JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

DANIEL SCHWARTZ

L'expérimentation en thérapeutique

Journal de la société statistique de Paris, tome 102 (1961), p. 16-25
<http://www.numdam.org/item?id=JSFS_1961__102__16_0>

© Société de statistique de Paris, 1961, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (http://publications-sfds.math.cnrs.fr/index.php/J-SFdS) 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/

VI

L'EXPÉRIMENTATION EN THÉRAPEUTIQUE

Connaître l'effet d'une thérapeutique est en médecine une nécessité majeure. L'information fournie par le laboratoire, ou l'expérimentation sur animal, ne peuvent apporter une réponse suffisante, et c'est finalement sur les malades eux-mêmes qu'il faut apprécier le traitement.

La comparaison de 2 groupes de sujets, soumis à des traitements A et B, est pour le statisticien un problème familier, celui sans doute à qui la méthode statistique doit son existence même, son prodigieux développement lié au perfectionnement du « *design of experiment* », et ses succès les plus nets. La solution peut être ainsi résumée :

— constitution de 2 groupes comparables par un procédé où intervient la répartition au hasard.

— comparaison des résultats dans les 2 groupes par un test de signification; comme les 2 groupes sont semblables à tous égards, sauf en ce qui concerne les traitements, toute différence significative peut être mise sur le compte de ces derniers, de sorte que c'est la causalité même qui peut être démontrée.

Cependant, le seul énoncé de cette méthode fait apparaître une grave difficulté : le tirage au sort, accepté de gaîté de cœur quand il s'agit de parcelles en culture ou de souris, choque d'emblée dès qu'il concerne l'homme. C'est pourquoi on doit d'abord se demander si cette exigence des statisticiens n'est pas trop stricte, et si on ne pourrait pas y renoncer.

Les médecins ont pendant longtemps éludé cette difficulté en admettant que les malades de 2 années successives, ou de 2 hôpitaux voisins, constituent des groupes comparables. Il a fallu un certain nombre de déboires pour mettre en évidence la fausseté de ces hypothèses.

Les maladies sont changeantes — tout particulièrement les maladies infectieuses; c'est cette variabilité qui, au moment de l'entrée en scène de l'auréomycine, fit croire à son efficacité dans le traitement de la pneumonie atypique; il fallut attendre 4 ans, pendant lesquels avaient été utilisés des centaines de kilos d'auréomycine, pour qu'un essai correct en démontrât la parfaite inutilité [4].

Le recrutement même des malades change d'une année à l'autre, d'un hôpital à l'autre. Nous pouvons citer le cas d'un grand centre anticancéreux où, de 1945 à 1950, la mortalité des cancers de la langue augmenta d'année en année d'une façon extraordinaire : c'est qu'un service spécialisé dans le traitement de ces cancers venait d'être installé, et attirait chaque année davantage les cas les plus graves. Quand une maladie peut être traitée par les radia-

tions ou par la chirurgie, l'introduction d'un nouvel équipement radiothérapique augmente la proportion de cas justiciables de la radiothérapie et diminue celle des cas opérables; l'arrivée d'un chirurgien réputé produit l'effet inverse. Le sens même de la modification n'est pas toujours prévisible. Le seul fait d'expérimenter un nouveau traitement peut d'ailleurs modifier les conditions de recrutement; c'est ainsi qu'en Grande-Bretagne l'introduction des anticoagulants dans le traitement de l'infarctus du myocarde provoqua, dans certains hôpitaux, un afflux de malades; le supplément consistait en cas limites, qui en temps normal n'auraient pas consulté, et de gravité moindre. De fait, il fut vérifié que l'ensemble des malades venus consulter à cette époque présentait en moyenne un pronostic meilleur que celui des années antérieures, et il n'est pas douteux qu'en jugeant l'effet du traitement par comparaison avec le passé on aurait commis une lourde erreur [1].

