains auteurs qui dès lors considèrent qu'il est inutile

Nous avons également montré que la qualité éthique des ECR estimée à partir des publications n'est modifiée ni dans le temps ni en fonction de la taille de l'essai. Ces résultats sont cependant à considérer avec réserve et mériteraient d'être confirmés sur un échantillon de taille plus importante. Enfin, l'analyse de corrélation entre qualité éthique et méthodologique permet de démontrer qu'il n'existe aucune relation entre ces deux standards de qualité. Un ECR de qualité scientifique très satisfaisante peut s'avérer très insuffisant en ce qui concerne les exigences d'ordre éthique.

Ces résultats permettent de prendre conscience de la nécessité d'améliorer la qualité éthique des essais contrôlés randomisés. Nous pensons que cette amélioration peut se faire essentiellement à trois niveaux : celui des commissions institutionnelles de contrôle (comités d'éthique), celui des comités de lecture des journaux médicaux et enfin au niveau de la formation médicale.

Les comités d'éthique ont, dans tous les pays où ils ont été mis en place, la possibilité de refuser leur approbation et leur caution aux protocoles de recherche clinique qui leurs sont soumis. Un des problèmes les plus débattus en ce qui concerne les missions des comités d'éthique est celui de savoir si ces derniers doivent faire la distinction entre la méthodologie et l'éthique proprement dite. Deux conceptions s'opposent quant au rôle des comités d'éthique. En effet, pour certains, il appartient aux comités d'éthique de juger non seulement de la qualité éthique mais aussi de la qualité méthodologique, partant du postulat qu'un protocole qui n'est pas scientifiquement correct n'est pas éthique puisque l'on peut faire courir des risques au malade sans aucun intérêt quant à la conclusion pour la collectivité. Pour May : "Une expérience pauvrement conçue, organisée ou planifiée est non éthique par définition et devrait être rejetée. Il est de la responsabilité des comités d'éthique de s'assurer que la conception et le

plan expérimental de l'étude respectent les canons de la méthode scientifique et cela d'autant plus que les individus participant à l'essai ne sont pas susceptibles d'en retirer un bénéfice individuel" (18). Les tenants de cette conception des attributions des comités d'éthique s'étonnent que des statisticiens ou des méthodologistes ne soient pas obligatoirement représentés dans certains pays au sein de ces institutions. D'autres, au contraire, suggèrent que le rôle des comités d'éthique se borne aux préoccupations de protection des personnes en s'assurant que les sujets ont reçu une information honnête, compréhensible et la plus complète possible avant de donner leur consentement.

Les comités de lecture peuvent, grâce à leur rôle de "verrou éditorial" contribuer à améliorer la qualité éthique des ECR en plaçant un peu plus haut la barre de leurs exigences dans ce domaine. En effet, il est communément admis qu'une plus grande exigence éditoriale des "grandes revues" a contribué à améliorer la qualité méthodologique des publications depuis le début des années 80. Il serait souhaitable que les comités de lecture exigent des auteurs de publications des garanties plus substantielles en ce qui concerne les précautions prises en faveur de la sécurité des patients inclus dans les essais et le respect des principes éthiques fondamentaux.

Enfin, l'amélioration de la qualité éthique des ECR passe par une formation à l'éthique médicale. Il serait souhaitable que toutes les structures de formation médicale contribuent à cet effort de formation éthique des médecins, tout spécialement en ce qui concerne les problèmes spécifiques posés par l'expérimentation sur l'homme.

CONCLUSION

Il est désormais admis que seule la réalisation d'un essai thérapeutique randomisé permet de fournir une évaluation non biaisée et convaincante de l'efficacité et de la tolérance d'un nouveau médicament. Bien évidemment, la mise en oeuvre de la méthode expérimentale doit respecter des standards de qualité au premier rang desquels figurent les exigences éthiques et méthodologiques qui sont les garants de la crédibilité des conclusions.

