

# Méthodologie des essais cliniques de petits effectifs

Muriel Vray,<sup>1</sup> Danièle Girault,<sup>2</sup> Natalie Hoog-Labouret,<sup>3</sup> Raphaël Porcher,<sup>4</sup> Jean-Christophe Thalabard<sup>5</sup> et les participants à la Table Ronde n° 1 de Giens XIX†

1 Institut Pasteur, Unité Epidémiologie Pathologies Emergentes, Paris, France

2 Laboratoires Takeda, Puteaux, France

3 Afssaps, Saint-Denis, France

4 Hôpital Saint-Louis, Paris, France

5 Hôpital Necker, Paris, France

## Résumé

Les essais cliniques de petits effectifs sont les essais dont le nombre de patients ne permet pas de répondre de manière correcte à l'objectif d'une étude avec les règles méthodologiques communément appliquées. Cette situation est fréquente dans le cas de maladies rares, en pédiatrie, dans certaines pathologies cancérologiques ou lorsque l'on souhaite limiter le nombre de patients exposés au traitement. Après avoir identifié les principaux problèmes méthodologiques, sont abordées les méthodes classiques (essai contrôlé, randomisé et réalisé en double insu sur deux groupes parallèles, essai croisé, plan factoriel, essai réalisé avec plusieurs mesures répétées dans le temps, plan additionnel, plan de retrait aléatoire ou plan d'échappement précoce) et les méthodes moins usuelles (approches séquentielles, méta-analyses, méthode « N of 1 » et autres méthodes d'aide à la décision ou de modélisation). Des recommandations sont ensuite proposées pour s'assurer que les résultats obtenus ne sont pas le fait du hasard et augmenter le niveau de preuve.

**Mots clés :** essais de petits effectifs, méthodes statistiques, niveau de preuve

Les essais cliniques de petits effectifs sont des essais menés et analysés avec un nombre de patients inférieur à celui que l'application des règles méthodologiques les plus communes ferait considérer comme nécessaire pour répondre à l'objectif de l'étude. La présente réflexion ne concerne donc pas les essais classiques de phases I et II, où le nombre de patients à inclure, même s'il est quantitativement faible, est méthodologiquement adapté à l'objectif de l'étude. Elle concerne tous les essais visant à établir la preuve de l'existence d'un effet donné, que ce soit en termes d'efficacité ou de tolérance.

Cette situation se rencontre fréquemment dans le champ des maladies rares, en pédiatrie ou en cancérologie, dès lors que des contraintes externes de disponibilité de patients ne permettent pas le recrutement d'un nombre de patients permettant de satisfaire à des critères de puissance classiques. C'est également le cas lorsque l'on souhaite limiter le nombre de patients exposés aux modalités thérapeutiques étudiées, soit en raison de contraintes

de protection particulière des personnes (enfants, femmes enceintes), soit du fait de la toxicité du produit (cancérologie) ou de son inefficacité (placebo, faibles doses). La réflexion méthodologique concernant les études de « preuve de concept », dont les résultats conduisent à la décision de poursuivre ou non le développement d'une molécule, quoique fondée sur un contexte différent, peut se rattacher au présent domaine.

La question essentielle autour de ces essais de petits effectifs est le niveau de preuve apporté par les résultats de telles études, qui renvoie à la question de la pertinence de l'utilisation de méthodes non standards.

En effet, réaliser, par nécessité, un essai avec de petits effectifs expose à plusieurs ordres de difficultés, certaines n'étant pas spécifiques mais accrues par le nombre réduit des patients :

1. Le risque de ne pas pouvoir conclure alors même qu'il existe une réelle différence entre les modalités thérapeutiques comparées, du fait de la grande variabilité des fluctuations

† Pour la liste des participants, voir en fin d'article.

aléatoires. Seul un effet quantitativement important est susceptible d'être mis en évidence.

