

Méthodologie des essais cliniques en cancérologie

Design of clinical trials in oncology

●● Brigitte Tranchand*, Sylvie Chabaud**, Olivier Colombar***, Sophie Dussart**, David Pérol**

► RÉSUMÉ

Lors de la mise en place des essais cliniques en oncologie, plusieurs questions se posent tant au niveau scientifique qu'au niveau éthique. Au cours des dix dernières années, différentes approches ont été proposées pour améliorer la manière de conduire les essais cliniques lors du développement des médicaments anticancéreux, dans le but de réduire le nombre de patients traités à des doses infrathérapeutiques et d'optimiser l'information obtenue durant ces essais. Cette revue présente les différentes approches disponibles, de la phase I à la phase III, leurs avantages et leurs limites. Un compromis entre toutes ces méthodes est suggéré.

Mots-clés: Essai clinique – Phase I – Phase II – Phase III – Oncologie.

Summary. *There are many scientific and ethical issues when designing clinical trials in oncology. Over the last 10 years, different proposals have been published to improve the way to conduct clinical trials in oncology in order to reduce the number of patients treated at subtherapeutic level, optimize the information gained during these studies and improve the efficiency of clinical trials. This review presents the different approaches available from phase I to phase III. The advantages and disadvantages are discussed, and a compromise between all these methods is suggested.*

Keywords: *Clinical trial – Phase I – Phase II – Phase III – Oncology.*

D'une manière générale, le développement d'un médicament peut être considéré comme un processus d'accumulation d'informations qui se termine lorsque l'information rassemblée est résumée et présentée aux agences d'autorisation de mise sur le marché. Ce processus comprend différentes étapes alternant apprentissage et confirmation (figure 1) [1]. En principe, chaque étape se fait après la précédente, mais avec possibilité de retour en arrière.

* Université de Lyon, Lyon. EA3738 ciblage thérapeutique en oncologie, faculté de médecine Lyon-Sud-Université-Lyon-1 et centre Léon-Bérard, Lyon.

** Unité de biostatistiques et d'évaluation des thérapeutiques, université de Lyon centre Léon-Bérard, Lyon.

*** Université de Lyon, Lyon. EA3738 ciblage thérapeutique en oncologie, faculté de médecine Lyon-Sud-Université-Lyon-1, Lyon.

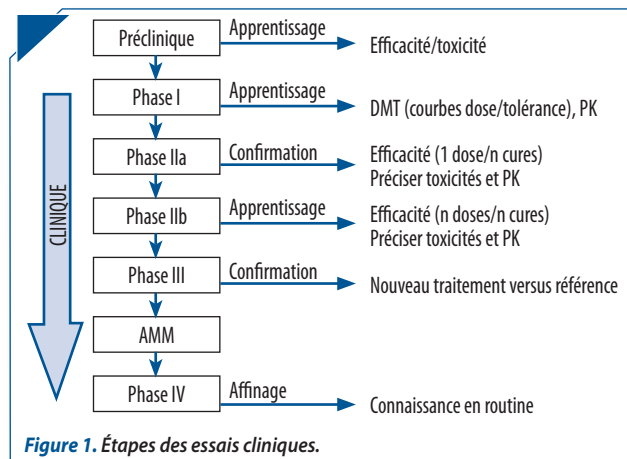


Figure 1. Étapes des essais cliniques.

Apprentissage et confirmation ayant des objectifs différents, tous deux nécessitent différents dessins d'études et différentes méthodes d'analyse. Il est clair que plus le dessin d'un essai sera efficace et plus l'information obtenue sera pertinente et précise.

Pour chaque étape du développement d'un médicament, le protocole doit spécifier les objectifs primaires et secondaires, les critères d'inclusion, le plan de traitement tout au long de l'étape, ainsi que les méthodes d'analyse de données.

ESSAIS DE PHASE I

Le dessin des essais de phase I en oncologie dépend du médicament à l'étude. En effet, pour diverses raisons, l'approche utilisée pour un agent cytotoxique est souvent différente de celle appliquée aux thérapeutiques ciblées :

- ✓ les cytotoxiques sont censés diminuer la taille de la tumeur, quand les cytostatiques sont censés ralentir ou arrêter la croissance tumorale. L'évaluation de la réponse au traitement devra donc se faire sur des critères différents;
- ✓ pour les cytotoxiques, la courbe dose-réponse (efficacité/toxicité) est généralement croissante monotone, quand pour certains cytostatiques tels que l'interféron γ , une cytokine, cette courbe dose-activité peut présenter une forme en cloche (2);
- ✓ enfin, les thérapeutiques ciblées étant supposées moins toxiques que les agents cytotoxiques du fait de leur action spécifique sur une cible, les études de phase I peuvent être réalisées sur des sujets sains quand les études précliniques chez l'animal ont montré que le médicament ne présentait pas ou peu de toxicité.

Un exemple de tels essais concerne l'erlotinib, un inhibiteur de tyrosine kinase (3).

Les essais concernant les thérapeutiques ciblées peu toxiques se rapprochent des essais des médicaments non anticancéreux. La méthodologie exposée ci-après concerne essentiellement les essais de phase I des agents cytotoxiques et des cytostatiques présentant un risque majeur de toxicité.

Objectifs des essais de phase I

Les essais de phase I sont généralement réalisés en monochimiothérapie, mais peuvent aussi servir à évaluer des combinaisons de médicaments déjà connus. Il s'agit de protocoles dont le but est :

- ✓ de déterminer les profils de toxicité et de sécurité d'emploi d'un nouvel agent ou d'une nouvelle combinaison d'agents connus ;
- ✓ de décrire le profil pharmacocinétique du médicament après son administration ;
- ✓ d'établir une dose optimale à administrer pour les études de phase II en découlant.

