



Les cafés de la statistique

"La statistique éclaire-t-elle
les questions de société" ?

soirée du 12 mai 2015

Thérapeutiques et risques : quelle transparence ?

synthèse des débats ^[*]

Au cours de ces dernières années, des accidents graves, très médiatisés, liés à des produits de santé, semblent avoir affecté la confiance que la population accordait jusqu'alors au secteur médical. Dans le même temps, les modalités d'appropriation et de transmission des savoirs chez les soignants, ainsi que les pratiques de soin, ont subi une profonde rupture, marquée, notamment, par la mise en place de divers outils où la méthodologie bio-statistique occupe une place singulière, mais aussi par les progrès de la connaissance du génome, la mise à disposition de "big data" (mégadonnées) et l'apparition de nouvelles classes de régulateurs, prometteurs de thérapies personnalisées.

La relation patient-soignant connaît aussi une période de mutation importante, ne fût-ce que par la facilité d'accès aux informations via Internet.

Ces évolutions majeures sont-elles à même d'améliorer la confiance des patients dans notre système de soin ?

Invités :

Pascal Maire, Pharmacien hospitalier aux Hospices civils de Lyon,

Jean-Christophe Thalabard, Professeur des Universités et praticien hospitalier, Université Paris-Descartes et Hôpital Cochin

Exposés introductifs :

Pascal Maire expose d'abord qu'il n'a pas de conflit d'intérêts, qu'il a sans doute un lien d'intérêts avec son employeur et qu'il reste sans doute sujet à des liens d'influence : d'où la nécessité de dire "d'où il parle". Etant pharmacien, il entend cependant se placer moins du côté de la pharmacody-

^[*] Tant l'exposé liminaire que le contenu des échanges sont structurés en quelques thèmes, sans suivre l'ordre chronologique. Par ailleurs, l'identité des intervenants n'était pas toujours connue et l'on a choisi de ne pas attribuer nominativement les propos. Au reste, ceux-ci ont été reconstitués à partir des notes du secrétariat sans reprendre leur formulation détaillée. Pour retracer le débat, les thèmes sont souvent introduits sous forme d'une question : ce qui vient ensuite n'est pas la seule réponse des invités, mais l'ensemble des contributions des participants.

namique que de celui du patient. Ce dernier passe au total peu de temps au contact des professionnels de la santé. Pour le reste, il vit dans un monde où circule une abondante information

P. Maire commence par cette petite histoire. Un voisin ami, à la campagne, lui demande avis (bien qu'il ne soit pas médecin prescripteur, ce voisin le suppose, en tant que pharmacien hospitalier, compétent). Ce voisin est traité depuis des années pour une affection cardiaque avec deux médicaments (un anticoagulant et un anti-hypertenseur) et astreint à un contrôle mensuel. Son cardiologue lui propose un nouvel anticoagulant, qui ne suppose pas de contrôle mensuel : ce serait plus confortable. En fin de consultation, alléguant de contraintes devenues réglementaires, le cardiologue lui demande de signer un formulaire reconnaissant être informé des conditions et risques. Le voisin est ensuite suivi par son généraliste. Après quelque temps, il éprouve quelques effets secondaires (tête lourde...). Il apparaît que le formulaire était en fait un consentement pour être inclus dans un protocole d'expérimentation du nouveau médicament. P. Maire est circonspect envers la stratégie des laboratoires mais aussi envers l'attitude des médecins "investigateurs". Ils sont incités et rétribués par les laboratoires pharmaceutiques pour inclure des patients dans une étude clinique ; mais ils sont aussi poussés eux-mêmes par la curiosité scientifique et le souci d'améliorer les pratiques. P. Maire en parle d'abord avec son adjoint. Celui-ci relève que s'il conseillait d'arrêter l'expérience et qu'un incident survienne, on pourrait l'en tenir responsable... Pour l'heure, on ne change donc rien. A quelque temps de là, une publication indique que le nouveau médicament a été autorisé parce qu'il « fait au moins aussi bien que ceux existant ». Dans ce cas, le risque est "équivalent" : moindre pour l'hémorragie cérébrale mais supérieur pour l'hémorragie digestive...

A quoi l'invité ajoute que, dans le cas évoqué là, le patient avait ajusté son traitement, fractionnant les prises et les déplaçant vers la soirée. Ainsi, souvent, les patients "bricolent". Cette question de **P'observance** n'est pas anecdotique : une étude estime que, en France, 50 % des patients modifient la prescription en fonction de leur perception des symptômes, de leurs craintes ou croyances, allant jusqu'à interrompre le traitement si les symptômes cessent. L'utilisation effective des médicaments (*effectiveness*) n'est pas toujours celle résultant des essais cliniques (*efficacy*) qui ont abouti à l'AMM¹.

Le second invité, Jean-Christophe Thalabard, commence lui aussi par déclinier ses éventuels conflits d'intérêts : absence de lien avec l'industrie du médicament, notamment. Sa spécialité (gynécologie-endocrinologie) lui fait prescrire des substances actives hormonales pour des situations qui ne concernent pas seulement des sujets malades mais qui touchent le plus grand nombre, en lien avec l'évolution de la société et la qualité de vie (contraception, ménopause, procréation médicalement assistée, etc.). A propos de la manipulation des stéroïdes sexuels, qui ont abouti à la contraception en 1960, il rappelle que leur histoire est intimement liée au développement de l'industrie chimique allemande dans la première moitié du 20^{ème} siècle, encouragé largement par le III^e Reich (premières Frauenklinik ; expérimentations humaines). Par la suite, ces expérimentations humaines se poursuivirent à Porto-Rico puis aux Etats-Unis, sur des populations particulières ; avant d'en arriver aux années 1960 où la diffusion de ces substances a été présentée comme une libération pour les femmes. Cette inscription de la contraception orale (la "pilule") dans les mouvements féministes peut expliquer qu'en France aucune étude épidémiologique prospective indépendante de suivi n'ait été faite entre 1960 et 2010, contrairement au Royaume-Uni ou aux pays scandinaves, sur d'éventuels effets à moyen et long termes. Ainsi, dans la loi de Santé Publique d'août 2004, les dispositions législatives ont ciblé essentiellement des aspects sociaux (tels que éviter les avortements d'adolescentes de 13 à 15 ans), mais en occultant les aspects de pharmacodynamique.

¹ AMM : autorisation de mise sur le marché.

J-C Thalabard considère aussi que les incidents qu'il y a eu, voire les scandales, ont généralement conduit, hormis de simplement changer symboliquement la dénomination des agences, à "bouger les lignes" : augmentant l'exigence quant aux conflits d'intérêts. La difficulté était alors de devoir se passer d'experts reconnus mais aux liens d'intérêt trop visibles et, de faire appel à des personnalités peut-être moins expertes ou tout du moins présentées comme telles par les tenants de l'ordre ancien.

Il évoque à son tour le problème d'une utilisation des médicaments différant de celle des essais thérapeutiques. Ajoutant que ceux-ci sont rarement conduits sur de longues durées : ainsi dans le cas de maladies communes (telles que : hypertension, diabète, maladies cardio-vasculaires, rhumatologie), cinq ans est souvent un maximum, tandis que les traitements dureront souvent plus de dix ans, voire la vie entière. On n'a donc aucune certitude *a priori* quant à leurs effets sur de longues périodes.

Le mot *transparence* est maintenant omniprésent, largement promu par les agences américaines (Food and Drug Administration, FDA) et européenne (European Medical Agency, EMA). La réalisation d'essais thérapeutiques nécessite leur conformité aux textes encadrant les recherches sur les personnes. Elle suppose leur autorisation préalable par les autorités de santé du ou des pays concernés. A l'initiative d'abord des grandes revues médicales, leur inscription est devenue obligatoire sur des registres d'essais accessibles à tous (<https://clinicaltrials.gov/>), ainsi que la mise à disposition rapide d'un rapport de fin d'essai, quel qu'en soit le résultat².