Peut-être pensera-t-on que de telles difficultés sont évitables si on constitue, d'après les dossiers des malades, des groupes cliniquement comparables. Mais dans la plupart des maladies les critères cliniques sont si imprécis, si inconstants d'une année à l'autre, ou d'un médecin à l'autre, qu'un tel dessein est manifestement utopique.

Ainsi, malgré tous les efforts pour constituer des groupes comparables par des moyens en quelque sorte humains, il a fallu se rendre à l'évidence : ces moyens laissent toujours subsister entre les groupes des différences, faibles peut-être, mais hélas du même ordre que celles qu'on escompte en général entre les traitements : le tirage au sort n'est pas une exigence superflue, mais bien une nécessité.

Cette vérité a maintenant été admise par de nombreux médecins, dans plusieurs pays, et notamment dans les pays anglo-saxons. Diverses expériences ont été conduites, avec une remarquable rigueur. Elles ont porté, non seulement sur des troubles bénins comme le rhume, l'insomnie, la toux, la douleur post-opératoire, mais aussi sur des maladies graves : l'infarctus du myocarde, le cancer. Elles ont utilisé des plans expérimentaux plus ou moins raffinés; des progrès en sont résultés. Ces développements ne datent certes que d'un petit nombre d'années, ils suffisent cependant pour affirmer qu'une ère nouvelle est née en thérapeutique, celle de l'expérimentation par opposé à l'observation [1, 2, 3].

C'est cette expérimentation que je voudrais vous dépeindre aujourd'hui; elle le mérite — car, si le plan expérimental est pour le statisticien un vieux problème, il prend, du jour où il concerne l'homme, un visage singulier. Je parlerai de trois aspects essentiels — qui sont trois difficultés : un aspect psychologique, un aspect matériel et un aspect moral.

*
* *

L'aspect psychologique concerne les phénomènes de suggestion, chez le malade et chez le médecin.

La suggestion chez le malade peut être ramenée à la notion générale d'effet « *placebo* »; le *placebo* (étymologiquement : pour faire plaisir) est un traitement absolument inactif administré au patient, à son insu, aux lieu et place d'un traitement actif. De telles médications ont depuis longtemps été administrées par les médecins — quelquefois à leur insu — avec des résultats remarquables; mais depuis quelques années ils ont fait l'objet d'études rigoureuses. Dans une série d'expériences portant sur plus de 1 000 sujets, traités par des placebos imitant divers remèdes, 35 % environ des sujets ont ainsi été guéris de maux tels que : douleur post-opératoire grave, douleur angineuse, céphalée, mal de mer, nausée, anxiété, toux, rhume, etc...

Les placebos peuvent également entraîner des effets toxiques et des réactions secondaires : palpitations, diarrhées, urticaire.

Ces troubles ne sont aucunement imaginaires; ils existent réellement, et s'accompagnent parfois de variations objectives, par exemple des modifications sanguines, non perçues par le malade. Ils sont d'ordre psychosomatique.

De tels phénomènes obligent à distinguer, pour un traitement, son *action* et son *effet :* il peut être efficace, comme le placebo, sans être actif. Disons d'une autre manière que l'effet global d'un traitement est analysable en 2 parts : la première *spécifique*, liée à ses particularités pharmacodynamiques ou autres, la seconde *générale*, liée au seul fait du traitement, ou même à son simulacre. Celle-ci est sans doute loin d'être inintéressante, mais c'est la première qu'il faut connaître pour apprécier une thérapeutique.

On remarquera que la liste des affections sensibles au placebo ne mentionne pas seulement des troubles manifestement liés au psychisme, comme l'anxiété, ou la douleur, mais d'autres comme le rhume ou la toux; encore cette liste n'est-elle pas limitative, et chacun sait combien de maladies, dans leur évolution, comportent une part subjective. Même dans le cas du cancer par exemple, si on apprécie l'effet d'un traitement d'après la proportion de survivants après un nombre déterminé d'années, qui pourrait affirmer que l'effet placebo ne prolongera pas de quelques semaines la vie de certains malades, augmentant ainsi le taux de survie?