Dans le cas d'un agent cytotoxique, pour lequel le mécanisme d'action n'est pas spécifique, la toxicité est un élément prépondérant à étudier afin d'établir des courbes dose-toxicité. Il faudra donc dans un premier temps définir la toxicité dose-limitante (TDL). Celle-ci pourra être constituée par un seul événement toxique ou par une combinaison d'événements ; par exemple, "tout événement toxique irréversible de grade supérieur ou égal à 2" ou "thrombocytopenie de grade 4 pendant plus d'une semaine ou neutropénie avec nombre de leucocytes inférieur à 1 000 pendant plus de 2 semaines ou neutropénie fébrile durant plus de 1 semaine".

Une fois la TDL déterminée, le principal objectif sera d'estimer la dose maximale tolérée (DMT), définie comme la dose induisant un risque prédéfini de toxicité sévère. Il s'agit donc d'un protocole d'escalade de doses (en général monochimiothérapie) pour lequel il faut fixer la dose de départ et le mode d'escalade de doses.

Sujets concernés

En oncologie, les sujets inclus dans les essais de phase I d'agents cytotoxiques ou de thérapeutiques ciblées à fort potentiel toxique sont des patients porteurs de tout type de tumeurs solides ou hématologiques ayant échappé aux thérapeutiques disponibles, ou des patients atteints de tumeurs pour lesquelles il n'existe aucun traitement. Ces patients acceptent de participer à l'essai d'un nouvel agent comme dernier espoir de traitement. D'un point de vue éthique, il est donc primordial que l'essai de phase I soit planifié de façon à minimiser le nombre de patients traités à des doses non thérapeutiques.

Les critères d'inclusion impliquent généralement des patients de plus de 18 ans et pas trop âgés, avec des valeurs biologiques normales. Pendant longtemps, la condition d'âge excluait des essais cliniques les enfants et les sujets âgés. Aujourd'hui, même s'il est de plus en plus courant de ne plus fixer de limite d'âge supérieure, des essais de phase I sont spécialement conçus pour ces populations particulières en utilisant l'approche de population.

Plan de traitement des essais de phase I

Le plan de traitement est défini dans le but d'atteindre les objectifs décrits ci-dessus. La planification de l'essai nécessite donc de choisir :

- ✓ la dose de départ ;
- ✓ la méthode d'escalade de doses dont, avec la détermination des paliers de dose, le nombre de patients par palier, l'affectation d'un patient à un palier de dose, le critère d'arrêt de l'escalade.

● La dose de départ

Pour des raisons éthiques, il ne faut pas encourir le moindre risque de toxicité lors de ce premier test sur l'homme. C'est pourquoi, le plus souvent, la dose de départ est déterminée comme étant un dixième de la DL_{10} chez la souris si elle n'est pas toxique chez le chien et, sinon, un tiers de la plus faible dose induisant une toxicité chez le chien.

● L'escalade de doses

Les principales différences trouvées entre les divers essais de phase I publiés en cancérologie concernent essentiellement le mode d'escalade de dose, le nombre de patients par palier et les critères d'arrêt de l'escalade.

La plus ancienne et la plus utilisée des méthodes d'escalade est celle bien connue basée sur la suite de Fibonacci ou la suite de Fibonacci modifiée. Cependant, l'élaboration du dessin des essais de phase I a bénéficié, ces dernières années, d'une avancée significative concernant la méthodologie utilisée pour escalader les doses. Ces méthodes plus efficaces sont encore malheureusement peu utilisées. Il s'agit entre autres de la méthode d'escalade accélérée, de la méthode par réévaluation continue (CRM), ou encore de la méthode guidée par la pharmacocinétique sur la base de relations allométriques.

Méthode "standard" selon la suite de Fibonacci modifiée

Mathématicien italien né à Pise, Leonardo Pisano ou Fibonacci (contraction du latin *filius Bonacci*) vécut à l'époque de la construction de la célèbre tour penchée vers les années 1200. À cette époque, l'Italie utilise encore les chiffres romains. Très jeune, Leonardo accompagne son père en Algérie pour être initié à l'arithmétique, utile à un futur marchand. Il découvre en Afrique du Nord la numération de position et le calcul indo-arabe, qu'il juge plus avancés que les chiffres romains. De retour en Italie, il se consacre pleinement aux mathématiques, et en particulier au problème suivant : "Combien de couples de lapins obtiendrons-nous à la fin de l'année si, commençant avec un couple, chacun des couples produisait chaque mois un nouveau couple, lequel deviendrait productif au second mois de son existence ?" La réponse a été donnée par la séquence où chaque nombre est la somme des deux précédents : 0, 1, 1 + 0 = 1, 1 + 1 = 2, 1 + 2 = 3, 3 + 2 = 5, 5 + 3 = 8, etc.

En outre, il est intéressant de noter que le rapport de deux nombres consécutifs de cette suite forme une autre suite qui converge vers le fameux nombre d'or, utilisé en architecture pendant la Renaissance ou encore par Léonard de Vinci, qui la verra dans les proportions du corps humain avec l'homme

de Vitruve. On en trouve aussi une illustration dans la fleur de tournesol, dans la formation de certains coquillages, sur l'ananas, sur la pomme de pin ou sur le chou romain, qui présentent tous une spirale d'or (figure 2).

La question qui se pose ici est de savoir sur quel fondement scientifique cette suite a été utilisée pour calculer des paliers d'escalade de dose en oncologie: est-ce en raison d'une suite en relation avec la perfection qu'est le nombre d'or?