Enfin, le coût des grands essais – ainsi que les promesses réelles ou supposées de la génomique pour une médecine personnalisée – incite fortement à rendre possible de réexploiter les données conservées des essais thérapeutiques (utilisation secondaire). Des dispositions récentes prises par l'EMA, la FDA – ainsi que par certaines Agences qui, de fait, promeuvent de larges essais – vont dans le sens d'une obligation de partager les données. La mise en œuvre reste cependant complexe³ et se heurte au climat très compétitif du milieu.

Concernant les liens et conflits d'intérêt, il peut subsister un écart entre ce qui est déclaré et la réalité des faits, d'autant plus que le professionnel de santé n'a pas toujours une vision claire des financements dont il bénéficie. Une grande partie des déplacements, réunions de travail, séminaires, ateliers, congrès, formations continues des professionnels de santé restent tributaires de financements émanant de l'industrie pharmaceutique, soit directement, soit par l'intermédiaire de fondations ou associations "loi de 1901". En ce sens, il a été demandé directement aux financeurs de déclarer aux autorités publiques, nominativement, les montants versés. L'ensemble de ces informations sont désormais accessibles à tous via Internet⁴.

Débat :

Les interventions sont reclassées en cinq grands thèmes :

² La tentation est en effet grande de ne faire état que des études montrant des effets positifs, les autres étant occultées.

³ La concurrence est ici un frein évident. Mais, dans d'autres domaines où les enjeux sont moindres, réexploiter les données d'autrui est en soi souvent problématique car il faut que leur créateur ait bien décrit leur conception et les éléments de fait qui peuvent en avoir limité ou altéré la signification ou la fiabilité. [NDR]

⁴ <https://www.transparence.sante.gouv.fr/flow/main;jsessionid=761A41D7C905D28E0B740AAF5A2B7DDD.sunshine-public?execution=e1s1>, ou aussi :

www.transparence.sante.gouv.fr/flow/main;jsessionid=99FF927EBD0A40A6398C7615EFA3465A.sunshine-public?execution=e1s1

- A. comment comparer deux médicaments ?
- B. qu'est-ce que la médecine fondée sur les preuves ?
- C. comment est-on inclus dans une étude clinique ?
- D. aspects économiques
- E. stratégies et décision.

A. Comparer deux médicaments

Plusieurs participants demandent de préciser comment se fait la comparaison de deux médicaments. Ils questionnent en particulier le critère “non-inférieur” : pourquoi “non-inférieur” plutôt que “supérieur” ?

Les deux invités signalent que, jadis, on se contentait d'une comparaison sur l'animal : ajouter l'essai clinique est plus exigeant. En fait, les industriels du médicament avaient généralement de meilleurs méthodologues. Ils ont très tôt compris l'importance de critères simples, objectifs et clairs pour évaluer l'efficacité des traitements ; de méthodes statistiques appropriées pour planifier des études, choisir des critères et analyser les observations. Ce sont eux qui ont demandé que la loi encadre les essais. Ils entendaient ainsi se protéger d'études hospitalières insuffisamment fondées : dont les auteurs percevaient comme secondaires les questions de responsabilité juridique et étaient également peu persuadés de l'intérêt d'une méthode rigoureuse pour inclure des patients dans les essais de façon à éviter les biais dans l'évaluation des stratégies thérapeutiques..

Le test de non-infériorité. Les deux invités expliquent que ce test ne se pose que pour introduire un nouveau traitement là où en existe déjà un, dit “de référence” : reconnu efficace envers une pathologie donnée. Parfois, ce traitement de référence s'accompagne (même s'ils sont très rares) d'effets secondaires gênants voire dangereux ou de contraintes (horaires de prise du médicament, contrôles fréquents, etc.) ou encore, d'un coût élevé : toutes choses qui peuvent en limiter l'usage. Il est intéressant d'en introduire un autre plus acceptable, pourvu qu'il soit au moins aussi efficace. On est là en aval des essais “preuve de concept” ou de supériorité, destinés à mettre en évidence l'effet d'un nouveau médicament. Il ne s'agit pas simplement de délivrer une AMM pour un produit concurrent qui serait en tous points comparable à ce qui existe. Ce n'est pas un simple « me too » (moi aussi). Dès lors qu'il présente moins d'inconvénients on va l'accepter même s'il n'est pas plus efficace quant à l'effet principal recherché. On l'acceptera bien sûr s'il se trouve aussi supérieur à cet égard ; mais on se contente qu'il soit seulement équivalent à celui de référence. On demande donc :

- que les effets secondaires du nouveau traitement ou d'autres caractéristiques jugées intéressantes cliniquement ou économiquement soient, elles, significativement meilleures qu'avec celui déjà autorisé. L'exemple donné en début de soirée montre que certains effets peuvent être atténués et d'autres au contraire accrus ;
- que l'effet recherché soit au moins équivalent, c'est à dire qu'il ne soit pas significativement⁵ inférieur.

La discussion fait apparaître que la notion d'*équivalence clinique*⁶ n'est pas déterminée de façon précise. On doit choisir, à dire d'expert, une plage d'équivalence. Ainsi, si l'effet recherché est évalué en pourcentage de succès, on demande qu'il soit à peu près le même qu'avec le traitement de référence. “A

⁵ Lors de cette soirée, la discussion n'a pas explicité la différence conceptuelle importante entre deux acceptions du terme “significatif”. Il est utile que le compte rendu aille au-delà et clarifie : voir l'annexe 1 à la fin du compte rendu.

⁶ Lors de cette soirée, cette marge de tolérance (équivalence pour l'utilisateur : ici, équivalence clinique) a été citée sans qu'on s'y attarde. Deux justifications, de nature différente, peuvent en être données : voir annexe 2.

peu près” peut vouloir dire même un peu inférieur : si cette efficacité un peu moindre est compensée par le gain sur les autres critères. Par exemple, le pourcentage en question parmi les patients recevant le traitement nouveau lors de l’essai pourrait être accepté jusqu’à 10% au-dessous de celui observé pour celui de référence.⁷

Dans cette visée, le test de non-infériorité semble moins exigeant qu’un test de supériorité⁸ : il est en fait plus difficile, car ce qu’il faut, outre l’équivalence sur l’effet recherché, c’est vérifier que le nouveau traitement comporte significativement moins d’effets secondaires. Or, ceux-ci sont relativement rares. Ils sont aussi de natures diverses (éventuellement même de sens inverse, comme il a été signalé) : souvent chacune très rare même si prises ensemble elles le sont moins. Il faudrait alors un échantillon de plus grande taille pour vraiment démontrer leur atténuation. Or, on calibre plutôt les essais (nombre de sujets à inclure dans l’essai) pour vérifier simplement l’effet principal recherché. Il y a là une “perte de puissance”, dont on ne tient pas compte : en fait, le test de non-infériorité, comme son nom l’indique, ne s’intéresse qu’à vérifier l’équivalence (approchée) de l’effet principal et les améliorations quant aux effets secondaires sont à démontrer par ailleurs. De surcroît, notamment pour les traitements longs, dans les maladies chroniques, il faut idéalement des essais longs : et là, les abandons en cours de route deviennent trop fréquents, ce qui compromet la représentativité initiale.

Cette “faible puissance” à déceler et comparer les effets secondaires restait un obstacle pour leur prise en compte effective : l’absence d’effets observés durant l’essai clinique faisant conclure à tort à leur absence en vie réelle. On laissait au système de pharmacovigilance et aux industriels le soin de les repérer en aval de l’autorisation de mise sur le marché⁹.

Un autre participant, qui connaît la pratique des laboratoires, convient que l’étude fondée sur la non-infériorité a peu d’utilité. Mais il expose qu’elle s’inscrit dans une stratégie de recherche plus générale, qui cherche la supériorité de l’effet principal. Aucun laboratoire ne fait des études de non-infériorité isolées. Cela étant, il admet qu’il y a eu des dérives.

Notant cependant que la discussion, ici, est plutôt dubitative envers les essais de non-infériorité, un participant s’étonne que l’on critique ce critère et souligne qu’offrir un médicament équivalent à un prix moindre permet une concurrence bénéfique. Et en effet, au moyen d’un test de non-infériorité, un concurrent peut introduire un produit équivalent à un prix inférieur.