2. La nécessaire prudence dans l'extrapolation des résultats observés à l'ensemble de la population susceptible de recevoir le traitement, dès lors que les conditions d'obtention de ces résultats s'écartent par trop des conditions ultérieures d'utilisation en situation clinique courante.

3. La *quasi*-impossibilité de prise en compte de covariables dans l'analyse de la réponse au traitement, rendant l'identification notamment des patients répondeurs difficile.

4. L'absence parfois de mécanisme biologique explicite permettant de s'appuyer sur un critère objectif ou marqueur biologique permettant d'étayer les résultats, alors que la taille réduite de l'effectif limite le niveau de preuve apporté par l'évaluation clinique.

5. Le contexte souvent très particulier de recours à de tels essais du fait notamment de l'absence d'alternative thérapeutique, de la relation privilégiée entre patients ou association de patients et thérapeutes, de l'absence de toutes données antérieures sur la variabilité des critères envisagés et les risques encourus, autant de sources de biais sur l'évaluation objective du traitement.

La somme de ces incertitudes et contraintes impose dans la conduite de ces essais une rigueur et une transparence égales voire supérieures à celles des essais classiques afin de s'assurer que les résultats obtenus ne sont pas simplement le fait du hasard.

Le plan retenu se propose d'évoquer, dans un premier temps, des méthodes qui ressortent de l'approche méthodologique classique mais qui peuvent prendre un aspect particulier dans ce contexte avant d'évoquer, dans un second temps, des méthodes jugées moins usuelles.<sup>[1]</sup>

## 1. Méthodes classiques d'évaluation

Quelle que soit la méthodologie utilisée, un essai de petits effectifs ne saurait déroger à la règle générale et ne peut être entrepris qu'à la triple condition d'une question médicale pertinente et importante, d'une conduite selon une méthodologie rigoureuse permettant de répondre à la question posée, d'un respect des règles et principes éthiques.

Le modèle de référence reste l'essai contrôlé, randomisé, en double insu mené sur deux groupes parallèles. Il devra être systématiquement envisagé et privilégié, y compris en ayant recours à des essais multicentriques internationaux lorsqu'un recrutement local ou national ne permet pas d'atteindre les effectifs classiquement requis. L'utilisation de protocoles et de cahiers en ligne peut faciliter de tels essais multicentriques comme le préconise par exemple une organisation telle que

GEREQ.<sup>[2]</sup> La réduction du nombre de sujets passe par une réduction de la variabilité de la réponse, qui peut être obtenue par des méthodes classiques de réduction de la variance.

### 1.1 Essai croisé ou essai en « cross-over »

Le sujet est ici son propre témoin et reçoit, dans un ordre aléatoire, deux ou plusieurs séquences de traitement. Dans le cas de la comparaison de deux traitements A et B, chaque sujet reçoit soit la séquence A puis B, soit l'inverse B puis A, l'attribution d'une séquence particulière résultant d'un tirage au sort équilibré. Les avantages sont : (i) pour le clinicien, une comparaison de l'effet sur un même patient, gommant les effets de la variabilité inter-sujets ; et (ii) pour le méthodologiste, un gain de puissance d'autant plus important que la variabilité du critère mesuré pour un même sujet (variabilité intra-sujet) est faible par rapport à la variabilité de cette même mesure entre sujets (variabilité inter-sujets) et que la différence d'effet attendu entre les deux traitements comparés est faible par rapport à la variabilité inter-sujets. Cependant, ce plan expérimental est contraignant : il concerne des pathologies chroniques stabilisées ; il suppose que chaque sujet aborde la deuxième phase de l'étude dans un état comparable à celui du début de la première phase, éliminant *a priori* les guérisons, décès ou perdus de vue en cours d'étude ; le critère principal de mesure de l'efficacité de l'intervention ne doit pas être modifié par des phénomènes d'apprentissage (exemple du tapis roulant dans l'angor) ou d'accoutumance (régime amaigrissant). De plus, l'effet du traitement reçu dans la première phase doit être épuisé lorsque le sujet débute la deuxième phase, imposant, en cas d'effet rémanent possible, une période dite d'élimination (« wash-out ») plus ou moins longue. La durée totale de traitement pour chaque patient est donc au moins deux fois plus longue que dans une étude à groupes parallèles, exposant au risque accru de perdus de vue, alors même que l'analyse finale ne peut porter que sur les sujets ayant un critère mesuré pour chacune des deux phases de traitement. Ceci impose donc des critères se modifiant rapidement sous traitement pour limiter la durée totale de l'étude. Enfin l'analyse classique impose d'éliminer un effet d'interaction entre la période et le traitement avant de pouvoir comparer les effets des traitements.<sup>[3,4]</sup>