Le problème est que plus on monte en paliers, plus les marches sont hautes et plus on risque des toxicités et ce, d'autant qu'on s'approche de la dose maximale tolérée. C'est pourquoi cette suite a été modifiée de façon à diminuer un peu la hauteur des marches des paliers élevés. Les paliers définis par la suite de Fibonacci et par la suite modifiée sont présentés dans le tableau.

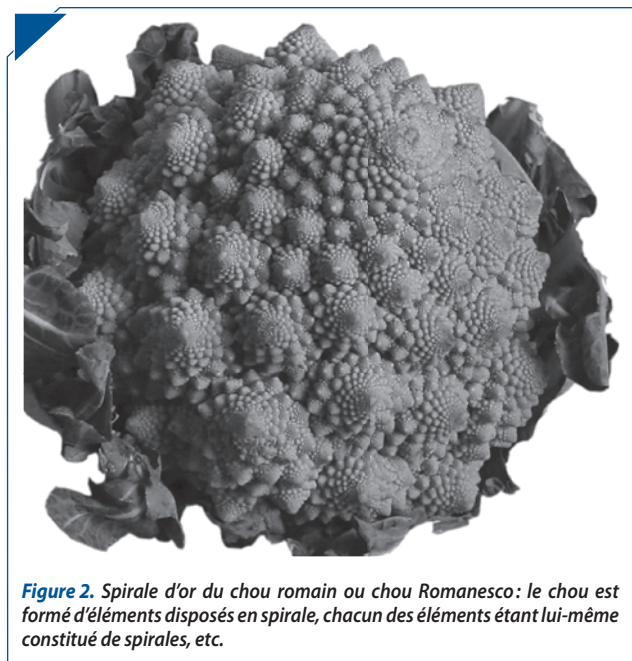


Figure 2. Spirale d'or du chou romain ou chou Romanesco: le chou est formé d'éléments disposés en spirale, chacun des éléments étant lui-même constitué de spirales, etc.

Tableau. Escalade de doses suivant la suite de Fibonacci, et la suite modifiée, compte tenu d'une dose de départ de 1 mg/m².

Palier	Série de Fibonacci		Série de Fibonacci modifiée	
	Dose (mg/m ²)	Incrément (%)	Dose (mg/m ²)	Incrément (%)
1	1	-	1	-
2	2	100	2	100
3	3	50	3,3	67
4	5	67	5	50
5	8	60	7	40
6	13	63	9	33
7	21	62	12	33
8	34	62	16	33

Toutefois, il faut savoir que, dans la littérature, cette modification varie avec les auteurs, et il est préférable, dans un protocole, de rapporter explicitement la série utilisée plutôt que de seulement indiquer "selon le schéma de Fibonacci modifié".

Le schéma standard prévoit d'inclure des cohortes de 3 patients par palier de dose, les doses étant le plus souvent exprimées en mg/m². L'escalade de dose est réalisée interpatients et aucune escalade n'est autorisée intrapatient. Le processus continue jusqu'à la survenue d'un pourcentage prédéfini d'une toxicité dose-limitante.

Généralement, la règle de décision d'un tel essai est la suivante (figure 3):

- ✓ si aucune TDL n'est observée chez les 3 patients de la cohorte, l'escalade continue;
- ✓ si une TDL est observée chez un seul patient de la cohorte, 3 patients supplémentaires sont inclus au même palier. Si aucun de ces 3 patients ne présente de toxicité, on continue l'escalade, sinon l'escalade est arrêtée;
- ✓ si on observe une TDL chez 2 ou 3 patients de la cohorte, on arrête l'escalade.

Une fois l'escalade arrêtée, 3 patients supplémentaires sont inclus à la dose précédant la dose d'arrêt. La dose maximale tolérée (DMT) est alors définie comme le palier précédant la dose d'arrêt de l'escalade et c'est souvent la dose recommandée pour les phases II subséquentes.

Même s'il n'y a aucune base scientifique sérieuse pour cette approche, l'expérience a montré qu'elle présentait peu de risques pour les patients. Cependant, ce schéma traditionnel, largement utilisé en oncologie, présente plusieurs inconvénients:

- ✓ le premier, d'ordre éthique, est que trop de patients reçoivent des doses infrathérapeutiques quand la dose de départ est loin de la dose maximale tolérée, d'autant que, l'escalade de doses étant interdite intrapatient, un patient qui entre au palier 1 (non thérapeutique) recevra toujours une dose non thérapeutique;
- ✓ chaque étape ne tient aucun compte de l'information obtenue aux précédents paliers;
- ✓ il est possible de ne jamais atteindre la DMT;
- ✓ si on atteint une DMT, elle peut être différente de la "vraie" DMT, car elle est obtenue avec une très mauvaise précision. En fait,

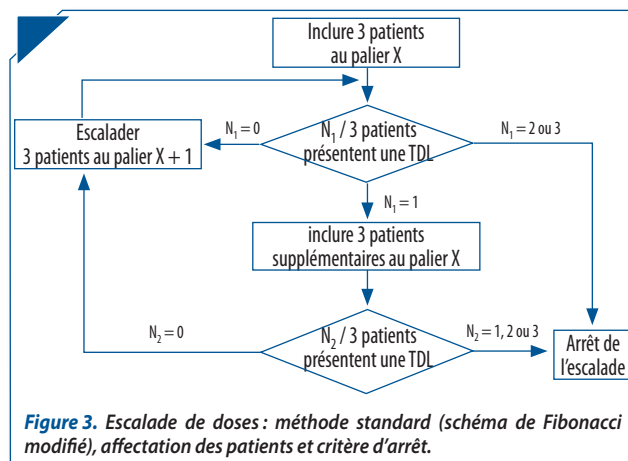


Figure 3. Escalade de doses: méthode standard (schéma de Fibonacci modifié), affectation des patients et critère d'arrêt.

la DMT obtenue avec de tels essais représente la dose pour laquelle le pourcentage de survenue d'une TDL est compris entre 15% et 70%. Ces pourcentages de survenue d'un événement toxique sont calculés sur seulement 3 patients d'une cohorte, ce qui est une hérésie statistique et peut expliquer pourquoi on a dû réviser les doses maximales tolérées dans de nombreuses études.