Un participant relève aussi – les deux invités le confirment – que les médicaments sont expérimentés généralement de façon indépendante : on n’a pas une vue complète des interactions entre deux médicaments, moins encore entre trois : or il est courant que les ordonnances dépassent largement ce chiffre¹⁰.

⁷ Pour illustrer la différence entre test de supériorité et test de non-infériorité, voir l’annexe 3 en fin de ce compte rendu. On y voit notamment comment intervient la marge d’équivalence dont il est question plus haut.

⁸ On est certes un peu moins exigeant quant à l’effet principal que pour le traitement de référence, mais celui-ci était déjà prouvé supérieur à l’absence de traitement : même légèrement en retrait, le traitement nouveau resterait efficace.

⁹ Les effets secondaires ne sont pas seulement constatés par des essais cliniques : ils sont aussi repérés par la pharmacovigilance, c’est à dire par les signalements de médecins traitants ou de pharmaciens au cours des années qui suivent la mise sur le marché. Ce n’est plus là un essai sur quelques dizaines de sujets, mais sur des milliers ou centaines de milliers ! Toutefois, outre qu’elle est tardive, on n’a pas là une très bonne mesure, car on est tributaire de l’initiative des patients et de leurs médecins. [NDR]

¹⁰ La question des comorbidités n’a été qu’effleurée : lorsqu’un patient souffre à la fois de deux pathologies ou davantage, elles peuvent modifier mutuellement leurs évolutions. Et, de surcroît, chacune donne lieu à des prescriptions différentes, qui s’ajoutent. Ou, parfois, se contraignent. [NDR]

Au total, hormis la vérification proprement statistique de l'efficacité, l'on se pose des questions

- sur la pertinence du critère requis pour comparer deux médicaments,
- sur l'aptitude des essais à saisir les effets d'un traitement prolongé,
- sur l'écart qui existe en pratique entre les conditions standardisées lors de l'essai et l'utilisation effective qui sera faite du médicament.

A quoi, sur un autre plan, s'ajoutent les enjeux commerciaux (voir plus loin : § D). Ils sont tels qu'il est tentant de gauchir les essais : par des protocoles insuffisants ou en écartant de l'échantillon des sujets mauvais répondeurs sans fournir d'explications claires. Ou encore, dans un processus de « pooling » de résultats (essais multicentriques, méta-analyses), en ne retenant que les études dont les résultats sont favorables et en occultant les autres¹¹. Les essais thérapeutiques sont le plus souvent à la charge des laboratoires qui proposent le nouveau médicament. Ceux-ci résistent à livrer leurs protocoles et leurs données de façon à permettre de les réanalyser. Les agences de certification font pression. Un argument qu'elles font valoir est que cette transparence est due aux sujets soumis à l'essai, en reconnaissance de leur participation.

Une illustration en sera donnée, en fin de soirée, avec l'évocation de l'homéopathie : mettant en lumière la complexité des liens entre les industriels et les expérimentateurs. Les sceptiques font valoir que les dilutions des remèdes homéopathiques sont telles que le plus souvent la substance supposée active ne se retrouve plus dans le granule ingéré. Un chercheur de renom (Jacques Benveniste) avait cependant produit des tests positifs, qu'il proposait d'expliquer par le fait que l'eau des solutions en aurait gardé la mémoire. Il ressort d'une double enquête menée par l'INSERM et par la revue *Nature*, qu'il aurait été abusé par ses propres expérimentateurs : lesquels, par complaisance, refaisaient le test jusqu'à avoir obtenu cinq positifs, les autres étant ignorés...

Tout cela bien sûr n'invalide pas les procédures de mise au point et validation, mais en marque les limites.

B. La médecine fondée sur les preuves.

La mise au point de médicaments et plus généralement de méthodes thérapeutiques, telle qu'évoquée ici, se fait selon un paradigme dit « *Evidence-Based Medicine* » (EBM), qu'on traduirait par « médecine fondée, ou basée, sur les preuves » (MBP) ou parfois aussi « médecine factuelle »¹². A partir des questions de la salle sur la comparaison entre médicaments, J-C Thalabard indique que ce paradigme a conduit à définir une échelle de *niveaux de preuve* : la preuve la plus solide, le “gold standard”, repose sur un essai randomisé¹³, en double aveugle^{14 15}. Ensuite, par force probante

¹¹ Les effets produits par une substance comportent toujours une part d'aléa : les patients sont différents, les pathologies ne sont pas homogènes, l'alimentation, l'hygiène et l'activité, voire même le moral du patient, sont autant d'éléments de variabilité. D'où une variabilité normale des résultats. La taille des échantillons permet de réduire sans la supprimer l'incertitude sur la mesure d'un effet moyen. Si l'on fait plusieurs tests, plusieurs essais thérapeutiques, il s'en trouvera peut-être certains qui concluront en sens inverse : et là, on peut sélectionner ceux qui, par hasard, se sont trouvés aller dans le sens souhaité. D'où, l'obligation signalée de tout publier. [NDR]

¹² Formalisé il y a plus de quarante ans par Archie Cochrane (UK) et Ian Chalmers (Canada), le paradigme EBM veut que la décision thérapeutique prenne en compte trois éléments : l'expérience clinique du praticien, les données de la recherche et les préférences du patient. Cette approche considère les seules données objectives en lien avec l'étiologie et la dynamique de la maladie, les moyens de diagnostic et les ressources thérapeutiques (médicaments, prothèses, protocoles opératoires, etc.) [NDR]

¹³ Comparaison des résultats obtenus pour deux échantillons (ou davantage) tirés au sort (*random* = hasard) et soumis à des traitements différents. Ou, entre un échantillon soumis au traitement étudié et un *échantillon-témoin* qui ne reçoit pas ce traitement. Les différences entre les deux échantillons sont dues au hasard de la répartition. Mais cette répartition ne doit

décroissante, on distingue : l'essai randomisé en simple aveugle (le patient ignore le traitement qu'il reçoit mais l'investigateur le sait), l'essai randomisé ouvert (où le traitement reçu n'est pas masqué), les études épidémiologiques (opérant sur des populations telles qu'elles s'offrent)¹⁶ et enfin les "histoires de chasse" (où chacun rapporte ce qu'il a vu ou ouï dire tel qu'il se le représente).

Ce paradigme, lent à s'installer, est devenu la culture de référence. Il est à la base de ce qu'on enseigne aux étudiants en médecine ainsi que des formations continues. Dans le principe de la *collaboration Cochrane* (<http://www.cochrane.org>), chaque professionnel de santé peut trouver les informations scientifiques adaptées à la situation singulière qu'il est amené à voir et, réciproquement, amener sa pierre à l'édifice en participant, avec l'expérience acquise auprès de ses patients, à des groupes de travail sur une question pratique bien ciblée. A une époque où il existait peu de substances thérapeutiques vraiment actives, Cochrane & Chalmers faisaient le pari généreux qu'il était possible de mettre à la disposition du plus grand nombre de médecins la liste des seules substances actives, après une évaluation critique des données de la science, constamment actualisées : ceci était d'autant plus important que le praticien pouvait avoir une pratique isolée ou une expérience limitée de pathologies rares¹⁷.

Est-ce que cela répond à ce dont les praticiens ont besoin ? Faut-il opposer le pharmacologue, qui serait plutôt sur le versant scientifique et le médecin, plus pragmatique, qui veut avant tout des recettes : apporter une solution au problème du patient qu'il traite. Ses choix sembleraient plus flous, plus "artistiques" (ne parle-t-on pas d'art médical) ?

rien aux choix délibérés ou inconscients de l'expérimentateur, ni au comportement des sujets entrant dans l'essai, ni non plus aux circonstances qui auraient formé des populations particulières que l'on comparerait sans autre précaution. Sans doute, les résultats peuvent fluctuer dans la mesure où la répartition s'est faite au hasard : un hasard pur, échappant aux contingences, mais hasard quand même. Un calcul de probabilités permet alors de quantifier cette fluctuation et d'appliquer une *test statistique* : on accepte ou non l'hypothèse - mise à l'épreuve par l'essai - que le traitement A est supérieur au traitement B. Cette décision est exposée à deux risques d'erreur : soit d'accepter à tort l'hypothèse alors qu'elle était fautive, soit de la rejeter alors qu'elle était vraie. Ces deux risques sont mesurés ; ils diminuent si l'on accroît la taille des échantillons. D'où un arbitrage à faire entre, d'une part, le coût et le temps de l'expérience et, d'autre part, de degré de certitude que l'on souhaite. Cette méthode d'essais comparatifs n'est pas propre à la recherche pharmacologique et on l'emploie dans d'autres domaines : contrôle de qualité dans l'industrie, recherche agronomique, etc.