### 1.2 Plan factoriel

Un tel plan est classiquement utilisé pour évaluer simultanément les effets propres de deux ou plusieurs traitements ainsi que leurs effets synergiques ou antagonistes. Ainsi un plan factoriel  $2 \times 2$  concernant deux produits A et B comporte quatre bras : le bras *verum* A + *verum* B ; le bras *verum* A + placebo B ;

le bras placebo A + *verum* B ; le bras placebo A + placebo B. Le bénéfice attendu en termes de puissance dépend du degré d'interaction entre les deux traitements. Une telle approche est souvent utilisée pour évaluer différentes stratégies de doses.

### 1.3 Essai à mesures répétées du critère dans le temps

Plutôt qu'une mesure unique sur un grand nombre de sujets, il peut être envisagé de mesurer plusieurs fois un critère donné chez un plus petit nombre de sujets. La puissance d'une telle étude, donc le nombre de sujets nécessaire, dépend du degré de corrélation entre les mesures individuelles dans le temps, du nombre de ces mesures et des variances intra- et inter-sujets.

Le souci de « non-perte de chance » a conduit à proposer quelques plans expérimentaux particuliers, quoique satisfaisant à une méthodologie classique.

### 1.4 Le plan additionnel ou « add-on design »

Les patients bénéficient tous du traitement de référence. Sous celui-ci, ou au décours de celui-ci selon les circonstances, ils sont randomisés entre traitement additionnel étudié versus placebo. Un tel plan, qui suppose que le traitement de référence et le traitement à l'étude ont des mécanismes d'action différents, est souvent le seul possible dans le cas de pathologie létale où un traitement de référence a fait preuve de son efficacité.<sup>[5]</sup>

### 1.5 Le plan de retrait aléatoire ou « randomised withdrawal design »

Tous les patients reçoivent initialement le traitement à l'étude mais seuls les patients répondeurs poursuivent l'essai. Ils sont alors tirés au sort entre poursuite du traitement ou placebo. Les patients des deux bras sont suivis régulièrement sur des critères cliniques ou biologiques jusqu'à apparition d'une rechute. Ils sont alors remis sous le traitement « efficace ». Cette approche permet de réduire la période pendant laquelle les sujets reçoivent du placebo, tout en n'exposant au placebo que les sujets répondeurs, optimisant la possibilité d'observer l'effet du traitement chez les répondeurs.<sup>[6]</sup> En revanche, dans le cas de maladies d'évolution lente et imprévisible à rémissions spontanées (sclérose en plaques [SEP] par exemple), ce plan peut exposer inutilement aux effets secondaires nocifs des traitements, sans possibilité d'attester de cette rémission.

Cette méthode est acceptée aux Etats-Unis dans le cadre d'obtentions de mise sur le marché.

### 1.6 Le plan d'échappement précoce ou « early escape »

Les patients sont « sortis » de l'étude s'ils répondent à un critère d'échec défini *a priori*. L'évaluation porte sur le taux d'échec, permettant de minimiser la durée d'exposition à un traitement inefficace.

Une des difficultés importantes de ces essais concerne la définition d'un critère binaire d'évaluation en échec/rechute, ainsi que la perte de puissance rapide de ces études en cas d'un nombre important de perdus de vue.