Une exploitation plus rationnelle et plus efficace de l'information disponible par le biais de méthodologies plus rigoureuses devrait permettre d'optimiser les essais de phase I. Cela devrait conduire à une meilleure évaluation de la relation dose-toxicité, et donc à une meilleure estimation de la dose recommandée pour les phases II, tout en diminuant le nombre de patients recevant un traitement insuffisant. L'amélioration devrait porter sur une dose de départ mieux appropriée, moins de patients par palier de dose et un schéma d'escalade plus efficace, tout en gardant à l'esprit la sécurité des patients.

Méthode d'escalade accélérée

La méthode d'escalade accélérée (*accelarated titration design, ATD*) est une méthode en deux étapes qui a été développée pour contourner certaines limites de la méthode selon un schéma de Fibonacci (4). Le principe est d'accélérer l'escalade pendant une première étape : un seul patient est inclus par palier de dose jusqu'à la survenue d'une TDL, et les doses sont doublées de palier en palier. Dès l'apparition d'une TDL ou de n'importe quelle toxicité de grade 2, on inclut respectivement 6 ou 3 patients à ce palier. La deuxième étape consiste à continuer l'escalade selon le schéma classique décrit plus haut. L'escalade de dose intrapatient est autorisée et se déroule comme suit : escalader en cas de toxicité de grade 0 ou 1 à la cure précédente, rester au même palier en cas de toxicité de grade 2, descendre de palier en cas de toxicité de grade supérieur à 2.

R. Simon et al. ont pu démontrer que cette approche réduisait de façon significative le nombre de patients traités à des doses infrathérapeutiques, du fait, d'une part, du nombre réduit de patients aux faibles paliers de dose pour chaque cohorte et, d'autre part, de l'escalade de doses chez un même patient, lui donnant ainsi une chance de recevoir une dose efficace (4). Ainsi, les sujets qui seraient sortis de l'étude pour cause de progression peuvent rester plus longtemps dans l'essai.

Un avantage non négligeable est une augmentation substantielle de l'information, particulièrement en ce qui concerne la toxicité cumulative. Cependant, cette escalade de dose plus agressive pouvant aussi être associée à un risque plus important pour le patient, il faudra porter un soin particulier à l'évaluation de ce qui est acceptable comme risque de toxicité.

R. Plummer et al. ont utilisé cette méthode lors d'un essai de phase I d'un analogue d'un taxane (5). L'étude incluait 53 patients, avec 153 cycles administrés (1 à 8 cycles par patient), dont un nombre non négligeable à des paliers plus élevés que les 3 premiers paliers, ce qui explique le taux inhabituel de réponses observées pour une phase I d'un cytotoxique. En outre, la quantité d'informations obtenues était plus importante au regard de la méthode classique en ce qui concernait les toxicités attendues à la dose préconisée pour la phase II.

Si cette méthode présente l'avantage de diminuer le nombre de patients sous-traités, elle ne répond pas au problème de la méthode standard concernant la détermination de la DMT obtenue avec une mauvaise précision. D'autres méthodes alternatives à la méthode classique ont été développées. Il s'agit de méthodes non paramétriques telles que les méthodes *up-and-down*, des méthodes mettant en œuvre des régressions isotoniques, ou encore des méthodes bayésiennes. Parmi ces dernières, la plus connue est la méthode par réévaluation continue.

Méthode par réévaluation continue (CRM)

La méthode par réévaluation continue (*continual reassessment method, CRM*) introduite par J. O'Quigley est la plus connue des approches bayésiennes dans le contexte des essais de phase I et a été largement publiée, dans sa forme originale ou modifiée (6). Cette méthode basée sur le calcul de la probabilité de survenue d'une toxicité à chaque palier se déroule en plusieurs étapes (figure 4) :

✓ avant que l'essai ne démarre, on construit une courbe dose-toxicité représentant la dose en fonction de la probabilité de survenue d'une toxicité, ou bien de la fréquence de survenue de cette toxicité. Cette courbe est établie à partir des données disponibles chez l'animal (plusieurs espèces étant généralement testées en préclinique) ou éventuellement à partir d'autres études de phase I déjà disponibles ;