¹⁴ Même si les échantillons (randomisés) sont constitués sans égard aux circonstances ni aux choix des investigateurs ni des patients, le fait que ceux-ci (patients ou équipe soignante) sachent quel est le traitement reçu peut modifier inconsciemment leur comportement ou leur appréciation des effets. Dans la méthode en *double aveugle*, personne – ni le patient ni ceux qui prescrivent et observent – ne sait quel traitement un patient donné a reçu. Dans le cas d'un échantillon-témoin ne recevant aucun traitement, il se voit administrer un produit de même apparence mais totalement neutre, appelé *placebo* : un simulacre de traitement. [NDR]

¹⁵ J-C Thalabard rapporte que cette méthode a même été essayée pour vérifier un éventuel effet de la "prière d'intention" : la publication de l'étude (Archives of Internal Medicine en 1999) a suscité un courrier des lecteurs tellement abondant qu'un numéro spécial de la même revue lui fut dédié neuf mois plus tard. {<http://archinte.jamanetwork.com/article.aspx?articleid=485161>}.

¹⁶ L'épidémiologie travaille sur des données d'observation, sans pouvoir contrôler les distorsions ou influences comme on peut le faire grâce à la randomisation. Elle garde une méthodologie aussi rigoureuse que possible dans le recueil des données et, au stade de l'analyse, s'efforce de prendre en compte les distorsions induites peut-être par la structure de la population (par âge, sexe, lieu, type d'activité, etc.). Le choix de sujets-contrôles, tenant lieu d'échantillon-témoin, est primordial et plus ou moins simple. Par exemple, on peut rechercher, pour chaque individu de la population d'intérêt, un "cas-témoin" : c'est à dire quelqu'un qui présente le plus possible les mêmes caractéristiques, les deux individus ne différant que par le fait d'être ou non affecté par la pathologie ou le symptôme étudié. [NDR]

¹⁷ Ceci reste-t-il bien actuel ? On sait que les médecins sont abonnés à des journaux qui les tiennent au courant des recherches. Et, avec Internet désormais, ils ont accès à un vaste corpus. De plus, en bien ou en mal, ils sont l'objet d'un démarchage des visiteurs médicaux. [NDR]

Ce mouvement initialement à vocation très professionnelle a progressivement été adopté par le milieu académique médical, dynamique difficile à remettre en cause. Simple au départ, le modèle (évaluer une procédure par rapport à une autre, de référence, avec deux groupes parallèles de patients aléatoirement assignés à l'une ou à l'autre des deux procédures) s'est complexifié. Toute une panoplie de méthodes a cherché à concilier des impératifs éthiques (exposer le moins longtemps possible des patients à des traitements toxiques ou inefficaces), économiques (développer le plus rapidement possible un produit actif), sanitaires (proposer des traitements au rapport coût/bénéfice optimal). Les méthodes génomiques et la promotion d'une médecine personnalisée ont rendu nécessaire de repenser l'essai thérapeutique traditionnel dans ses phases précoces : ainsi sont apparus récemment les *basket-trials* (le traitement est alloué à des patients atteints de maladies différentes mais porteurs d'une signature génomique tumorale particulière témoignant d'une possibilité d'action ciblée à un certain niveau cellulaire), les *umbrella-trials* (différentes molécules actives sur des cibles tumorales différentes testées chez des patients atteints d'un même processus morbide). Les essais peuvent devenir séquentiels (avec possibilité d'arrêt en cours d'étude si la supériorité est manifeste ou si, au contraire, elle n'est plus susceptible d'être démontrée), adaptatifs (lorsque les critères d'inclusion/exclusion, le rapport d'allocation entre les groupes, les doses administrées, voire les critères principaux, sont modifiés en cours d'étude). Une difficulté est alors de planifier les analyses intermédiaires, tout en préservant l'issue, la possibilité même de mener l'étude jusqu'à son terme prévu afin de conclure. Des essais, dits "N of one", ont également été proposés où les divers traitements sont administrés en séquences aléatoires répétées sur un même sujet.

Un participant demande comment une revue telle que *Prescrire* peut se faire une opinion sur les médicaments dont elle parle. Elle n'a sans doute pas les moyens de refaire tous les essais ! Comment procède-t-elle ? Quel crédit lui accorder ? Certes, est-il répondu, elle ne peut refaire les essais ; elle s'inscrit néanmoins dans cette logique MBP. Totalement indépendante de l'industrie pharmaceutique, la revue propose une analyse critique des études publiées dans la "littérature" au sens large – par des experts normalement indépendants.

Ce paradigme MBP est ainsi largement prévalent. Il est réclamé par les revues scientifiques et par les autorités sanitaires. Celles-ci ont endossé la demande des grandes revues : que les essais aient été enregistrés *a priori*, afin d'éviter que les auteurs ne sélectionnent *a posteriori* les résultats positifs, issus de pseudo-études pas vraiment organisées prospectivement¹⁸ selon les bonnes pratiques de la recherche clinique. Il y a aussi une discussion pour ouvrir (rendre accessibles) les données des essais (pas seulement donc la méthodologie, mais bien les résultats individuels) de sorte que l'on puisse en refaire l'analyse. Les laboratoires résistent ; et les agences font pression.

Tout cela est très exigeant, très coûteux, et suppose des compromis entre protocole scientifique idéal et démarche pragmatique. Avec le risque parfois que cette quête apparente de rigueur ne passe à côté du problème tel qu'il se pose en vie réelle.

Un participant relève que l'autorisation de mise sur le marché est parfois assortie de critères très politiques. Il souligne aussi (voir plus haut) que, si l'on sait à peu près les effets d'un médicament, on est moins sûr de ce que produisent deux médicaments ensemble ou davantage. A quoi s'ajoute l'écart important (déjà évoqué) entre les conditions de l'essai et celles du traitement (observance,

¹⁸ En recherche sur des populations réelles, on distingue les analyses *prospectives*, dont les méthodes sont définies avant que les observations ne soient faites, et les analyses *retrospectives*, où l'on traite des données déjà existantes sans donc en maîtriser les processus de production. [NDR]

etc.), mais aussi l'exposition parallèle à de nombreuses substances – médicamenteuses ou non – comme par exemple la consommation de compléments alimentaires.

Est soulignée aussi la très grande diversité des pathologies. Par exemple, la différence entre maladies aiguës et maladies chroniques¹⁹ : pour ces dernières, on ne dispose pas d'essais de longue durée : les effets à long terme sont repérés plus tard, par surveillance "post-AMM", avec toutes les incertitudes quant aux conditions d'observation et de signalement par les médecins.

Les deux invités s'attardent un instant sur cette question. Autrefois, en France du moins, la pharmacovigilance reposait sur les "remontées" spontanées d'incidents. On recevait ainsi des alertes : mais seulement en cas de troubles, d'effets secondaires indésirables plus ou moins graves. Mais on n'avait pas le nombre total de traitements auquel rapporter ces signalements : le dénominateur demeurait inconnu. Aujourd'hui, notamment après l'affaire du Mediator²⁰, on a cherché à favoriser cette remontée d'information en ouvrant la possibilité à tout un chacun – aux patients notamment – de signaler directement les effets indésirables sur un site internet. Or, à ce jour, ils ne l'ont pas beaucoup ; et, l'augmentation des signalements ne résoudrait guère la difficulté, le dénominateur n'étant pas mieux connu qu'auparavant. En outre, ce canal risque d'être "bruité" par des campagnes à l'encontre de certains médicaments.