## 2. Méthodes moins usuelles

### 2.1 Les essais séquentiels

La technique des essais séquentiels a été développée dans le but de réduire le temps avant décision, tout en minimisant le nombre moyen de sujets inclus dans l'étude. On distingue les essais séquentiels où les sujets rentrent « par paire » dans l'étude avec une évaluation décisionnelle portant sur la comparaison des réponses de chaque paire et les essais séquentiels groupés, plus adaptés à la pratique.<sup>[7,8]</sup> Ces derniers (« group sequential design ») sont largement diffusés en pratique. Ils introduisent des analyses intermédiaires soit par groupe de patients inclus, soit par nombre d'événements observés dont le nombre et la répétition plus ou moins régulièrement espacée dans le temps sont planifiés *a priori*, qui garantissent des risques d'erreur de type I et II préfixés. Deux types de procédures d'analyses répétées ont été développés : l'approche par « frontières » et l'approche par « fonction de consommation » (« spending function »). Le test triangulaire<sup>[9]</sup> relève de la première approche. En pratique, le suivi consiste à reporter sur un graphique comportant deux zones frontières délimitant une zone de continuation le point (x, y) correspondant aux valeurs de deux statistiques calculées à partir des données observées à chaque analyse intermédiaire. Si le point est au dehors de la zone de continuation, l'essai s'arrête avec rejet de l'hypothèse nulle ou de l'hypothèse alternative selon la position respective du point par rapport aux courbes frontières. Dans le cas contraire, l'inclusion des patients continue. Dans la deuxième approche, des tests conventionnels sont répétés lors des analyses intermédiaires, avec des seuils ajustés au nombre de répétition de ceux-ci. Citons la procédure d'O'Brien et Fleming<sup>[10]</sup> et le cadre général de l'approche par « fonction de consommation du risque de type I » («  $\alpha$ -spending function ») de Lan et DeMets.<sup>[11]</sup>

## 2.2 Les essais flexibles

Ils représentent des extensions des essais séquentiels, dans lesquels le plan expérimental et/ou le schéma thérapeutique sont modifiés en fonction des données provenant de l'étude voire de données extérieures à l'étude lors des analyses intermédiaires.<sup>[12,13]</sup> La différence majeure réside dans la possibilité, en cours d'essai lors d'étapes planifiées, de réviser le protocole et de remplacer le plan d'expérience original par un autre plan d'expérience sous réserve que le risque d'erreur de type I conditionnellement aux observations disponibles ne soit pas modifié. Les règles d'adaptation n'ont pas besoin d'être fixées *a priori*. Les modifications du plan d'expérience peuvent être fondées sur les données recueillies dans l'essai (avec ou sans levée d'aveugle) ou extérieures à l'essai. Cependant, cette approche ne conduit pas forcément à une réduction du nombre de sujets, même si cela peut être le cas lorsque les données d'une étude pilote interne sont utilisées. Dans ce cas, les données de l'étude pilote peuvent non seulement servir à planifier la suite de l'étude mais aussi être prises en compte dans la mesure de l'effet traitement final.

## 2.3 Les méthodes de randomisation adaptative (par extension, essais adaptatifs)

Ils répondent, d'une part, au souci éthique d'attribuer le traitement le plus efficace à une majorité des patients inclus, alors même qu'en début d'étude le principe d'ambivalence (« équilibre ») justifiant la raison d'être de l'essai ne permet pas de favoriser l'allocation d'un des traitements et, d'autre part, à un souci de maximisation de la puissance de l'essai en cas de variances inégales entre les deux bras de traitements.<sup>[14]</sup>

Plusieurs méthodes d'allocation adaptative ont été proposées dont la plus connue est la méthode « play the winner »<sup>[15,16]</sup> reposant sur le modèle de l'urne initialement remplie d'un nombre égal de boules de deux couleurs différentes, correspondant respectivement aux deux traitements à comparer. En fonction de la réponse de chaque patient inclus, le ratio d'allocation des traitements pour les patients suivants est modifié en ajoutant, en cas de réponse favorable du traitement, une boule dans l'urne de la couleur correspondant au traitement reçu et, en cas de réponse défavorable, à l'autre couleur. Des schémas plus complexes peuvent tenir compte de facteurs pronostiques avec des allocations par strates.