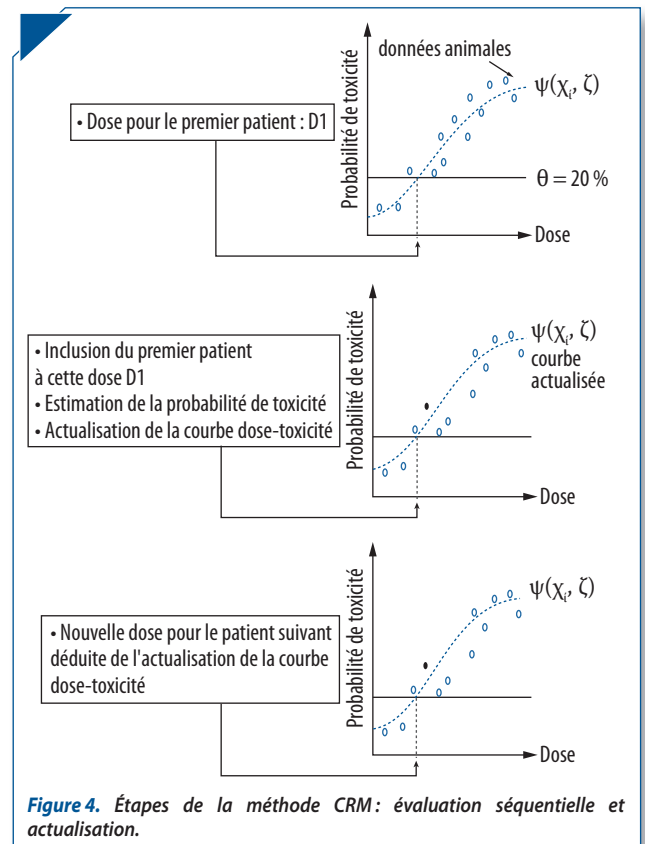


Figure 4. Étapes de la méthode CRM: évaluation séquentielle et actualisation.

✓ à partir de cette courbe, on détermine la dose de départ comme celle correspondant à une probabilité prédéfinie de toxicité inacceptable (généralement 20%). Ce pourcentage peut varier en fonction de la gravité de la toxicité encourue;

✓ on inclut 1 patient. Une fois évaluée la toxicité induite chez ce patient, on réajuste la courbe dose-toxicité par estimation bayésienne et on détermine la dose à administrer au patient suivant, qui correspond à la probabilité prédéfinie de toxicité inacceptable;

✓ on recommande jusqu'à obtenir une dose stable.

Toutefois, en raison de considérations éthiques, plusieurs modifications ont été proposées dans le but de réduire le risque pour un patient de recevoir une dose induisant une toxicité inacceptable. Elles portent sur le choix de la dose de départ, le nombre de patients à affecter par palier de dose et les critères d'arrêt. Parmi les principales modifications, S.N. Goodman et al. (7) suggèrent de ne pas utiliser la distribution a priori pour déterminer le niveau de dose initial mais plutôt le niveau le plus bas dans l'ensemble de doses; on parle alors de "modified CRM" (mCRM). En outre, ils proposent de démarrer une étude non pas avec un sujet par niveau de dose mais avec des cohortes de 2 ou 3 patients, et d'utiliser une procédure d'escalade de type CRM dès l'observation d'une TDL. Enfin, les augmentations de dose sont limitées par les niveaux de dose prédéterminés (c'est-à-dire que l'on ne permet pas de grandes escalades de dose).

Les méthodologies que nous avons détaillées sont univariées car elles portent uniquement sur des modèles dose-toxicité, mais il est intéressant d'introduire en phase I la notion d'efficacité. Ainsi, certains auteurs proposent de modifier la CRM en incluant la durée du suivi des patients ou bien en réalisant une modélisation simultanée ou séparée des relations dose-toxicité et dose-efficacité. Il s'agit alors de CRM bivariée (bCRM) [8-10].

Les avantages de la méthode CRM, univariée ou bivariée, sont multiples. À chaque instant, on utilise toute l'information préclinique disponible, ainsi que celle émanant d'autres études et de l'étude en cours, ce qui est un bénéfice majeur pour les patients. On réduit ainsi le nombre de patients qui reçoivent ainsi un traitement infra-thérapeutique. En outre, du fait même de la méthodologie, la DMT est obtenue avec son intervalle de confiance et non plus comme une valeur résultant de pourcentages calculés sur 3 ou 6 patients.

Malgré ses avantages, cette méthode est encore peu utilisée, notamment en raison de sa complexité nécessitant des biostatisticiens et des moyens de calcul élaborés. Les cliniciens ont l'habitude d'une méthode simple et considérée comme sûre pour les patients. Toutefois, devant le manque d'efficacité de la méthode selon la suite de Fibonacci modifiée et compte tenu des moyens statistiques dont disposent les centres pouvant mettre en place les essais de phase I en oncologie, les investigateurs devraient tirer profit des avancées dans ce domaine, comme le montre une étude de phase I du topotécan administré par voie intrapéritonéale ainsi que divers exemples rapportés par E.A. Eisenhauer et al. (11, 12).

Méthode d'escalade guidée par la pharmacocinétique

Il est largement reconnu que toutes les informations obtenues à partir des études précliniques peuvent être d'un apport substan-

tiel pour construire des essais de phase I qui soient informatifs et pertinents. Pourtant, il y a une donnée qui est souvent laissée de côté: l'information pharmacocinétique-pharmacodynamie (PK-PD), qui tient compte du fait que la variabilité de toxicité et/ou d'efficacité observée entre les patients est due pour une grande part à des différences dans le devenir du médicament après son administration (distribution dans les tissus, métabolisme, élimination). Une analyse PK-PD lors des essais devrait permettre:

✓ d'identifier des agents avec mécanismes saturables pour lesquels une escalade de dose peut induire une surexposition inattendue au médicament;

✓ d'identifier d'éventuels métabolites non observés lors des études précliniques ou l'apparition de nouveaux métabolites aux fortes doses;

✓ de mettre en évidence différentes covariables permettant l'identification des sujets à haut risque de toxicité;

✓ de transposer des résultats aux analogues du médicament, en incluant l'information dans les nouveaux essais;

✓ de simuler des essais cliniques pour optimiser le dessin des phases II (doses et schémas d'administration optimaux).

Concernant la planification de l'escalade de doses, partant de l'hypothèse que des concentrations plasmatiques identiques induisent le même effet biologique chez l'animal et chez l'homme, on procédera comme précédemment et, au lieu de déterminer une dose maximale tolérée, on cherchera une exposition maximale tolérée, évaluée par l'aire sous la courbe des concentrations plasmatiques (ASC), par une escalade classique, ou selon une CRM en remplaçant la courbe dose/effet par une courbe ASC/effet. Cette approche a été mise en œuvre pour différents cytotoxiques (13-17).