Au passage, est relevé que les notices présentes dans chaque boîte de médicament énumèrent un nombre inquiétant d'effets secondaires éventuels : tout ce qui a pu être rapporté au cours de la période de prise du médicament, sans explication visible autre que celui-ci, mais sans certitude non plus qu'il en soit la cause... Certains participants suggèrent qu'il vaudrait mieux ne pas les lire !

Est aussi rapidement évoqué que les mobilisations massives de données (phénomène "BigData") permettrait d'exploiter des gisements d'information. Déjà, en France, une base de données reprend les prestations de l'assurance-maladie²¹. De nombreuses recherches cliniques s'appuient sur l'analyse de cette information. L'ouverture de ces données à d'autres utilisateurs est en discussion : le prin-

¹⁹ Lors d'un précédent café de la statistique (*L'avenir de l'assurance-maladie*, 9 octobre 2012) avait été signalée l'évolution considérable de la morbidité générale : les progrès de l'hygiène et de la médecine ont grandement diminué les maladies aiguës et passagères (létales ou non), tandis que le vieillissement de la population amène davantage de pathologies chroniques et de co-morbidités, c'est à dire la simultanéité de plusieurs affections et donc celle de leurs traitements respectifs. [NDR]

²⁰ Le Médiator a été autorisé en 1976, pour des traitements visant à réduire le taux de cholestérol sanguin, autorisation étendue ensuite pour contrôler le sucre sanguin, comme adjuvant dans le traitement du diabète de type 2. Son efficacité restait cependant limitée sur ces critères, tandis que le risque cardiovasculaire ne faisait pas parti des critères évalués. Ce n'est que secondairement que des doutes sont apparus sur son innocuité à cet égard. (Ceci peut illustrer ce qui est dit de la difficulté d'essais concernant les traitements de longue durée.) Or, ce traitement s'est trouvé largement prescrit en dehors de son indication : comme coupe-faim. Plus de 20 ans plus tard seulement, quelques études alarmistes, difficiles à mener, ont fini par faire émerger un risque suffisamment important pour que scandale généré conduise à le retirer en 2009. On estime que 4 à 5 millions de patients ont dû l'utiliser. Et il y aurait eu, selon diverses estimations, de 500 à 2 000 morts.

²¹ En France – ce dont la plupart des autres pays ne peuvent disposer – on a constitué une base de données couvrant les remboursements de tous les régimes d'assurance-maladie : le Sniiram (Système national inter-régimes d'assurance-maladie). Les données relatives aux hospitalisations y sont également intégrées. Ceci permet de suivre sur d'assez longues périodes les prescriptions dont a bénéficié un même malade. Sans doute ne connaît-on pas là les achats non présentés au remboursement ; et, l'on ne sait pas si ni comment les médicaments délivrés ont été consommés. De plus, on ne connaît pas les pathologies, qui ne peuvent qu'être déduites des prescriptions : de façon incertaine, donc. Malgré ces limitations on a une assez bonne vision. On sait également l'âge et le sexe des patients, leur localisation, qui étaient les prescripteurs et l'historique des soins en relation ou non avec la pathologie traitée (vaccinations, opérations, etc.) L'on connaît aussi le prix des soins et la part remboursée aux assurés. Cette base de données couvrant exhaustivement la population, on peut descendre à un grand détail d'analyse. [NDR]

cipe en semble acquis, mais les modalités restent à déterminer. Il pourrait en aller de même pour les données des essais thérapeutiques (comme cela a été signalé). Pour l'instant, l'accès aux données de santé est très contrôlé, soumis à l'avis d'un comité indépendant, au vu de la qualité du demandeur et de ses objectifs scientifiques ou médicaux.²²

La pharmacovigilance reste un domaine difficile. D'une part, on a des échelles de temps différentes : le temps de l'effet principal n'est pas celui des effets secondaires. D'autre part, le médicament n'est pas seulement un objet thérapeutique, mais aussi un objet économique et un objet social.

C. Comment est-on inclus dans une étude clinique ?

Une large part du débat porte sur cette question. La règle, de nos jours, est que les sujets ne soient inclus dans un essai qu'avec leur consentement. Et, que celui-ci repose sur une information suffisante. C'est ce qu'on appelle le "consentement informé" ou "éclairé"²³. Les comités appelés à se prononcer sur les projets de recherche s'en assurent et regardent comment est formulée l'information. Les médecins expérimentateurs ont parfois du mal à intégrer ces prescriptions, qui leur semblent mettre en doute leur manière d'informer. Il se peut aussi que le médecin expérimentateur fasse, même inconsciemment, pression pour que son patient accepte : soit parce qu'il y a un intérêt direct financier (rétribution selon le nombre de patients inclus), soit tout simplement parce qu'il s'intéresse à la recherche en cours. Autre question : en principe, l'on peut se retirer de l'essai à tout moment : mais le peut-on vraiment, en pratique ?²⁴ Si l'on a une information initiale, lors de l'inclusion, il n'y a souvent pas d'information des patients en cours d'essai²⁵.

Mais, que veut exactement dire *informer* ou *éclairer* ? Les comités de protection des personnes (CPP), instaurés par la loi, essaient de donner une place au patient : c'est de son point de vue qu'on peut dire s'il est suffisamment informé. D'autant qu'on a devant soi aujourd'hui des gens très informés (par la presse, par Internet). Informés ou inquiétés ! Et, il ne convient pas non plus de vouloir en dire trop : faut-il donner des informations non souhaitées ? Et ceci, outre le fait qu'elles soient ou non compréhensibles, selon la personne à qui on s'adresse.

Une participante fait état de son expérience : souvent, au départ, le protocole (les informations sur l'essai, ses buts et conditions) est en anglais. Parfois il est traduit in extenso et parfois pas du tout.

²² La perspective "BigData" va toutefois bien au-delà de ces données déjà constituées : il pourrait s'agir de tout ce qui est collecté par les "objets connectés" utilisés quotidiennement par un nombre croissant de personnes, qui contrôlent leur rythme cardiaque, leur exercice physique, etc. Cela suscite des convoitises ; mais on ne sait guère encore quelle serait la signification de telles données et de leur exploitation. [NDR]

²³ "consentement informé" est la francisation de *'informed consent'*. La loi Informatique & Libertés dit *consentement exprès*. Certains, de manière fautive, prononcent le 's' final ou, même, écrivent "express" : le terme *exprès* n'a rien à voir avec la rapidité (ni avec les trains ni avec le café !) et signifie « Qui est exprimé de manière à ne laisser aucun doute possible » (Littre). La jurisprudence l'interprète comme "écrit et préalable" : ce qui peut poser problème pour certaines recherches et appelle parfois une dérogation. [NDR]

²⁴ Relevons cependant que le volontariat du sujet – initial et maintenu – introduit un biais de subjectivité (différent de celui que le double aveugle cherche à éviter) qui atténue l'objectivité scientifique que l'essai est censé assurer. C'est donc un compromis entre rigueur scientifique et considérations éthiques. [NDR]

²⁵ Il faut reconnaître que l'on ne sait pas bien quelle information on pourrait donner, alors que les observations – souvent faites en des lieux différents – ne sont pas encore centralisées. Ceci, hormis le fait aussi que les expérimentateurs ont le souci légitime de limiter les abandons. Ou alors, il faudrait concevoir cette information au fil de la recherche de façon que les sujets, se sentant associés à celle-ci, partagent le souci que leur abandon ne la compromette pas. On peut certes le souhaiter, mais cela entraîne aussi un travail additionnel pour l'équipe de chercheurs. [NDR]

Ces protocoles ne sont pas rédigés par les investigateurs, par des médecins, mais par des techniciens et des juristes²⁶. Même lorsque les notices sont claires, reste une différence de cultures individuelles et il faudrait toute une pédagogie. Ce n'est pas là un problème juridique. Les patients s'en remettent au médecin qui leur dit « signez là » ! Cette intervenante estime que l'information est censée protéger le patient, dûment averti du risque qu'il encourt ; mais qu'en fait, cette supposée information et le consentement qui s'ensuit ont pour fonction véritable de protéger l'investigateur.