La réalisation d'essais de haute qualité permettrait d'assurer un meilleur impact des publications et, par voie de conséquence, la modification rapide des habitudes de prescription médicale pour les malades concernés.

L'amélioration de la qualité des E.C.R passe certainement par l'appropriation par la communauté médicale à tous ses niveaux, de la méthodologie et des principes éthiques propres aux essais sur l'homme. En facilitant la lecture critique, les grilles d'évaluation des qualités éthique et méthodologique contribuent à atteindre cet objectif.

Il est manifeste que l'amélioration de la qualité éthique d'une expérimentation peut, à partir d'un certain degré, remettre en question le choix d'un plan expérimental comportant par exemple une randomisation et un groupe contrôle. S'il semble assez clair que l'E.C.R est un instrument très précieux et que son application minutieuse et rigoureuse peut protéger d'innombrables patients et promettre des avantages à beaucoup d'autres, le débat relatif à son application convenable et à ses critères de rigueur doit en même temps se poursuivre. Des méthodologies autres que l'E.C.R doivent être développées, notamment pour des situations dans lesquelles les contrôles, la randomisation ou la technique en aveugle ne sont pas appropriés ou applicables.

REFERENCES BIBLIOGRAPHIQUES

- (1) -POCOCK S.J. : Clinical trials : a practical approach. New York, John Willey & Sons, 1983.
- (2) -SCHAPIRO S.H., LOUIS T.A. : Clinical trials : issues and approaches. New York, Marcel Dekker, 1983.
- (3) -CHALMERS T.C., SMITH H., BLACKBURN B., SILVERMAN B., SCHROEDER B., REITMAN D., AMBROZ A. : A method for assessing the quality of a randomized control trial. *Controlled Clinical Trials*, 1981, 2 : 31-49.
- (4) -DER-SIMONIAN R., CHARETTE L.J., BUCKNAM Mc PEEK B.A., MOSTELLER F : Reporting on methods in clinical trials. *The New England Journal of Medicine*, 1982, 306, 22 : 1332-1337.
- (5) -CHALMERS T.C., CELANO P., SACKS H.C., SMITH H. Jr. : Bias in treatment assignment in controlled clinical trials. *The New England Journal of Medicine*, 1983, 309 : 1358-1361.
- (6) -CHALMERS T.C. and discussants : How to turn off an experiment In : *Ethical Safeguards in Research on Humans*, ed. Cooper J.D. & Lley H.D. 1976. Washington D.C. : Interdisciplinary Communications Associated. : 119-143.
- (7) -REITMAN D., SACKS H.S., CHALMERS T.C. : Technical quality assessment of randomized control trials (RCTs). *Controlled Clinical Trials*, 1987, 8 : 282.
- (8) -HELLMAN S., HELLMAN D.S. : Of mice but not men. *Problems of the randomized clinical trial*. *The New England Journal of Medicine*, 1991, 324, 22 : 1585-1589.
- (9) -PASSAMANI E. : Clinical trials : are they ethical ? *The New England Journal of Medicine*, 1991, 324 : 1589-1592.
- (10) -ADAMS C.E., LEFEBVRE C., CHALMERS I. : Difficulty with MEDLINE searches for randomized controlled trials. *The Lancet*, 1992, 340 : 915-916.
- (11) -HEWITT P., DICKERSIN K., CHALMERS T.C. : More on MEDLINE searches. *Controlled Clinical Trials*, 1988, 9 : 85-87.
- (12) -RUMEAU-ROUQUETTE C., BLONDEL B., KAMINSKI M., BREART G. : *Epidémiologie : Méthodes et pratique*. Flammarion ed. Paris, 1993.
- (13) -FERMANIAN J. : Mesure de l'accord entre deux juges : cas quantitatif. *Revue d'Epidémiologie et de Santé Publique*, 1984, 32 ; 408-413.
- (14) -POYNARD T. : Evaluation de la qualité méthodologique des essais thérapeutiques randomisés. *La Presse Médicale*, 1988, 17, 7 : 315-318.
- (15) -ANTCZAK A.A., TANG J., CHALMERS T.C. : Quality assessment of randomized control trials in dental research I. Methods. II. Results : Periodontal research. *Journal of Periodontal Research*, 1986, 21 : 305-321
- (16) -ADAMS M.E., Mc CALL N.T., GRAY D.T., ORZA M.J., CHALMERS T.C. : Economic Analysis in randomized control trials *Medical Care*, 1992, 30 : 231-243.
- (17) -NICOLUCCI A., GRILLI R., ALEXARIAN A. : Quality, evolution and clinical implications of randomized controlled trials on the treatment of lung cancer. *JAMA*, 1989, 265 : 2101-2107.