Précisons enfin que l'effet placebo ne s'attache pas seulement à des médications administrées par voie orale ou parentérale; il peut résulter de la radiothérapie, d'une intervention chirurgicale, ou même — la suite nous le montrera — des relations avec le médecin.

Un point très important pour la conduite de l'expérience est le suivant : puisqu'une certaine proportion seulement des sujets sont sensibles au placebo, est-il possible de les connaître à l'avance? On pourrait alors les éliminer de l'essai, et apprécier ainsi l'effet *spécifique* du traitement.

On a d'abord essayé de relier la sensibilité au placebo à différentes caractéristiques des sujets; on a effectivement trouvé une association à quelques-unes d'entre elles, mais dont aucune n'est simple : ni le sexe, ni l'âge, ni le niveau d'instruction ne paraissent jouer un rôle; on doit faire intervenir un examen psychologique détaillé, complété par des tests psychologiques; en outre la liaison ainsi observée est faible; de sorte que cette recherche explicative, malgré son intérêt indéniable, ne peut en aucun cas permettre de prédire la réaction d'un individu donné de façon à l'éliminer, le cas échéant, d'un essai thérapeutique.

On a ensuite recherché si la sensibilité au placebo est une propriété constante dans le temps. Quelques expériences faites à ce sujet apportent des conclusions contradictoires :

— Dans un premier travail [5], des sujets atteints de fréquentes migraines ont, au cours d'une longue période, été traités au moment de leurs diverses crises avec des produits actifs A, B, C, ou avec du placebo, dans un ordre aléatoire. On a compté ensuite pour chaque sujet le nombre de guérisons obtenues avec le placebo; ce nombre n'était pas distribué de façon homogène avec un seul mode, mais selon une courbe en U faisant apparaître 2 catégories de sujets : ceux qui n'avaient presque jamais été guéris, et ceux qui l'avaient été presque toujours. On a pu alors éliminer, dans l'analyse statistique, les malades du 2e groupe, et comparer les produits A, B, C, sur les sujets pratiquement insensibles au placebo; on a ainsi décelé entre ces produits des différences qui, sur le groupe total de sujets, n'apparaissaient pas.

— Par contre, dans un travail de même nature, portant sur l'administration de vomitifs, précédée de médications préventives ou de placebo [4], on n'a pu observer aucune constance dans l'effet placebo, celui-ci étant distribué parmi les sujets au cours d'essais

répétés, selon une loi binomiale : les sujets sensibles au moment de l'essai i n'avaient ni plus ni moins de chances que les autres d'être sensibles au moment de l'essai $(i + 1)$.

En vérité il est peu probable que la sensibilité au placebo caractérise un groupe de sujets bien déterminé, constant dans le temps et quelle que soit l'affection considérée; la réalité est sûrement plus complexe, et on ne peut envisager comme règle générale l'élimination pour tout essai des sujets sensibles au placebo. Il reste alors une autre méthode pour tenir compte de cet effet : elle consiste à administrer le traitement étudié à un groupe, et le placebo à un groupe témoin : c'est la méthode « anonyme » ou *blind* des anglo-saxons; elle permet de connaître la différence entre l'effet « global » et l'effet « général » du traitement, c'est-à-dire son effet « spécifique ».

Sans doute cette méthode n'est-elle pas toujours possible, pour des raisons morales ou matérielles; mais il faut savoir que l'effet placebo est maintenant un phénomène suffisamment démontré, et suffisamment général — dans le domaine des maladies comme dans celui des traitements — pour qu'on ne puisse plus l'ignorer dans un essai : si on n'a pu en tenir compte dans le plan expérimental, il faudra au moins s'en souvenir dans l'interprétation critique des résultats.