Si une telle méthode paraît séduisante sur le plan conceptuel, elle reste relativement peu utilisée en pratique. La difficulté logistique majeure est, sans doute, la capacité d'évaluation rapide d'un critère qui se doit d'être binaire, au fur et à mesure des suivis

de chaque patient inclus. Elle requiert une planification très soignée.<sup>[17]</sup>

## 2.4 La méthode « N of 1 »

Il arrive souvent que le clinicien se trouve face à un patient pour lequel des traitements sont nécessaires alors qu'il n'existe aucun essai ayant démontré leur efficacité sur ce type de patients. Dans cette situation le clinicien peut être tenté d'administrer successivement plusieurs traitements reconnus efficaces pour d'autres types de patients pour évaluer leur effet sur ce patient particulier et choisir le plus efficace. Dans cette stratégie « N of 1 », inspirée du domaine des sciences du comportement, le patient est son propre témoin et le résultat lui est directement adapté. La multiplication de tels « micro-essais » sur une population de patients réputés hétérogènes peut permettre d'avancer quelques conclusions plus générales secondairement sur tel ou tel traitement.

Guyatt et al.<sup>[18]</sup> ont proposé une stratégie contrôlée, randomisée où chaque patient reçoit des paires de traitement sur des périodes consécutives, selon des modalités d'allocation rappelant l'essai en « cross-over ». Chaque séquence de paires de traitement est répliquée plusieurs fois jusqu'à ce que le patient et son médecin soient convaincus du meilleur traitement en termes de bénéfices (tolérance et efficacité).

Cette approche est en théorie intéressante lorsque des essais ont préalablement démontré que le traitement était totalement inefficace chez certains patients ou lorsqu'il persiste un doute sur la possibilité d'une efficacité chez un patient donné, ou encore lorsqu'un patient souhaite essayer un traitement jugé plus ou moins subjectivement dangereux ou inefficace par le praticien. L'approche est adaptée à une pathologie de type chronique nécessitant un traitement long, avec survenue possible d'effets indésirables liés au traitement apparaissant rapidement sous traitement et disparaissant rapidement à l'arrêt du traitement. Le nombre de séquences de paires d'administration n'est pas préfixé mais la décision sera d'autant plus robuste que ce nombre de paires est élevé.

L'expérience dans l'utilisation d'une telle approche reste très limitée.

## 2.5 L'essai pragmatique

Il nous a paru utile de rappeler ici la méthodologie des essais pragmatiques, même si elle a été développée il y a plus de 30 ans par Schwartz et Lellouch,<sup>[19,20]</sup> qui obéit à la même motivation d'orienter les praticiens en les aidant à prendre des décisions thérapeutiques adaptées à la situation réelle de prescription, la plupart du temps après démonstration d'une efficacité intrinsèque.

Dans une telle démarche expérimentale, le traitement ou la stratégie retenue *in fine* correspondent à celui ou celle donnant les meilleurs résultats évalués sur des critères « pragmatiques », qui se doivent de résumer les avantages et inconvénients des stratégies thérapeutiques comparées (succès/échec ou échelle de qualité de vie). L'analyse est en théorie réduite à sa plus simple expression puisque, sans aucun calcul, le traitement retenu est celui qui donne le meilleur résultat. Les difficultés concernent essentiellement la mesure de la différence entre les traitements qui peut être ignorée ainsi que la construction du critère d'évaluation qui doit refléter le « service médical rendu » dans sa globalité.

#### 2.6 Les méthodes de synthèse *a posteriori*

Nous regroupons ici des méthodes reposant sur l'analyse de données existantes et n'imposant pas de mener de nouvelles études.