- (18) -MAY W.W. : The composition and function of ethical committees. *Journal of Medical Ethics*, 1975, 1 : 23-29.
- (19) -DE VAHL DAVIS V. : How informed is informed consent ? *Bulletin of Medical Ethics*, 1992, 3 : 13-18.
- (20) -BICKNELL J. : Consent and people with mental handicap. *British Medical Journal*, 1989, 299 : 1176-1177.
- (21) -HORTON R. : The context of consent. *The Lancet*, 1994, 344 : 211-212.
- (22) -TUFFS A. : Ethics of clinical experiment questioned. *The Lancet*, 1992, 340 : 782.
- (23) -SIMES R.J. , TATTERSALL M.H.N., COATES A.S., RAGHAVAN D., SOLOMON H.J., SMARTT H. : Randomized comparison of procedures for obtaining informed consent in clinical trials for cancer. *British Medical Journal*, 1986, 293 : 1065-1068.
- (24) -MILLER L.L. : Pitfalls in the drug approval process: dose-effect, experimental design, and risk-benefit issues. *Drug Information Journal*, 1992, 26, 251-260.
- (25) -SCHAFER A. : The ethics of the randomized clinical trial. *The New England Journal of Medicine*, 1982, 307 : 719-724.
- (26) -SHAW L.W., CHALMERS T.C. : Ethics in cooperative clinical trials. *Annals of the New York Academy of Science*, 1970, 169 : 487-495.
- (27) -OH V.M.S : The placebo effect: can we use it better? *British Medical Journal*, 1994, 309 : 69-70.
- (28) -ROTHMAN K.J., MICHELS K.B. : The continuing unethical use of placebo controls. *The New England Journal of Medicine*, 1994, 331 : 394-398.
- (29) -SCHMUCKER D.L., VESELL E.S. : Underrepresentation of women in clinical drug trials. *Clinical Pharmacology & Therapeutics*, 1993, 54 : 11-15.
- (30) -HORTON R. : Trials of women. *The Lancet*, 1994, 343 : 745-746
- (31) -CASCHETTA M.B., CHARKIN W., McGOVERN T. : FDA policy on women in drug trials. *The New England Journal of Medicine*, 1993, 331 : 1815.
- (32) -MERKATZ R.B., TEMPLE R., SOBEL S., FEIDEN K., KESSLER D.A. and the Working Group on Women in Clinical Trials. : Women in clinical trials of new drugs. A change in Food and Drug Administration policy. *The New England Journal of Medicine*, 1993, 329 : 292-296.
- (33) -BENNETT J.C.: Inclusion of women in clinical trials. policies for population subgroups. *The New England Journal of Medicine*, 1993, 329 : 288-291.
- (34) -SIMON R. : A decade of progress in statistical methodology for clinical trials. *Statistics in Medicine*, 1991, 10 : 1789-1817.
- (35) -POCOCK S.J. : When to stop a clinical trial. *British Medical Journal*, 1992, 305 : 235-240.
- (36) -McNAMEE D. : Stopping trials early. *The Lancet*, 1994, 344 : 327.
- (37) -KROL W.F. : Closing down the study. *Controlled Clinical Trials*, 1983, 4 : 505-512.