Combinaison de médicaments dans les essais de phase I

En oncologie, il est d'usage d'administrer des combinaisons de cytotoxiques. Parce que ces combinaisons peuvent entraîner des interactions médicamenteuses, telles que des synergies au niveau des toxicités, des essais de phase I doivent être réalisés pour déterminer la DMT de chaque agent administré au cours d'une même chimiothérapie. Dans ce cas, la situation est plus complexe qu'en monochimiothérapie, car il n'y a pas une seule solution au problème, mais un ensemble de plusieurs combinaisons de valeurs de DMT possibles.

Le plus souvent, pour chaque médicament, l'espace de posologies à administrer est établi en fonction des résultats des phases I conduites quand il est administré en monochimiothérapie. Puis, l'escalade de doses est planifiée. Divers schémas empiriques sont utilisés; ainsi, pour deux médicaments, on peut monter les doses du premier en gardant une dose fixe pour le second, ou bien escalader les doses alternativement ou simultanément.

E.L. Korn a proposé en 1993 un outil graphique pour définir l'escalade de doses dans le cas de combinaisons de deux agents (figure 5) [18]. Un diagramme de tolérabilité a été construit à partir de l'information disponible sur chaque médicament pour les différents types de toxicité concernés, en tenant compte de

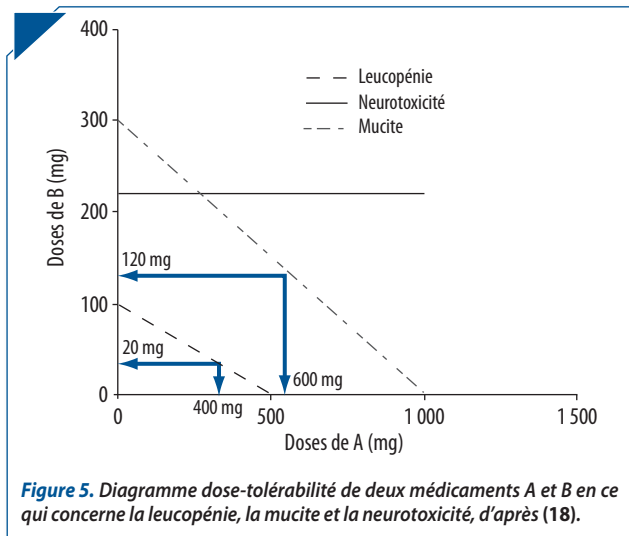


Figure 5. Diagramme dose-tolérabilité de deux médicaments A et B en ce qui concerne la leucopénie, la mucite et la neurotoxicité, d'après (18).

la dose maximale pour chaque agent au regard d'une toxicité donnée et en introduisant un facteur de pondération de toxicité relative de A par rapport à B. Les points au-dessus de chaque ligne représentent des combinaisons induisant une toxicité inacceptable. Dans cet exemple, 400 mg de A administrés avec 20 mg de B n'induisent aucun problème de toxicité, même si on ne prévoit pas de protection hématologique. En revanche, si on administre 600 mg de A et 120 mg de B, il faut prévoir des facteurs de croissance tels que le G-CSF pour prévenir la toxicité hématologique; concernant les autres toxicités, il n'y aura en principe aucun problème, sauf en cas de synergie de toxicité, celle-ci n'étant pas prise en compte lorsqu'on utilise ces fonctions linéaires. Cette méthode peut aussi être employée avec trois agents en utilisant un diagramme en trois dimensions. Une fois le diagramme dose-tolérabilité établi, les auteurs proposent divers schémas d'escalade de doses en fonction de l'activité relative des deux médicaments et en tenant compte d'éventuelles synergies au niveau de la toxicité.

Si cette approche est plus rationnelle que les démarches empiriques évoquées précédemment, les méthodes plus sophistiquées et plus efficaces proposées pour les monochimiothérapies peuvent être étendues à ces combinaisons. Comme pour un seul agent, on doit, dans un premier temps, établir des courbes dose-toxicité à partir des données des études de phase I réalisées en monochimiothérapie et, éventuellement, à partir des données de phase II et III disponibles, puis construire des isobologrammes. De tels plans expérimentaux ont été appliqués en phase I, dont un essai portant sur la capécitabine administrée en combinaison avec le cyclophosphamide et l'épirubicine, et un essai combinant gemcitabine et cyclophosphamide (19, 20).

Conclusion

La mise en place des essais de phase I selon l'approche standard de Fibonacci est empirique et sans réel fondement scientifique.

Ces essais exposent trop de patients à des doses infrathérapeutiques et manquent d'efficacité, car ils laissent de côté une énorme part de l'information disponible.

Plusieurs méthodes, basées sur des concepts scientifiques solides, sont maintenant disponibles. Si certaines semblent trop complexes ou paraissent, du point de vue éthique, présenter trop de risques pour les patients, elles peuvent toutefois être utilisées dans un premier temps:

- ✓ lors d'essais de phase I d'analogues de molécules bien connues;
- ✓ dans les chimiothérapies moins toxiques comme certaines thérapeutiques ciblées.

Cette étape permettrait aux investigateurs de se familiariser avec un outil performant. Pour des agents plus toxiques, une approche combinant la prudence et les concepts scientifiques de la CRM pourrait ainsi être mise en place, en concertation avec les cliniciens, les statisticiens et les pharmacologues.