#### 2.7 Les méta-analyses d'essais de petits effectifs

Sous réserve d'études bien menées, de qualité suffisante, les techniques de synthèse quantitative peuvent permettre d'atteindre une puissance suffisante pour dégager un effet thérapeutique. La difficulté majeure réside dans l'identification des protocoles et des données correspondantes non publiées, indispensables pour s'assurer de l'homogénéité des critères d'évaluation et des patients inclus. La perspective d'une méta-analyse *a posteriori* ne doit pas dispenser d'une réflexion *a priori* sur la possibilité même d'un essai multicentrique tenant compte des spécificités de chaque centre.

#### 2.8 L'analyse décisionnelle

L'intérêt de cette approche purement analytique repose sur la possibilité de réaliser des analyses de sensibilité sous diverses hypothèses en s'appuyant sur des données issues de la littérature et des avis d'experts du domaine. Elle vise à évaluer la modification possible des décisions en fonction des hypothèses sur les modèles d'effet, de leurs probabilités de réalisation et des utilités correspondantes attribuées. La qualité d'une analyse décisionnelle dépend fortement de sa planification initiale afin de structurer la question posée, et de la lisibilité des étapes de son exécution.

### 3. Recommandations

Au terme de cette revue des méthodologies existantes, et partant du principe que ce n'est pas parce que l'effectif d'un essai est petit que les critères de qualité ne sont pas les mêmes, bien au

contraire, les recommandations suivantes peuvent être retenues :

- *Favoriser les protocoles multicentriques internationaux*

Il convient de se donner, chaque fois que possible, la possibilité d'inclure un nombre de patients assurant une puissance raisonnable pour démontrer l'effet attendu, en favorisant les études multicentriques, en favorisant l'accès aux informations concernant les essais cliniques en cours ou prévus, en impliquant les associations de patients quand elles existent. Les essais internationaux, plus difficiles sur le plan logistique, devraient être toujours encouragés et peuvent être facilités par des plateformes de protocoles *on-line* et la diffusion des e-CRF (case report forms [cahiers d'observation]) [cf. GEREQ<sup>[2]</sup>].

- *Assurer une qualité méthodologique irréprochable*

La même qualité méthodologique que pour les essais « classiques » s'impose :

- un protocole écrit avec des objectifs clairement définis ;
- une méthodologie et des critères de jugement adaptés aux objectifs ;
- un suivi conduit conformément aux bonnes pratiques cliniques ;

– des données validées et une analyse planifiée dans le protocole. Si des analyses différentes de celles prévues initialement sont nécessaires, elles devront être clairement justifiées.

- *Eviter la multiplication des expérimentations limitées non protocolisées a priori*

La même rigueur méthodologique s'impose à tous les essais qu'ils soient institutionnels ou industriels, en particulier pour le recueil des données en fonction de l'objectif visé. Ce point est particulièrement important car il impose de différencier l'expérience « sauvage » réalisée par des cliniciens confrontés à la nécessité de prendre en charge leur patient d'un protocole de recherche, qui, même s'il repose sur un effectif très faible, aura été planifié, rédigé et soumis aux autorités compétentes pour validation.

- *Toujours intégrer une composante prospective*

Une composante prospective est nécessaire pour garantir un niveau de preuve suffisant. Les données rétrospectives ou les informations issues de données préalablement enregistrées dans lesquelles les événements étudiés sont survenus antérieurement à la décision de mise en place de l'étude sont souvent entachées de biais plus ou moins faciles à contrôler, particulièrement sur des effectifs faibles.

- *Toujours identifier des hypothèses de départ reposant sur un rationnel aussi développé que possible et explicité*

Les hypothèses doivent être clairement définies *a priori*, quitte à utiliser des méthodes flexibles qui permettent des réadaptations en cours d'études. Les prérequis disponibles à

l'initiation des petits essais sont souvent limités et les hypothèses posées souvent peu documentées. Dans ces conditions, il paraît justifié de pouvoir adapter les procédures de suivi ou d'analyse en fonction des résultats obtenus en cours d'étude. Toutefois, les hypothèses de départ doivent être identifiables et toutes les modifications en cours d'étude doivent être clairement explicitées. De même, la valeur arbitraire d'un  $p < 0,05$  peut être rediscutée.