Les deux invités acquiescent. L'un note qu'il y a, d'une génération à l'autre, de notables différences tant chez les médecins que chez les patients. L'autre considère que cela pose la question du statut de la signature en milieu médical. En théorie, c'est le médecin investigateur qui présente l'essai, explique, puis recueille l'accord du patient et sa signature. En fait, c'est souvent une personne différente qui fait signer et l'on ne sait s'il y a vraiment information. (Sans compter qu'il y peut y avoir parfois de fausses signatures ...) La signature du patient ne dégage pas la responsabilité du médecin. L'innovation réglementaire française, en 1988, fut de mettre la protection des personnes qui se prêtent à des recherches au centre des préoccupations juridiques. Il serait intéressant, vingt ans plus tard, si cela n'a été fait, de connaître les éventuelles actions en justice et leur résultat. On pense qu'en cas d'incident, il y a généralement une transaction entre avocats. Mais on ne sait pas vraiment ce qui s'est passé.

Comment un patient invité à entrer dans un essai va-t-il prendre l'annonce qu'un placebo pourrait lui être administré ?²⁷ Dès lors qu'il accepte le risque inhérent au produit expérimenté, il peut souhaiter aussi profiter de son éventuelle efficacité ! Mais, même sans cela, le simple fait de savoir qu'on participe à cette loterie n'induit-il pas un état mental qui biaise l'expérience ? On sait en effet combien la suggestion peut influencer sur l'état du patient (voir Knock²⁸ !). La notion de placebo est une vue de l'expérimentateur, elle satisfait le médecin, mais il peut être insupportable au malade de se voir proposer (et, même au médecin d'avoir à proposer) l'alternative entre un possible remède et rien... Sauf si l'on n'en sait rien. Cela dit, il semble que les patients atteints d'une maladie grave soient moteurs au point de passer sur cette question de placebo et de vouloir participer à l'essai : du moins ont-ils une chance de bénéficier du traitement s'il s'avérait efficace. Et puis, il faut compter avec la générosité : beaucoup veulent servir.

Enfin, au chapitre de la constitution d'échantillons en vue d'évaluer un médicament, est signalé que les grands opérateurs sur Internet (Google, Facebook, ...) ont le moyen de rassembler très rapidement de nombreuses personnes susceptibles de participer à un essai. Ces opérateurs entrent ainsi sur le "marché" des essais²⁹.

D. Les aspects économiques

Ceci n'a été que rapidement évoqué. Mais, c'est une dimension importante du problème car des enjeux considérables sous-tendent là les positions des divers acteurs.

²⁶ Hors séance, elle précisera que ce sont généralement des ATC (assistants techniques de clinique) qui établissent les notices : ils ne sont pas médecins et ont une connaissance du droit sans être pour autant juristes.

²⁷ Bien entendu, on ne lui dit pas qu'un placebo *va* lui être administré : ce serait contraire au principe même du placebo. On lui dit seulement qu'un placebo *pourrait* lui être administré. En lui expliquant alors la comparaison de deux échantillons tirés au hasard. [NDR]

²⁸ *Knock ou le Triomphe de la médecine*, pièce de Jules Romains.

²⁹ voir par exemple : <https://www.patientslikeme.com/>, google-health.

Les invités indiquent que, pour qu'une molécule originale devienne un médicament disponible sur le marché, le coût total est de l'ordre du milliard d'euros voire davantage. La moyenne se situerait plutôt vers 1,2 ou 1,3 G€. Et, cela prend dix à douze ans. Il faut bien que cela soit financé. Face à quoi, la politique publique fait pression pour que les prix soient modérés. Les prix des médicaments sont plutôt inférieurs en France à ceux observés dans d'autres pays. Ce qui, au passage, ferait que les génériques n'offrent pas, en France, un avantage de prix considérable. Certains estiment que le prix élevé des médicaments contribue pour moins de 10 % au "trou" de la Sécurité Sociale ; et que, si la pression sur les laboratoires devenait trop forte, le risque serait que la recherche et la fabrication se délocalisent : la France devrait alors importer ses médicaments.

Les stratégies actuelles de découverte et mise au point de nouveaux principes actifs se sont modifiées, avec le développement des techniques notamment de génie génétique. Beaucoup de grandes firmes pharmaceutiques n'ont plus de département recherche en interne, mais entretiennent une veille scientifique intense pour détecter et racheter les produits de start-up lorsqu'elles découvrent un nouveau traitement.

La fixation du prix de remboursement d'un médicament très actif sur de petites populations bien ciblées devient plus difficile car il s'agit d'un marché de niche, forçant les spécialistes à imaginer de nouvelles stratégies en conditionnant le remboursement d'un fort prix d'achat à la « satisfaction » du « patient- client » via les organismes de sécurité sociale.

E. Comment se prennent les décisions ?

Il semble à un participant que, il y a dix ans, la discussion que nous avons ce soir aurait beaucoup plus porté sur la chimie et la biologie. On s'intéresse davantage aux effets : d'où des protocoles observationnels où la statistique prend une position dominante. Là, on prend en vue des aspects plus psychologiques ou politiques. Y a-t-il encore une recherche fondamentale, demande-t-il ? Pascal Maire en convient tout à fait : signalant toutefois que l'on revient aussi parfois à des considérations plus positives (voulant dire par ce mot : prenant en compte la matérialité des choses et non seulement leur dimension subjective.) Par exemple, jadis on pourfendait l'ostéopathie ; aujourd'hui, elle s'inscrit parmi les méthodes thérapeutiques acceptées.

Jean-Christophe Thalabard précise qu'il faut avoir à l'esprit la différence entre une recherche d'aval, très appliquée, vouée à développer des stratégies thérapeutiques, et une recherche d'amont, plus fondamentale, dont les structures gardent un esprit très "Claude Bernard" : elles doivent se nourrir de l'autre. La recherche s'attaque à des questions fondamentales mais renouvelées, telles que les nouveaux outils de manipulation de l'ARN³⁰, les couplages physique/chimie pour rendre perméable à certaines molécules la barrière entre sang et cerveau ou encore les perspectives de thérapie cellulaire.

D'une façon générale, les participants s'interrogent sur la dynamique de "tout ce système" et sur son pilotage effectif. Quel est le poids respectif des services marketing et commercial, des services techniques et des données scientifiques ? Comment fonctionne le système que forment les *entreprises du médicament*³¹ ? Quel est le poids du Comité Economique des Produits de Santé (CESP) [

³⁰ Les ARN (acides ribonucléiques) assurent diverses fonctions entre l'ADN porteur de l'hérédité (et popularisé par les tests d'identification des personnes et de leur filiation) et la production, dans les cellules de l'organisme, des diverses substances nécessaires au fonctionnement de celui-ci. [NDR]

³¹ Voir l'organisation qu'elles se sont donnée, LEEM (LEs Entreprises du Médicament) : « faciliter les échanges entre ses membres, à défendre leurs intérêts collectifs et à resserrer les liens avec les autres professions de santé. » (www.leem.org/).

<http://www.sante.gouv.fr/comite-economique-des-produits-de-sante-ceps.html>] qui contribue à fixer les prix de remboursements par l'assurance maladie et qui donc contribue à la diffusion des innovations, celui de la Haute Autorité de Santé (HAS) et celui de l'Agence Nationale de Sécurité des Médicaments et des produits de santé ANSM (<http://ansm.sante.fr/>) ? Quel est le poids spécifique des patients et de leurs demandes éventuelles, plus ou moins rationnelles ou influençables ? Enfin quel est le poids des professionnels de santé, médecins en premier lieu, et pharmaciens ?

Le médicament est un objet étrange : entre science, santé et société...