ESSAIS DE PHASE II

Objectifs des essais de phase II

Les essais de phase II jouent un rôle important dans le processus de développement d'un nouveau traitement. En effet, seules les molécules ayant montré une efficacité antitumorale suffisante à l'issue de cette phase justifieront la poursuite de leur développement dans des essais comparatifs de phase III, essais à beaucoup plus grande échelle nécessitant de gros moyens. Les plans expérimentaux de phase II doivent donc être suffisamment performants pour éliminer rapidement les traitements inefficaces en utilisant au mieux les ressources disponibles, tout en minimisant le risque d'écarter une molécule ayant un effet antitumoral potentiellement intéressant.

Les essais de phase II ont très souvent été élaborés à l'aide de plans expérimentaux à un seul bras, sur un nombre limité de patients, et en utilisant un critère d'évaluation binaire observable à court terme (typiquement, la réponse tumorale). Récemment, de nouvelles méthodes ont été utilisées afin de rendre cette phase importante du développement encore plus efficiente.

Méthode générale (21)

Le plan d'expérience le plus simple utilisé en phase II est un essai à un seul bras (bras expérimental), où le taux de succès observés est comparé à un taux théorique. Les plans expérimentaux de phase II sont basés sur l'hypothèse suivante:

Soit π la probabilité réelle de succès du traitement expérimental:

- ✓ test de l'hypothèse nulle : $H_0 : \pi \leq p_0$
- ✓ versus l'hypothèse alternative : $H_1 : \pi \geq p_1$

Avec: p_0 : pourcentage en dessous duquel la molécule sera considérée d'efficacité insuffisante (inefficacité maximale);
 p_1 : pourcentage au-dessus duquel la molécule sera considérée d'efficacité suffisante pour justifier des études ultérieures (efficacité minimale).

La **figure 6** illustre les zones d'efficacité et d'inefficacité pour π . En pratique, p_0 et p_1 doivent être clairement définis avant le début de l'étude et servent à déterminer le nombre de patients qui devront être inclus dans la recherche afin de pouvoir conclure avec une puissance suffisante. Le pourcentage p_0 est classiquement choisi proche de la probabilité de succès du traitement standard; p_1 est supérieur à la probabilité de succès du traitement standard. Pour des raisons éthiques, les plans expérimentaux les plus couramment utilisés sont bâtis en étapes, chacune permettant d'arrêter précocement l'essai en l'absence d'efficacité et donc de limiter le nombre de patients traités par une substance inactive. Les plans les plus couramment utilisés sont présentés ci-après.

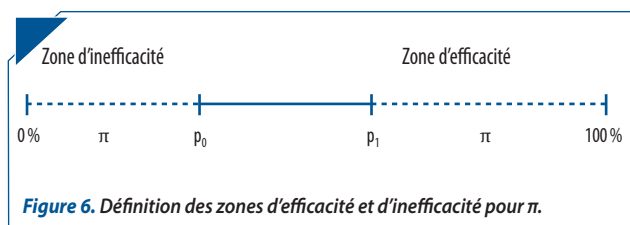


Figure 6. Définition des zones d'efficacité et d'inefficacité pour π .

● Plan en une étape: T.R. Fleming (22)

C'est le plan le plus simple. Le nombre de patients à inclure dans l'essai est défini à l'aide de tables, en fonction des valeurs de p_0 et p_1 fixées a priori, et des risques α (risque de conclure à l'efficacité d'une molécule qui est en fait inefficace) et β (risque de ne pas conclure à l'efficacité d'une molécule qui est en fait efficace). Une seule analyse est réalisée en fin d'essai, dans laquelle le nombre de succès observés (n_0) est comparé à un nombre de succès théoriques (n). Si $n_0 \geq n$, le traitement est considéré comme potentiellement efficace. Dans un essai de phase II, il peut être plus important de minimiser le risque β afin de ne pas rejeter une molécule à tort et de ne pas passer à côté d'une opportunité thérapeutique. On choisira ainsi plutôt $\beta = 5$ ou 10% . De même, il n'est pas toujours nécessaire de prendre un risque $\alpha = 5\%$, comme cela est préconisé en phase III. Un risque d'erreur $\alpha = 10\%$ peut ainsi permettre d'inclure moins de sujets, tout en maintenant dans des limites acceptables le risque d'exposer des patients à une molécule inefficace dans un essai de phase III ultérieur.

Exemple 1:

On considère le traitement A comme inintéressant pour un développement ultérieur si le taux de succès est inférieur à $p_0 = 35\%$, et comme intéressant si ce taux est supérieur ou égal à $p_1 = 50\%$.

Sous ces hypothèses, il est nécessaire d'inclure dans l'étude **72 patients**, en posant $\alpha = 10\%$ et $\beta = 10\%$ (on privilégie la puissance de l'essai). À l'issue du recrutement, le traitement A sera jugé efficace si au moins 31 patients sont non progressifs à 3 mois.

Ce plan d'expérience présente l'avantage d'être facile à mettre en œuvre. Sa principale limite réside dans le fait qu'aucun arrêt précoce d'étude n'est envisagé, ce qui risque d'exposer un nombre important de patients à une molécule inefficace. Ainsi, dans cet exemple, on pourrait s'apercevoir au moment de l'analyse qu'aucune réponse ni stabilité de la maladie à 3 mois n'avait été observée après l'inclusion de 42 patients, ce qui permettrait de conclure dès ce moment, et de manière définitive, à la non-efficacité du traitement et d'arrêter l'essai. C'est cet inconvénient qui a été pris en considération dans le développement de plans en plusieurs étapes.