ANNEXE I

GRILLE D'ÉVALUATION DE LA QUALITÉ MÉTHODOLOGIQUE DES ESSAIS CONTRÔLÉS RANDOMISÉS

(Traduction et adaptation française de la grille de T.C. CHALMERS)
New England Journal of Medicine, 1993, 329 : 288-292.

- 2.1 : Description de la population sélectionnée
- 2.2 : Description de la population éligible mais non sélectionnée
- 2.3 : Retraits après randomisation
- 2.4 : Description des traitements
- 2.5 : Description du placebo
- 2.6 : Randomisation à l'aveugle
- 2.7 : L'aveugle pour les patients
- 2.8 : L'aveugle pour les médecins vis-à-vis du traitement
- 2.9 : L'aveugle pour les médecins et les patients vis-à-vis des résultats
- 2.10 : Nombre de sujets nécessaires
- 2.11 : Distribution des variables avant traitement
- 2.12 : Test de l'efficacité de l'aveugle
- 2.13 : Compliance
- 2.14 : Equivalent biologique
- 3.1 : Signification statistique des critères de jugement principaux
- 3.2 : Estimation a posteriori du risque β en cas de différence non significative
- 3.3 : Inférence statistique
- 3.4 : Analyse statistique appropriée
- 3.5 : Prise en compte des sorties d'essai
- 3.6 : Discussion statistique des événements indésirables
- 3.7 : Analyses rétrospectives
- 3.8 : Situation en aveugle du statisticien
- 3.9 : Problème des analyses multiples
- 4.1 : Dates de l'étude
- 4.2 : Résultats de l'analyse des données avant randomisation
- 4.3 : Tabulation des événements employés comme critères de jugement
- 4.4 : Délais d'apparition des événements dans les courbes de survie

La cotation de chacun des items décrits ci-dessus aboutit à l'obtention d'un score quantitatif, reflet de la qualité globale de chaque essai contrôlé randomisé. Le score total, obtenu en faisant la somme des scores de tous les items, est ensuite divisé par le score maximal attribuable.

ANNEXE I I

GRILLE D'ÉVALUATION DE LA QUALITÉ ÉTHIQUE DES ESSAIS CONTRÔLES RANDOMISÉS

La grille d'évaluation de la qualité éthique des E.C.R. que nous proposons se compose de 10 items. Chacun de ces items entre pour le même poids dans le score total de la grille. A l'exception de deux items (items N° 8 et 9) qui sont susceptibles d'être "non attribuables", les réponses sont de type binaire : "oui" ou "non". Les 10 critères constituant cette grille sont indépendants. En majeure partie inspirés des recommandations des codes d'éthique internationaux (Code de Nuremberg, Déclarations d'Helsinki et de Tokyo), ces critères recouvrent l'ensemble des standards éthiques que l'on est en droit d'attendre de toute expérimentation sur l'être humain.

1. Recueil des consentements
2. Soumission du protocole à un Comité d'éthique
3. Evaluation préalable du rapport bénéfice/risque individuel et collectif
4. Respect de la clause d'équivalence a priori
5. Taux de refus de consentements
6. Justification éthique de l'emploi ou du non emploi du placebo
7. Critères d'exclusion abusifs ou non justifiés
8. Existence de règles d'arrêt précoce
9. Justification de la mise en place de règles d'arrêt précoce pour des raisons éthiques
10. Existence d'un Comité indépendant de surveillance de l'essai

Critères 1 et 2 : critères éthiques réglementaires ou légaux

Critères 3 et 4 : critères éthiques théoriques fondamentaux

Critères 5, 6, 7, 8, 9 et 10 : critères éthiques accessoires

ANNEXE III

DESCRIPTION DE LA METHODE D'EVALUATION DE LA QUALITE ETHIQUE DES ESSAIS CONTROLES RANDOMISES

INTRODUCTION :

La grille d'évaluation de la qualité éthique des E.C.R que nous proposons se compose de 10 items. Chacun de ces items entre pour le même poids dans le score total de la grille. A l'exception de deux items (items N° 8 et 9) qui sont susceptibles d'être "non attribuables", les réponses sont de type binaire : "oui" ou "non". Les 10 critères constituant cette grille sont indépendants. En majeure partie inspirés des recommandations des codes d'éthique internationaux (Code de Nuremberg, Déclarations d'Helsinki et de Tokyo), ces critères recouvrent l'ensemble des standards éthiques que l'on est en droit d'attendre de toute expérimentation sur l'être humain.