Un participant avait noté – à partir de son expérience de la Commission des stupéfiants – que les décisions sont parfois très “politiques”. Un autre interroge sur le déploiement de l'AZT pour soigner le Sida, alors que son effet semblait plutôt réduit. Ce médicament, répond Pascal Maire, n'apportait peut-être en effet pas grand-chose : il semblait avoir plus d'effet sur un marqueur³² de la maladie que sur la maladie même. Mais il y a eu un mouvement des patients, qui le réclamaient ; et cela a entraîné une dynamique d'essais thérapeutiques. La même chose avait pu s'observer pour le traitement de la tuberculose par la streptomycine, à la fin de la 2^e Guerre mondiale, en Grande Bretagne : ce fléau, l'état économique du pays et la volonté de stimuler la reprise industrielle ont amené les autorités à se lancer dans un essai randomisé. Le cas du Sida est intéressant, car c'est la première fois que les patients ont pris directement un rôle majeur dans la lutte contre la maladie.

L'invité poursuit en évoquant la maladie d'Alzheimer. On s'interrogeait sur l'intérêt de quelques petits médicaments qui avaient des effets secondaires incertains, plutôt qu'une prise en charge médico-psycho-sociale. Là, le médicament a été un catalyseur pour développer une politique de prise en charge plus globale. Un participant précise à ce point, que la présence d'un conseiller du ministre, personnellement concerné par la maladie d'Alzheimer, avait été déterminante. On a vu ainsi apparaître des méthodologies d'essais “clusterisés” : la randomisation se faisant à l'intérieur même de l'équipe soignante, pour différencier des gestes ou des écoles...

Pour l'autisme, la situation reste très confuse. On a entendu tout et son contraire.

Une autre intervention d'un participant citait une étude qui avait mis en évidence une fréquence d'appendicectomies deux fois plus élevée à Marseille qu'ailleurs : mais, qui relevait aussi qu'il s'y trouve deux fois plus de chirurgiens. Les traitements seraient-ils plus dictés par l'offre médicale que par le besoin de traitement ? Jean-Christophe Thalabard constate certes que la cartographie peut montrer des concentrations surprenantes. Mais il ne faut pas en tirer de conclusions hâtives. Corrélations n'est pas cause. Des analyses plus fines – épidémiologiques, géographiques, démographiques, sociologiques – de l'habitat, du travail et des déplacements urbains et péri-urbains suggèrent qu'en Ile de France, par exemple, pour des raisons de commodité d'accès, des concentrations d'activités médicales, centres hospitaliers, cliniques, etc. le long des grands axes de circulation.

Quelques questions portent sur l'autorisation de mise sur le marché (AMM). L'autorisation par l'Agence européenne suffit-elle ? – oui : elle permet la commercialisation dans tous les pays de l'Union et tend donc à être demandée plutôt qu'à la seule agence française.

³² Un marqueur est une substance dont la présence ou l'élévation permettent de diagnostiquer ou de suivre l'évolution d'une maladie, sans que soit nécessairement élucidé comment elle intervient dans cette évolution.

Une interrogation demeure cependant, quant à l'indépendance des experts. Un participant remarque que, à ce point du débat, l'on n'a pas évoqué les conflits d'intérêts³³. (Quoique les deux invités aient justement commencé par cela !) L'un note que les Anglo-saxons exposent sans détour leurs liens d'argent ; tandis que chez nous, on en est plus gêné, on tend à les cacher. C'est une question de culture. Et, ajoute l'autre, la question n'est pas exclusivement d'ordre pécuniaire : il faut ajouter les liens d'influence. Le problème n'est pas seulement de savoir si l'expert est expressément acheté par un groupe d'intérêt : il est intrinsèquement conditionné par son milieu. Quelle peut, par exemple, être la réaction d'un pharmacien, scientifique positif, à qui l'on enseigne l'homéopathie ? On l'invite là à sortir de ses représentations et de son milieu.

Saisissant cet exemple, un participant rapporte que Jacques Benveniste a été sévèrement sanctionné³⁴ : mais, tardivement, alors que ses liens avec les laboratoires Boiron étaient notoires³⁵. Il considère que cela montre qu'existe une censure. Il évoque, dans la même ligne le soutien apporté, en 1984, aux "médecines parallèles" par Georgina Dufoix, ministre des affaires sociales.

*

Quelques mots de conclusion

« Faut-il avoir peur des médicaments ? » titrait une tribune libre émanant d'un groupe actif de médecine générale, proche de la revue *Prescrire*, dans le quotidien « Le Monde » en avril 2015.

Oui, en général, puisque, rappelons que, dans la Grèce antique le *pharmakon* désignait aussi bien le médicament que le poison. Le danger associé à toute prise médicamenteuse justifie sa manipulation par des professionnels dûment formés et expérimentés. Les nouvelles thérapies, toutes ciblées qu'elles soient, n'y échappent pas et imposent de déterminer le véritable rapport bénéfice/risque des individus qui y sont exposés.

Mais, beaucoup de progrès a été accomplis, ces trente dernières années, tant au niveau national qu'europpéen et international, vers une volonté de transparence pour le développement, l'évaluation et la mise sur le marché de nouvelles substances. Sans doute reste-il encore beaucoup de chemin à accomplir et il existera toujours des possibilités de s'arranger avec les règles instaurées. Elles sont actuellement relativement consensuelles, mais la difficulté demeure d'en contrôler le respect effectif. Développer une culture de l'évaluation du risque chez les professionnels de santé, d'une part ; et,

³³ On désigne ainsi le fait qu'une personne appelée à donner un avis ou à prendre une décision puisse cumuler divers rattachements ou activités. Elle adopte alors des points de vue différents selon qu'elle se trouve dans l'une ou l'autre de ces positions ; mais l'on n'a pas la garantie qu'elle sache, lorsqu'elle se trouve dans l'une de ces positions, faire abstraction des circonstances et conditions des autres. En particulier, qu'étant payée pour l'une, elle ne soit (délibérément ou inconsciemment) portée à favoriser les intérêts du payeur lorsqu'elle devrait arbitrer entre des solutions ou des décisions où celui-ci est partie prenante. Or, souvent, la compétence d'un expert tient justement au fait qu'il a acquis une expérience dans un domaine ou une institution différente de celle où il est appelé à rendre un avis supposé neutre. [NDR]

³⁴ Si la sanction a été sévère, il ne faut pas pour autant jeter un discrédit total sur le personnage, dont l'hypothèse sur une supposée « mémoire de l'eau » a été moquée. En matière scientifique, des idées d'abord considérées impensables se sont parfois révélées être des intuitions remarquables (la physique quantique en fournit maint exemple). D'autres scientifiques incontestés (tel le professeur Luc Montagnier) ont prêté attention aux hypothèses de J. Benveniste. [NDR]

³⁵ Le commentaire d'un des invités sur les tests présentés par ce chercheur est reclassé, pour le présent compte rendu, à la fin du § A.

d'autre part, une communication objective, ouverte, adaptée et citoyenne envers les patients ou leurs associations demeurent des objectifs importants.

De nouveaux défis contraignent à repenser les méthodes anciennes et en stimulent l'apparition de nouvelles, perpétuellement remises en cause. Le statisticien, longtemps isolé face aux professionnels de santé dans les comités de suivi d'essais a été rejoint par des représentants de la société civile, des associations de patients. C'est peut-être encore trop peu mais déjà beaucoup dans la progression d'un regard citoyen et sociétal sur les pratiques. Les scandales qui ont emplis les médias au cours des trente dernières années – certains aux conséquences dramatiques : sang contaminé, hormone de croissance, hypocholestérolémiant, anti- inflammatoires, coupe-faim, traitement hormonal de la ménopause, contraception orale, vaccination, ... pour n'en citer que quelques-uns – ont évité de s'endormir sur l'illusion d'une médecine toute puissante et d'un risque zéro. Les discussions constructives autour de graves incidents survenus lors d'essais chez l'homme tout comme le débat de ce soir n'auraient sans doute pas existé sous cette forme il y a vingt ans.



Vous avez dit “significatif” ?

Le terme “significatif” a un sens très différent selon qu’on se réfère à la preuve d’une propriété au vu d’un résultat ou que l’on envisage l’utilisation.