*
* *

Les phénomènes de suggestion ne concernent pas seulement le malade, mais aussi le médecin, et ce fait entraîne des difficultés de deux ordres :

a) L'évolution de la maladie d'abord risque de ne pas être appréciée objectivement. Il faut rappeler ici que ce jugement fait intervenir, soit des critères entièrement subjectifs — comme la variation d'intensité d'une douleur — soit un examen clinique apparemment objectif, mais qui laisse encore une large part à l'impression personnelle du médecin. Des examens paracliniques, comme la radiographie, n'échappent pas entièrement à cette règle, et une expérience célèbre, faite aux U.S.A. par un groupe de radiologistes et de phtisiologues réputés, a montré combien l'évolution de la tuberculose, jugée d'après des radiographies successives, prêtait matière à interprétations diverses [7]. La possibilité d'apprécier l'évolution d'une maladie d'après les examens de laboratoire entièrement objectifs est finalement beaucoup plus restreinte qu'on ne le croit. Dans ces conditions, le médecin qui préfère un des 2 traitements à comparer risque fort d'être influencé, et son jugement tendancieux enlèvera toute valeur à l'essai.

b) La deuxième difficulté, qui est beaucoup moins connue, est que l'opinion du médecin peut parfaitement influencer l'évolution même de la maladie, d'une manière objective, par suite de l'effet placebo.

La littérature cite quelques exemples démonstratifs [4] : deux groupes d'ulcéreux traités de la même façon, mais avec un succès très inégal, selon que le traitement était administré par des personnes persuadées ou non de son efficacité; une forte corrélation entre le degré de confiance de divers médecins dans une médication tranquillisante, et la proportion de succès obtenue chez leurs patients.

Un cas tout à fait extraordinaire est rapporté par Wolf [4] : il s'agit d'une malade, atteinte d'asthme chronique depuis de longues années, et habituée à essayer, sans succès d'ailleurs, toutes les médications; enfin parut un produit doué d'une réputation merveilleuse; son médecin se procura le médicament miracle auprès du laboratoire qui le fabriquait, et le fit essayer à sa malade : la crise d'asthme fut aussitôt guérie. Après plusieurs succès, le médecin décida de remplacer le médicament par du placebo : cette fois la crise ne céda pas. Le médecin alterna médicament et placebo, et obtint régulièrement un échec avec celui-ci,

un succès avec celui-là. Il courut alors apporter au laboratoire le résultat de sa rigoureuse expérience, pour y apprendre la révélation la plus stupéfiante : on ne lui avait jamais donné que du placebo... La différence entre les succès et les échecs ne traduisait donc que la confiance ou le scepticisme du médecin !

Ainsi la suggestion du médecin, comme celle du malade doivent être prises en considération dans le plan expérimental. La parfaite rigueur exige que le traitement actif et placebo soient administrés selon un code inconnu à la fois du malade et du médecin : c'est la méthode « doublement anonyme » ou *double blind* des anglo-saxons, qui a déjà été utilisée avec succès dans plusieurs essais thérapeutiques.

Il va de soi que, dans bien des cas, cette méthode ne peut être utilisée, le médecin devant obligatoirement connaître le traitement pour soigner correctement son malade; mais alors cette lacune devra être prise en considération dans l'interprétation critique des résultats.

*
* *

Après l'aspect psychologique, j'aborde l'aspect matériel, c'est-à-dire pratiquement celui du nombre de malades nécessaire. Celui-ci peut être calculé à l'avance par des formules dans le genre de la suivante, donnée ici sous sa forme la plus rigoureuse :

$$[1] \qquad 2\,n = \frac{(K_{2\alpha} + K_{2\beta})^2}{(\text{arc sin.}\sqrt{p'} - \text{arc sin.}\sqrt{p})^2}$$

avec :

. n = nombre de sujets dans chacun des 2 groupes.
. $K_{2\alpha}$ et $K_{2\beta}$ = écarts réduits correspondant, dans la loi normale, aux probabilités $2\,\alpha$ et $2\,\beta$ (α = risque de 1re espèce, ou de déclarer avantageux un traitement qui ne l'est pas, β = risque de 2e espèce, ou de rejeter un traitement avantageux).
. p' et p = proportions de guéris (ou de survivants...) dans les 2 groupes; ces proportions sont naturellement inconnues, mais on doit avoir une valeur approximative de p pour le traitement témoin, et se fixer la différence $(p' - p)$ qu'on prétend déceler.