- *S'assurer d'un comité indépendant*

Le comité indépendant (CI), « garde-fou » de l'essai, doit être constitué ici encore plus qu'ailleurs de membres non impliqués dans le recrutement et le traitement des patients. Sa mission doit être de surveiller le déroulement de l'étude, de proposer des décisions susceptibles de rendre l'étude la plus performante possible, en évitant aux investigateurs d'être juges et parties. Il doit constituer une protection pour les patients et rassurer les cliniciens.

- *Présenter les données individuelles de façon exhaustive et transparente*

Les données, y compris individuelles, doivent être présentées de façon exhaustive et transparente. Faute de données quantitativement importantes, il est nécessaire que toutes les informations recueillies dans ces essais soient présentées à l'ensemble de la communauté susceptible d'en retirer un intérêt. La représentation graphique des données individuelles permet de mieux comprendre les résultats obtenus.

- *Recommander systématiquement des suivis de cohorte de patients traités, en cas d'obtention d'AMM à partir d'essais de petits effectifs, surtout lorsqu'ils reposent sur des critères intermédiaires*

En cas d'obtention d'une autorisation de mise sur le marché (AMM) à partir d'essais de petits effectifs, des études ultérieures (type suivi de cohorte) sont recommandées pour confirmer les résultats observés, sur des critères cliniques, d'autant plus que l'AMM aura été donnée sur des critères intermédiaires. La plupart du temps, ces critères intermédiaires n'ont pas été validés car ils ne répondent pas aux exigences permettant de prédire l'effet obtenu sur des critères cliniques. Si le recours à des critères intermédiaires est particulièrement justifié dans le contexte des petits essais, il est encore plus nécessaire qu'un suivi de tous les patients traités soit ensuite mis en place afin de confirmer le bénéfice apporté aux patients sur des critères cliniquement pertinents mesurés sur le long cours. Ceci permet par ailleurs de confirmer l'efficacité et la tolérance au long terme du traitement, alors que l'enregistrement a été accepté à partir d'essais de durée limitée.

- *Ne jamais figer les règles du jeu*

Les règles du jeu ne doivent jamais être figées, les informations recueillies pouvant modifier les décisions. Par exemple,

dans une maladie rare sans traitement disponible, la mise en évidence d'un effet bénéfique inattendu avec un traitement doit autoriser à modifier le plan d'analyse, compte tenu de la difficulté à mettre en place un nouvel essai sur ce nouveau critère.

- *Impliquer les associations de patients à tous les stades des études*

Les associations de patients doivent être impliquées, notamment dans le cas des maladies rares, et à tous les stades de l'étude. Elles permettent d'aider au recrutement en informant les patients de la réalisation de l'étude et en facilitant également le recueil du consentement par l'investigateur. Elles peuvent aussi avoir un rôle de conseil lors de la rédaction du protocole pour la faisabilité des examens prévus en fonction du handicap éventuel des patients, et en aidant à la diffusion des résultats lorsqu'ils sont positifs pour l'accès au traitement d'autres patients.

- *Ne jamais oublier l'intérêt direct du patient présent et futur*

L'intérêt du patient présent et futur reste en effet l'élément fondamental de ces études.