1. *significativité pour l'utilisateur*

D’un point de vue pratique, pour ce que l’on a à faire, il n’est pas nécessaire que les propriétés d’un objet soient rigoureusement précises. Si l’on achète une table de 1,20 mètre, on ne demande pas cette longueur au millième de millimètre près. D’une façon générale : les contrôles industriels se font avec une marge de *tolérance*. Le poids, la pureté, la durée, etc. admettent, par rapport à la valeur théorique v , un petit écart, en plus ou en moins, inférieur à un maximum que nous notons e . Cet écart maximum admissible dépend de l’utilisation qu’on a en vue. Tant que la mesure constatée demeure dans l’intervalle $[v - e \text{ à } v + e]$, on dit que l’écart n’est pas significatif.

2. *significativité de la mesure*

La mesure elle-même n’est pas infiniment précise et comporte une marge d’erreur. Si en répétant plusieurs fois la même mesure, l’appareil de mesure donne en moyenne la vraie valeur, on dit qu’il est *juste* (pour une balance) ou *sans biais* (pour une statistique sur échantillon). Et, donnant d’une fois sur l’autre des résultats légèrement différents : on dit qu’il est *précis* s’ils ne sont pas trop dispersés. Ils se distribuent généralement selon une loi de probabilité dite *normale* (ou, “de Laplace-Gauss”).

(Bien entendu, il est inutile que la mesure soit beaucoup plus précise que la tolérance de l'utilisateur (peser au milligramme près lorsqu’il admet une plage d’équivalence de 10 grammes, par exemple.) Mais bien sûr, il ne faut pas non plus que l'imprécision de la mesure soit presque aussi grande que la marge de tolérance, ni *a fortiori* supérieure !

Des techniques de tirage au hasard (*randomisé*) donnent des échantillons sans biais ; mais il y a toujours une fluctuation d’échantillonnage. On sait toutefois la calculer, dans le cas randomisé, en la caractérisant par l’écart-type de la distribution. La mesure est dite *significativement* supérieure à une valeur donnée (ou souhaitée) v si, au cas où la vraie valeur serait seulement égale à v , on n’aurait eu qu’une faible probabilité d’obtenir la valeur observée.

Pour faire le test, on se fixe le risque d’erreur que l’on accepte de courir, c’est à dire une probabilité α . Faisant l’hypothèse que v serait la vraie valeur et on calcule un seuil tel que la valeur observée à partir de l’échantillon aurait alors la probabilité α de dépasser ce seuil. Si elle le dépasse effectivement, on décide d’admettre que la vraie valeur est supérieure à v . Autrement dit, on répond “oui” à la question “le traitement testé est-il d’efficacité supérieure à v ?”. Si en fait il n’a pas cette efficacité, on se sera trompé en prenant cette décision : cela peut se produire avec la probabilité α ; et c’est un risque qu’on assume. Souvent, on se la fixe à 5 %, c’est à dire qu’on accepte de se tromper une fois sur 20 en prenant pour efficace un traitement qui ne l’est pas. On dit aussi que l’efficacité est supérieure à v avec un “niveau de confiance” de 95 %. Ce risque de 5 % est commode car le seuil d’acceptation se situe alors à peu près à 2 écarts-types (exactement, 1,96) au-dessus de v . Mais, selon les risques admissibles dans la situation considérée, on choisit aussi bien toute autre valeur de α .

Au total, la significativité pour l'utilisateur a à voir avec l’écart admissible sur la valeur de la variable considérée ; la significativité de la mesure, avec le risque admissible d’une décision erronée

L'équivalence clinique

On considère comme cliniquement équivalents deux traitements dont l'efficacité (pourcentages de succès, niveau d'un paramètre biologique, etc.) est assez voisine.

Deux justifications, de nature différente, peuvent en être données :

⇒ les cliniciens n'ont pas besoin de mesurer la performance d'un traitement de façon précise : de leur point de vue, des fluctuations modérées ne sont pas significatives compte tenu de la variabilité que comportent tous les processus vitaux ;

⇒ si les inconvénients mentionnés sont majeurs, une efficacité légèrement moindre peut être jugée acceptable – comme c'est évoqué dans le présent débat – en contrepartie d'un gain sur ces inconvénients. C'est un jugement discrétionnaire, à dire d'experts, sanctionné évidemment par l'autorité régulatrice.

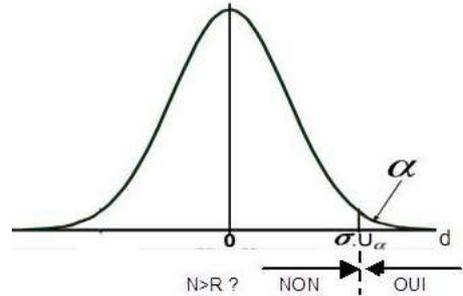
Certains auteurs donnent une troisième justification : l'efficacité du traitement de référence n'est elle-même connue qu'approximativement, puisqu'elle a été déterminée sur un échantillon. C'est un argument : en ce sens qu'on ne peut refuser de considérer comme équivalents deux traitements dont la différence d'efficacité mesurée serait plus petite que la plage d'incertitude de la mesure. En revanche, pour un test de non-infériorité, on n'a pas à additionner la plage d'équivalence du clinicien à celle de l'incertitude sur l'efficacité du traitement de référence, comme certains semblent le faire en juxtaposant des intervalles de confiance. En effet, la procédure correcte de comparaison statistique considère la différence des performances constatées pour les deux traitements : cette différence est affectée d'un aléa dont la variance est la somme des variances observées sur les deux échantillons. (Ce sont les variances qui s'ajoutent et non les écarts-types.) L'incertitude sur la référence est ainsi prise en compte dans un test qui confronte directement le nouveau traitement et celui de référence. Ceci, que l'on se donne ou non une plage d'indifférence clinique telle qu'évoquée ici.

En pratique donc, on prendra comme zone d'équivalence au moins la plage d'incertitude sur le traitement de référence, mais éventuellement plus si les experts considèrent que la différence n'est pas cliniquement significative ; ou que la légère perte que l'on accepte est justifiée par la réduction des inconvénients lorsqu'on substitue le nouveau traitement.

Tests de supériorité ou de non-infériorité

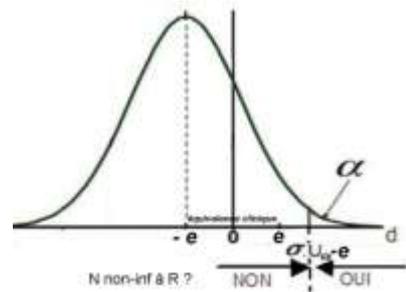
Test de supériorité : le traitement N (nouveau) est-il supérieur à R (celui de référence) ?

La différence d'efficacité (d), mesurée sur échantillons, suit une loi de probabilité figurée ci-contre par une "courbe en cloche". Si les deux traitements avaient la même efficacité, d serait en *espérance* (en moyenne) nulle. Et, la valeur observée de cette différence aurait une probabilité α de dépasser un seuil critique : obtenu en multipliant l'écart-type (qui se déduit lui aussi des observations) par la valeur d'une "variable gaussienne centrée réduite", notée u , donnée par une table pour une probabilité de dépassement égale au risque α qu'on a choisi. Si elle franchit effectivement ce seuil, on considère que cela avait peu de chance de se produire (pour deux traitements d'égale efficacité) et que donc d est significativement positive, c'est à dire que N est supérieur à R. (Ceci, avec une probabilité α de se tromper : et l'on aura choisi α selon le risque qu'on accepte de courir).



Test de non-infériorité : le traitement N (nouveau) est-il au moins équivalent à R (celui de référence) ?

Ici, on ne demande pas une supériorité stricte, mais on accepte une plage d'équivalence clinique : on accepte une différence qui pourrait aller jusqu'à e en plus ou en moins. La différence d'efficacité vraie reste acceptable tant qu'elle est supérieure à la limite inférieure de la plage d'équivalence, c'est à dire $-e$. Le test se présente comme celui de supériorité, mais en se référant à une espérance de d égale à $-e$ et non plus à zéro. La courbe de probabilité est décalée de e vers la gauche.



[*nota* : dans tous les cas, la différence d entre les résultats des deux échantillons a pour variance (c'est à dire le carré de l'écart-type) la somme des variances respectives de ces deux résultats, au dénominateur desquelles figure le nombre d'observations de l'échantillon correspondant.]