Admettons que l'on consente :

$$\alpha = 5\,\% \qquad \beta = 5\,\%$$

et que pour le traitement classique $p = 50\,\%$.
Si on prétend déceler une différence de 10 % ($p' = 60\,\%$ pour le traitement à éprouver), la formule [1] donne alors pour le nombre de sujets nécessaire :

$$2\,n = 1\,100$$

Cette exigence de sujets est en vérité dramatique. Il suffit pour en être persuadé de la confronter par exemple aux effectifs annuels d'un des plus grands centres français de traitement du cancer :

— Nombre de nouveaux malades 7 000
— Nombre de malades atteintes d'un cancer du sein (cancer qui est le plus fréquent dans le centre). 400
— Parmi celles-ci, nombre de malades n'ayant pas reçu de traitement antérieur, et acceptant le traitement proposé par le centre 200

Il faut préciser ensuite que ces 200 cas sont cliniquement différents; leur gravité, leur traitement, obligent à distinguer au moins 4 classes, ce qui donne, pour un groupe homogène, une cinquantaine de cas. Si on ajoute enfin qu'un essai thérapeutique oblige à éliminer de tout groupe les sujets présentant certaines contre-indications, ou qu'on ne pourra suivre facilement, il ne restera plus finalement qu'*une trentaine de cas par an.* Le temps voulu pour en réunir 1 100, augmenté du délai de quelques années nécessaire pour apprécier l'effet du traitement, font que le résultat de l'essai sera certainement inutile au moment où on le connaîtra...

Sans doute existe-t-il des essais thérapeutiques moins restrictifs, portant par exemple sur tous les cancers du sein, ou même sur tous les cancers, mais l'exemple cité reste dans la limite des problèmes usuels : l'écart entre le nombre de sujets désiré et le nombre possible est alors si considérable qu'il devient indispensable de chercher des remèdes. Nous allons en examiner quelques-uns :

a) *Exigences provenant de la formule* [1]

On peut d'abord se montrer moins sévère pour les risques, et moins ambitieux pour la différence $(p' - p)$ à déceler.

Si on décide par exemple d'admettre $\beta = 50$ %, le nombre de sujets nécessaire deviendra 4 fois plus faible que ci-dessus; il faut seulement savoir qu'on aura 1 chance sur 2 de rejeter un traitement intéressant. Ceci n'est pas grave si on dispose de beaucoup de traitements à essayer, car en prenant 4 fois moins de sujets par essai, on pourra faire 4 fois plus d'essais. Si au contraire les traitements disponibles sont rares, il serait déraisonnable de consentir un risque de 2e espèce aussi élevé.

On peut également augmenter le risque de 1re espèce : peut-être n'est-il pas si grave de déclarer avantageux un traitement qui ne l'est pas, car on s'apercevra vite de l'erreur commise. On peut enfin se borner à chercher les traitements qui apporteront 20 % de guérisons en plus, au lieu de 10 %.

Ces divers choix ne sont en réalité pas indépendants : le nombre de sujets qui donne 50 chances sur 100 de rejeter un traitement peu avantageux $(p' - p = 10$ %) ne donne qu'un faible risque de rejeter un traitement très avantageux ($\beta < 0{,}001$ pour $p' - p = 30$ %). Le bilan pourrait donc être fait en fixant, par exemple, la valeur de α et la variation de β en fonction de $(p' - p)$.

Il semble qu'en définitive le médecin, ayant évalué le nombre de sujets disponibles, doive demander au statisticien l'espoir de succès correspondant, sous la forme qui vient d'être définie : il reste alors à décider si cette espérance justifie ou non l'entreprise de l'essai.