DESCRIPTION DES ITEMS :

1. Mention du recueil des consentements éclairés :

Ce critère est accepté presque à l'unanimité comme étant l'une des prémisses de l'éthique en recherche sur l'être humain, les discussions se concentrant sur des questions secondaires comme l'importance des révélations nécessaires (19), la capacité des sujets à comprendre ou à donner un consentement véritable, le bien-fondé des exceptions à la règle du consentement éclairé, comme dans le cas des enfants ou des malades mentaux (20). Il s'agit d'autre part d'un prérequis très souvent obligatoire sur le plan légal ou réglementaire, cette exigence ayant en effet été intégrée dans de nombreux textes régissant l'expérimentation sur l'homme. Il serait souhaitable que figurent dans le compte rendu de l'étude une reproduction du formulaire d'information ou du moins des précisions quant à l'exhaustivité de l'information qui a été délivrée au patient. En effet, plusieurs études ont montré l'existence de négligences dans l'information apportée au patient en ce qui concerne notamment la randomisation, l'existence d'un placebo ou l'évaluation des risques encourus.

2. Mention de la soumission du protocole à un comité d'éthique :

Des commissions institutionnelles de contrôle ou des comités d'éthique ont été créés dans un nombre croissant de pays. Même si les points de vue divergent parfois sur leur rôle, leur attribution exacte et leur composition, tout le monde s'accorde sur le fait que ces comités indépendants sont une garantie indispensable pour juger de la validité du protocole sur le plan éthique. Comme le critère précédent, la soumission du protocole à un comité d'éthique est souvent considérée comme un prérequis réglementaire. Le fait de ne pas se soumettre à cette exigence prive les chercheurs d'une caution éthique et risque de les exposer tôt ou tard à des critiques de la part de leurs pairs comme cela vient d'être récemment le cas (21, 22).

3. Mention du taux de refus de consentement :

Les taux de refus de consentement ainsi que les raisons qui motivent le refus de consentement des patients sollicités pour participer à un ECR sont intéressantes à connaître. Très peu de travaux se sont penchés sur ce problème, certainement parce qu'il est délicat de persuader des individus de répondre à des questions une fois qu'ils ont manifesté leur intention de ne pas participer à l'étude. Par contre on sait que la façon d'informer les malades et de leur demander leur consentement est liée au taux d'acceptation ou de refus (23). D'autre part si le taux de refus devient important et si les raisons du refus sont liées d'une façon ou une autre au critère de jugement, il peut en résulter un biais qui hypothèque la validité des résultats.

4. Mention de l'évaluation préalable du rapport bénéfice/risque individuel et collectif :

S'il est inconcevable que l'on puisse exiger des investigateurs d'un essai qu'ils n'exposent leurs patients à aucun risque (on sait bien qu'il existe toujours une part de risque dans les soins médicaux), de manière réaliste et pragmatique, le devoir éthique d'un médecin en tant que clinicien investigateur, est de répondre à la question : quel degré de risque additionnel est-il justifié de faire subir à des malades ?