## Participants

C. Auriche (Afsaps, Saint-Denis), J. Bernard (Alliances Maladies Rares, Marly-le-Roi), D. Boichut (DRC, Paris), M. Bouhassira (Lilly, Suresnes), P. Boutouyrie (HEGP, Paris), P. Chaumet-Riffaud (AFM, Ivry), C. Chiron (Hôpital Saint-Vincent de Paul, Paris), Y. Costa (DRC, Paris), F. de Crémiers (Wyeth-Lederlé, Puteaux), S. Etienne (CEREP, Rueil-Malmaison), P. Fender (CNAMTS, Paris), R. Giorgi (Hôpital de la Timone, Marseille), F. Girard (Sanofi-Synthelabo, Le Plessis-Robinson), D. Girault (Laboratoires Takeda, Puteaux), C. Grosskopf (Roche, Neuilly-sur-Seine), F. Gueyffier (Hôpital Cardiologique, Lyon), N. Hoog Labouret (Afsaps, Saint-Denis), Y. Juillet (LEEM, Paris), P. Lechat (Hôpital Pitié-Salpêtrière, Paris), S. Leroy (LEEM, Paris), L. Parmentier (Serono, Boulogne-Billancourt), J. Ropers (Afsaps, Saint-Denis), B. Salanave (CNAMTS, Paris), T. Simon (Faculté de Médecine Saint-Antoine, Paris), E. Van Ganse (Hôpital Edouard Herriot, Lyon), M. Vray (Institut Pasteur, Paris).

## Références

1. Evans CH, Ildstad ST. Small clinical trials: issues and challenges. Washington, DC: National Academy Press, 2003
2. GEREQ. Gestion électronique des données cliniques [online]. Available from URL: [www.gereq.ca](http://www.gereq.ca) [Accessed 2004 July 1]
3. Matthews JNS. Small clinical trials: are they all bad? *Stat Med* 1995; 14: 115-26
4. Senn SJ. Cross-over trials in clinical research. Chichester: John Wiley & Sons, 1993
5. Chiron C, Marchand MC, Tran A, et al. Stiripentol in severe myoclonic epilepsy in infancy: a randomised placebo-controlled syndrome dedicated trial. STICLO Study Group. *Lancet* 2000; 356: 1638-42
6. Temple RJ. A Regulatory authority's opinion about surrogate endpoints. In: Nimmo WS, Tucker GT, editors. *Clinical measurement in drug evaluation*. Chichester: John Wiley & Sons Ltd, 1995
7. Armitage P. *Sequential medical trials*. 2nd ed. Oxford: Blackwell Scientific Publications, 1975
8. Whitehead J, Jones DR. The analysis of sequential trials. *Biometrika* 1979; 66: 105-13
9. Whitehead J. *The design and analysis of sequential trials*. 2nd ed. Chichester: Ellis Horwood, 1992

10. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics* 1979; 35: 549-56
11. Lan KKG, DeMets DL. Discrete sequential boundaries for clinical trials. *Biometrika* 1983; 70: 659-63
12. Posch M, Bauer P, Brannath W. Issues in designing flexible trials. *Stat Med* 2003; 22: 953-69
13. Jennison C, Turnbull BW. Mid-course sample size modification in clinical trials based on the observed treatment effect. *Stat Med* 2003; 22: 971-93
14. Rosenberger WF, Lachin JM. *Randomization in clinical trials*. New York: John Wiley & Sons, 2002
15. Zelen M. Play the winner and the controlled clinical trial. *J Am Stat Assoc* 1969; 64: 131-46
16. Wei LJ, Durham SD. The randomized play-the-winner rule in medical trials. *J Am Stat Assoc* 1978; 73: 840-3
17. Rosenberger WF. Randomized play-the-winner clinical trials: review and recommendations. *Control Clin Trials* 1999; 20: 328-42
18. Guyatt GH, Sackett D, Adachi J, et al. A clinicians' guide for conducting clinical trials in individual patients. *CMAJ* 1988; 139: 497-503
19. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in clinical trials. *J Chronic Dis* 1967; 20: 637-48
20. Vray M. Pragmatic and explanatory trials: ask and answer different questions. *Applied Clin Trials* 1999; 8: 42-50

---

Correspondance et offprints : *Muriel Vray*, Institut Pasteur, Unité d'Epidémiologie Pathologies Emergentes, 28 rue du Dr Roux, 75015 Paris, France.  
E-mail : vray@pasteur.fr