Rappelons que ce bilan devra souvent être examiné avec l'optique d'une stratégie d'ensemble, portant sur plusieurs essais consécutifs, en fonction du nombre des traitements à expérimenter.

b) *Modifications du plan expérimental*

L'amélioration du plan expérimental constitue naturellement la première tâche du statisticien. Le fait humain n'apporte pas nécessairement des complications dans ce domaine, et on a pu utiliser des plans expérimentaux plus ou moins raffinés, par exemple :

— la méthode des couples : le malade est utilisé comme son propre témoin, et reçoit successivement 2 traitements dans un ordre aléatoire. La méthode est généralisable à plusieurs traitements (méthode des blocs).

— les plans factoriels : dans le cas où le critère d'évolution est un critère quantitatif (durée de la maladie, régression d'une image radiologique) on a utilisé des plans factoriels permettant, soit de réduire la variance en contrôlant certains facteurs de variabilité, soit d'étudier simultanément plusieurs traitements, et de connaître ainsi l'effet propre de chacun et les interactions éventuelles. C'est ainsi qu'on a pu étudier simultanément, sur un groupe de malades atteints d'ulcère de l'estomac, l'effet propre du phénobarbital, de l'acide ascorbique, et de l'alitement, et l'interaction éventuelle de ces traitements entre eux [1].

— enfin les plans séquentiels revêtent une importance croissante. Ils présentent l'avantage d'une réduction du nombre de sujets (30 à 40 % en moyenne par rapport aux plans classiques) sans que l'échelonnement dans le temps soit forcément un inconvénient, du fait que les malades se présentent à l'hôpital l'un après l'autre. Ceci suppose naturellement que le résultat de la thérapeutique soit connu rapidement. S'il faut un long délai, le plan séquentiel entraînera certes une réduction du nombre de malades, mais aussi une augmentation de la durée de l'essai.

c) *Nombre de malades disponibles*

Si les considérations précédentes permettent de diminuer le nombre de sujets nécessaire, il reste à augmenter au maximum l'effectif de malades disponibles. Pour n'être que d'ordre purement pratique, ce remède ne doit pas être oublié, car c'est sans doute le plus intéressant.

Il va de soi qu'on peut arriver au résultat en multipliant le nombre des centres; si ceux d'un même pays ne suffisent pas (ce qui arrive car le nombre de centres bien équipés, et recevant assez de malades, est relativement restreint), on peut aller jusqu'à une collaboration internationale : c'est ainsi qu'un essai sur la valeur du traitement hormonal dans le rhumatisme articulaire aigu a été effectué en commun par la Grande-Bretagne, les U.S.A. et le Canada. Mais les problèmes de coordination deviennent vite si complexes, que l'essai risque de ne plus être réalisé correctement.

*
* *

Pour sérieuses que soient les difficultés psychologiques et matérielles, elles ne font que compliquer l'essai; bien plus grave est l'obstacle moral, qui peut l'interdire complètement.

En raison de sa gravité, le problème a fait l'objet d'infinies discussions, dont témoigne une abondante littérature. Des « règles d'or » ont été émises, dont le seul nombre suffit à prouver qu'aucune ne satisfait : certains n'admettent l'essai que pour les maladies bénignes, d'autres que s'il porte sur des volontaires; d'autres enfin exigent que l'essai apporte à la société un bénéfice indiscutable, et qu'on ne pourrait obtenir par aucune autre méthode...

Ce n'est pas le lieu ici de discuter amplement ce problème, dont nous ne voulons retenir que certains aspects d'ordre statistique.

Quand le médecin a pris en considération toutes les données qui caractérisent son patient, il choisit, en son âme et conscience, la thérapeutique qu'il suppose la plus appropriée. Peut-on lui demander d'inclure, dans son bilan, l'élément aléatoire du coup de pile ou face?

D'emblée nous dirons non si les 2 traitements diffèrent sur le plan du bénéfice attendu : nous excluons donc évidemment les essais portant sur 2 traitements connus comme inégaux, et ayant pour but d'évaluer leur différence.