Avant de commencer une étude sur l'homme, les chercheurs doivent établir une sorte d'équation montrant les rapports entre les risques entraînés par le programme de recherche et les avantages escomptés (24). L'évaluation de ce rapport bénéfice/risque à la fois sur le plan individuel et sur le plan collectif doit faire également partie de l'information fournie au patient lorsque l'on sollicite son consentement. Les débats se poursuivent quant à la façon de répartir ces risques et ces bénéfices entre les sujets étudiés et la société. En effet, Il semble que l'essai contrôlé randomisé tende à privilégier l'éthique collective au détriment de l'éthique individuelle et que les impératifs méthodologiques se substituent au jugement du médecin, jugement qui ne peut avoir d'autre objectif que le bénéfice individuel de chaque patient. Mais toute la difficulté réside dans le fait qu'il n'existe pas de distinction conceptuelle précise entre éthique individuelle et éthique collective.

5. Mention du respect de la clause d'équivalence a priori (Hypothèse H_0 crédible)

Tout essai comparatif ne peut être qualifié d'éthique que s'il satisfait à la clause d'équivalence a priori des traitements comparés. En effet, afin de concilier les deux impératifs que sont d'une part le soin des malades et d'autre part l'évaluation objective des traitements, le médecin investigateur doit se trouver dans une situation d'incertitude quant aux mérites respectifs des traitements comparés. Si le médecin pense que les traitements ne sont pas équivalents, sa conscience et son éthique de soignant lui commandent d'administrer à ses patients le traitement qu'il considère comme étant le plus performant. L'équivalence a priori des traitements, autrement dit la crédibilité de l'hypothèse d'égalité entre les traitements comparés (hypothèse nulle) au moment de débiter l'essai est une condition éthique indispensable et préalable à toute recherche (25). Dans les essais comprenant plusieurs bras, l'équivalence a priori doit exister entre tous les bras de l'essai, sans quoi le protocole doit être modifié pour exclure le ou les traitement(s) moins performant(s). Certains auteurs vont même plus loin en écrivant qu'un clinicien qui "sait, ou a de bonnes raisons de croire" qu'un bras de l'essai est préférable, devrait renoncer à participer à l'étude (26). D'autre part cette situation d'équivalence a priori est la seule qui justifie sur le plan éthique la répartition des traitements par tirage au sort (randomisation), seul mode d'allocation qui autorise ensuite le calcul des probabilités et l'analyse statistique.

6. Justification éthique de l'emploi ou du non emploi du placebo

On s'inquiète de plus en plus de l'utilisation de placebos dans les ECR, même parmi les auteurs qui seraient prêts à accepter une comparaison avec une méthode classique de traitement. On a, en particulier, considéré qu'il était amoral de donner aux malades des comprimés ne contenant pas de médicaments actifs et en ce sens de leur refuser un traitement qui serait peut-être bénéfique. Par ailleurs, il faut comprendre que de nombreux symptômes subjectifs désagréables peuvent disparaître ou apparaître spontanément, simulant un effet thérapeutique ou au contraire indésirable. Beecher parlait déjà, il y a 35 ans, du puissant placebo, et depuis, de nombreuses autres investigations ont confirmé la réponse aux placebos, aussi bien dans un sens positif que négatif.

D'un point de vue éthique les placebos sont acceptables si leur emploi est limité au traitement de symptômes subjectifs et non à celui de maladies plus graves pour lesquelles il existe un traitement réellement efficace. Il faut très soigneusement réfléchir sur la décision d'inclure pour comparaison dans une expérimentation randomisée un groupe sous placebo ou un groupe traité de manière classique. Il faut prendre garde de ne pas abuser des expérimentations cliniques randomisées contre placebo (27,28) et la méthode ne devrait être appliquée qu'en cas de nécessité réelle et si l'on ne peut espérer de résultats satisfaisants de la part d'autres méthodes de contrôle.