Mais imaginons le cas où le médecin hésite : il connaît d'une part l'effet du traitement

classique; d'autre part le traitement nouveau lui semble plus prometteur, mais ce gain n'est pas certain; la différence pourrait bien être en sens inverse; il faut penser aussi à des réactions secondaires imprévues. Dans ce bilan, qui fait la part de l'inconnu, les 2 traitements offrent la même espérance de gain. Si la situation se présente ainsi, — et le cas est vraisemblable, réel, fréquent — en quoi la conduite du médecin diffère-t-elle, finalement, d'un tirage au sort? Ainsi l'objection morale disparaît.

Cependant la situation se complique à l'échelon d'un groupe de médecins. Si 2 traitements offrent des espoirs équivalents dans l'état actuel des connaissances, on ne trouvera pas hélas les médecins unanimes dans ce classement d'*ex-æquo*, mais seulement avec une égale fréquence des médecins convaincus de la supériorité de chacun des 2 traitements. Le problème d'éthique revient alors à savoir si cette conviction peut s'incliner devant une opinion collective. Dans la négative, on ne peut confier l'essai qu'à des médecins acceptant individuellement la situation d'incertitude.

Cependant une nouvelle difficulté se présentera ensuite : la comparabilité, initialement admise, peut disparaître en cours d'essai devant une information nouvelle; c'est ainsi que, dans le plan séquentiel, en suivant les résultats au jour le jour, on risque de voir pencher la balance en faveur d'un des traitements, sans que la signification soit atteinte; arrêter l'essai interdirait toute conclusion valable, mais sa poursuite n'est plus en accord avec la clause de comparabilité.

Cette difficulté se présente moins dans les plans non séquentiels, parce que le résultat n'est connu qu'en fin d'expérience. Il reste cependant qu'il y a intérêt à raccourcir le plus possible la durée de l'essai, si on veut éviter que des informations extérieures viennent corriger l'idée initialement admise de comparabilité. De toute manière la rapidité de l'essai a un autre avantage, c'est que les résultats obtenus peuvent être utilisés plus vite au profit de la généralité des malades.

*
* *

Ainsi la thérapeutique, du jour où elle est entrée dans l'ère de la rigueur expérimentale, a posé des problèmes majeurs. Nous ne prétendons pas, dans ce bref exposé, les avoir résolus; nous pensons au moins les avoir posés.

Il importe que les médecins, les statisticiens, les administrateurs chargés de l'organisation hospitalière, cherchent ensemble à les résoudre.

Daniel SCHWARTZ
Directeur de l'Unité de Recherches Statistiques
de l'Institut National d'Hygiène à l'Institut Gustave-Roussy

BIBLIOGRAPHIE GÉNÉRALE

[1] *Controlled clinical trials* (Ed. Blackwell — Oxford) 1960.
[2] *Les essais thérapeutiques cliniques* (Ed. Masson et Cie) 1960, D. SCHWARTZ, R. FLAMANT, J. LELLOUCH, C. ROUQUETTE.
[3] *Clinical évaluation of new drugs* (Ed. Waife et Shapiro) 1959.

RÉFÉRENCES PARTICULIÈRES

[4] WOLF S. — *The pharmacology of Placebos* (Pharmacology Reviews *11*, n° 4, 689-704, 1959).
[5] JELLINEK. — *Clinical tests on comparative effectiveness of analgesic drugs* (Biometrics *2*, n° 5, 87, 1946)
[6] WOLF S., DOERING C. R., CLARK M. L., HAGANS J. A. — *Chance distribution and the placebo « reactor »*. (J. Labor. and Clin. Med. *49*, n° 6, 837-841, 1957).
[7] YERUSHALMI et *al.* — (American Rev. Tub. *64*, 225, 1951).

DISCUSSION

M. F. Rosenfeld. — L'application simultanée de deux traitements soumis à l'étude, à un grand nombre de malades, au lieu d'une application progressive suivant un plan séquentiel, en vue d'obtenir rapidement des résultats concluants, ne me paraît pas supprimer les cas de conscience; on risque au contraire de procéder à une application massive d'un traitement qui s'avérera moins bon. Il semble que lorsque des problèmes d'éthique se posent, les tests séquentiels offrent l'avantage de limiter des risques de cet ordre, mais alors, comme M. Schwartz l'a bien expliqué, on peut en arriver à arrêter le test avant d'être en mesure de conclure de manière suffisamment précise. La solution à ce dilemme ne me paraît pas résider dans l'élimination de la progression dans l'expérimentation, mais doit pouvoir se trouver dans l'arsenal de la statistique où l'on devrait rechercher ou affiner une méthode pouvant s'appliquer aux cas étudiés.

M. D. Schwartz. — Lorsque le plan séquentiel est applicable, il faut, avant de l'adopter, faire un bilan entre ses avantages (notamment la réduction du nombre de sujets) et ses inconvénients (notamment la durée; celle-ci dépend du temps nécessaire pour juger l'effet du traitement). Quant au problème de conscience posé en cours d'essai par la tendance qui se dessine, je ne vois guère comment le supprimer.

M. E. Ventura. — Dans le test progressif de Wald, il y a une part arbitraire qui est quelque peu gênante dans le choix des risques de première et deuxième espèce, α et β; dans les problèmes industriels que je rencontre en recherche opérationnelle, je m'efforce toujours d'éviter un choix à priori et considère que ces valeurs α et β doivent être incorporées dans la valeur d'une fonction objective que l'on cherche à rendre optimale.

Certes, dans le problème de thérapeutique présenté par M. Schwartz, le choix de la fonction objective est beaucoup plus délicat qu'en matière industrielle où l'introduction de la notion de valeur est beaucoup plus immédiate. Toutefois, je pense qu'il doit être possible de se rapprocher de ce concept de fonction objectif en cherchant une échelle de valeur représentative des risques contre lesquels on entend se prémunir.

M. D. Schwartz. — Il ne paraît pas facile de trouver en thérapeutique une fonction permettant d'objectiver les risques.

M. le Président. — M. Schwartz nous a exposé à quelles difficultés se heurtaient les observations statistiques telles que celles qui le préoccupent et quelles étaient, notamment, les difficultés de toute estimation. J'ai écouté l'exposé de M. Schwartz avec un très grand intérêt et je me suis demandé si l'estimation n'était pas rendue difficile, ou peut-être même impossible par le fait que les faits statistiques qu'il se proposait d'observer échappaient à toute mesure.

J'ai été dès lors conduit, en pensant à un exemple pourtant fort éloigné, à me demander si une solution ne pourrait pas être trouvée dans la considération du prix de revient de la maladie ou, peut-être, du prix de revient sinon de sa guérison, du moins de la thérapeutique applicable.

On retomberait alors dans le cas, très étudié, de la probabilité d'un accident dont le prix est connu et dont on peut espérer se couvrir au moyen d'une assurance.

M. P. F. Denoix. — Le problème est malheureusement beaucoup plus complexe

que celui du prix de revient du traitement ou de la guérison. Le rôle du médecin n'est-il pas de s'efforcer de guérir le plus grand nombre de malades?

M. Chevry signale qu'il a été très intéressé par les expérimentations qui consistent à comparer les effets d'un médicament supposé efficace et d'un simple « placebo ». Mais il demande comment les expérimentateurs concilient de telles expériences avec les soucis d'éthique professionnelle sur lesquelles le conférencier a insisté.

M. D. Schwartz. — Au problème fondamental d'éthique, l'introduction de placebo n'ajoute qu'une difficulté purement formelle; si le médecin a admis que, pour une affection donnée, dans l'état actuel des connaissances, il est permis de constituer un groupe témoin non traité, le fait d'administrer à celui-ci du placebo n'aggrave pas, sur le fond, le problème de conscience.