7. Critères d'exclusion abusifs ou non justifiés

Jusqu'à présent, les préoccupations éthiques autour des essais thérapeutiques avaient pour finalité de protéger l'individu sollicité à renoncer à son mieux être actuel pour un mieux être futur et collectif. Cette protection était envisagée uniquement comme un garde-fou contre d'éventuels abus d'enrôlement des malades. L'accent était mis unilatéralement sur le risque que constituait pour l'individu son inclusion dans un protocole de recherche. Or il se trouve que les problèmes éthiques peuvent être considérés sous un autre angle, à savoir celui des critères d'exclusion. Ces critères d'exclusion qu'il serait préférable d'appeler "critères de non inclusion" doivent être définis avec précision dans le protocole. Ils permettent de sélectionner ceux qui se verront proposer la possibilité de participer à l'étude. Les sources

possibles d'exclusion sont nombreuses mais ne sont pas toujours justifiées sur le plan méthodologique. Les exclusions non justifiées peuvent poser un problème d'ordre éthique si on considère que la participation à un essai thérapeutique peut être perçue par le malade comme une chance de recevoir un traitement plus efficace ou mieux toléré que le traitement de référence (cette situation s'est présentée dans certains ECR menés chez des patients atteints de SIDA). Dans ce cas, la non-inclusion de certains malades dans des protocoles comparant un nouveau produit à un traitement de référence peut être vécue par ceux-ci comme une sorte de discrimination (cas des patients alcooliques, toxicomanes ou psychopathes), une interdiction au "droit à la chance". La sous-représentation des femmes dans de nombreux protocoles procède de l'exclusion abusive (29,30,31,32,33)

8. Le protocole prévoit-il des règles d'arrêt précoce ?

Une des difficultés éthiques propre aux essais contrôlés randomisés concerne les modalités d'arrêt de l'essai thérapeutique, que le bénéfice escompté existe (on souhaiterait généraliser dès que possible le nouveau traitement) ou n'existe pas (on souhaiterait interrompre l'essai). Cette préoccupation éthique d'arrêt précoce s'oppose à la conception "dogmatique" de l'analyse d'un ECR : dans sa forme la plus pure, ce "dogme" stipule, pour comparer les deux groupes thérapeutiques, la réalisation d'un test statistique au terme de l'inclusion du nombre de malades nécessaire pour mettre en évidence la différence escomptée aux risques d'erreur consentis. Une recherche biostatistique active a permis, ces quinze dernières années, d'élaborer des méthodes statistiques adaptées à la décision d'arrêt d'un essai thérapeutique (34, 35). Il faut souligner que la décision d'interrompre l'inclusion dans un essai thérapeutique ne saurait être prise par la seule application d'une méthode statistique, même si ses propriétés en sont correctes : la décision effective d'arrêt de l'essai doit résulter d'une réflexion approfondie, pour laquelle l'assistance d'un comité de surveillance peut être utile.

9. En cas d'utilisation de règles d'arrêt précoce, sont-elles mises en place pour des raisons éthiques ?

La raison essentielle d'instaurer des règles d'arrêt précoce devrait être d'ordre éthique. En effet, ces mesures sont avant tout destinées à éviter de soumettre des patients à un risque qui devient amoral si l'on est en droit de conclure à une signification sur le plan statistique ou si il s'avère que le rapport bénéfice/risque a de fortes chances d'être préjudiciable au patient (36).

10. Existence d'un Comité indépendant de surveillance de l'essai

L'existence d'un comité de surveillance de l'essai, totalement indépendant du promoteur et des investigateurs est une garantie éthique supplémentaire. Constitués d'experts, ces comités de surveillance ("steering committee") ont pour rôle de veiller au maintien de la qualité et de la pertinence de la recherche entreprise, à l'émergence de l'évidence de la supériorité thérapeutique du traitement étudié (ou a contrario du non bénéfice de ce traitement), ainsi qu'à la survenue de problèmes de tolérance majeurs. Réunis régulièrement au cours de l'essai, les membres de ces comités devront statuer sur la base de résultats rapportés par groupe de traitement pour finalement conseiller au comité directeur de l'étude soit d'interrompre prématurément l'essai (37), soit de le poursuivre (avec le cas échéant modification du protocole). De tels comités devraient exister pour tous les essais thérapeutiques indépendamment de leur taille.