● Plan multi-étapes: T.R. Fleming (22), E.A. Gehan (23), R. Simon (24), L.G. Ensign et al. (25)

Les plans multi-étapes sont basés sur une méthodologie identique à celle utilisée dans le plan à une seule étape, mais en y ajoutant des analyses intermédiaires planifiées avant le début de l'étude qui permettent de vérifier que le nombre de succès obtenus à l'issue de chaque étape est supérieur au nombre théorique calculé. Si cette condition est vérifiée à la fin d'une étape, on poursuit l'essai jusqu'à l'étape suivante; sinon, la molécule est considérée comme inefficace et l'essai est définitivement arrêté.

Exemple 2 (26):

Il s'agit d'évaluer l'intérêt de l'ajout de thalidomide au docétaxel pour améliorer la survie des patients atteints de cancer de la prostate métastatique. Le nombre de sujets est calculé en utilisant un plan en 2 étapes de R. Simon (plan "minimax").

Les hypothèses sont les suivantes: $H_0: p_0 \leq 25\%$ et $H_1: p_1 \geq 45\%$ avec $\alpha = 5\%$ et $\beta = 10\%$.

On obtient alors le schéma d'étude suivant:

– Étape 1:

✓ si moins de 7 réponses sont observées sur les 26 premiers patients inclus, l'étude est arrêtée et le traitement est considéré comme inefficace;

✓ si 7 réponses ou plus sont observées sur les 26 premiers patients inclus, l'étude est poursuivie et 23 patients supplémentaires sont inclus afin d'arriver à la deuxième étape.

– Étape 2:

✓ si moins de 18 réponses histologiques sont observées sur les 49 patients inclus, le traitement est considéré comme inefficace;

✓ si 18 réponses histologiques ou plus sont observées sur les 49 patients inclus, la molécule est considérée comme suffisamment efficace pour que son développement soit poursuivi dans une phase postérieure.

La probabilité d'arrêter l'étude pour inefficacité du traitement expérimental peut être calculée à chaque étape. Ainsi, dans cet exemple, en cas d'inefficacité de l'association thérapeutique, avec $\pi = 25\%$, l'essai a 52% de chances d'être arrêté dès la première étape. Il existe également des plans en 3 étapes basés sur le même principe.

La différence entre les plans de T.R. Fleming, E.A. Gehan, R. Simon ou L.G. Ensign réside dans l'optimisation en termes

de nombre de sujets à inclure lors de chaque étapes. Le plan de E.A. Gehan, par exemple, est optimisé pour arrêter un essai le plus vite possible si aucun succès n'apparaît chez les premiers patients. En cas de poursuite, le nombre de patients à inclure dans l'étape 2 tient compte des résultats de l'étape 1. Pour le plan multi-étapes de T.R. Fleming, la répartition du nombre de sujets par étape est faite de manière arbitraire. Ce plan a surtout la particularité de permettre l'arrêt à la fin de l'étape 1 pour inefficacité ou efficacité précoce. Dans ce dernier cas, l'étude est arrêtée et la molécule sera jugée suffisamment efficace pour être étudiée dans des essais de phase III. R. Simon a développé deux types de plans, dans lesquels le nombre de patients par étape est défini par la méthode. Le plan "optimum" minimise le nombre moyen de sujets exposés à un traitement inefficace, alors que le plan "minimax" minimise le nombre total de sujets à inclure dans l'essai. Enfin, le plan de L.G. Ensign est un plan en 3 étapes qui combine la méthode de E.A. Gehan pour l'étape 1 et celle de R. Simon pour les étapes 2 et 3.

Une des limites posées par ces plans en étapes est liée au critère de jugement binaire (réponse au traitement), qui n'est pas toujours adapté à la problématique. Pour les nouvelles molécules cytostatiques (thérapeutiques dites "ciblées") qui permettent un retard à la progression tumorale sans pour autant provoquer de réponse tumorale, l'utilisation de ces plans est difficile et pose plus largement le problème de l'inadéquation des critères d'évaluation des drogues cytotoxiques fondés sur la réduction du volume tumoral. Par ailleurs, le délai nécessaire entre l'inclusion et l'évaluation d'un patient pose parfois problème dans les plans multi-étapes de phase II, avec le dilemme suivant : soit stopper temporairement les inclusions une fois le nombre de patients du palier atteint, au risque d'avoir du mal à reprendre une fois que le passage à l'étape supérieure est effectif, soit prendre le risque que tous les patients soient déjà inclus au moment des résultats de la première étape, ce qui remet en cause l'intérêt de ce type de plan. Enfin, ces plans expérimentaux présentent deux faiblesses : d'une part, l'impossibilité de prendre en compte la toxicité du traitement évalué dans les règles de décision, et ce bien que la toxicité soit un critère quasiment systématiquement étudié dans les essais de phase II ; d'autre part, l'impossibilité de prendre en compte un éventuel biais de sélection du fait de l'absence de bras contrôle.

Essais de phase II randomisés (27, 28)

L'introduction d'un bras contrôle dans les essais de phase II a été guidée par la difficulté de l'interprétation des résultats. En effet, un traitement est jugé inefficace dans un essai de phase II par confrontation d'un taux de réponse à un taux théorique qui peut être assimilé à une comparaison historique. Or, rien ne permet d'affirmer que le risque de base des patients inclus dans l'essai était comparable à celui des patients ayant permis de déterminer le taux de référence, et ce même si les critères d'inclusion/exclusion semblaient comparables. Ainsi, à l'issue d'un essai de phase II sans bras contrôle, un faible taux de réponse peut être lié à une réelle inefficacité du traitement expérimental, ou bien à la sélection de patients avec une faible probabilité de réponse. L'introduction d'un bras de référence jouant le rôle de

bras contrôle, ainsi que la répartition des patients entre traitement expérimental et traitement de référence par tirage au sort permettent de résoudre ce problème en limitant les effets des biais de sélection. En effet, le taux de réponse observé dans le bras de référence sert alors à valider la mesure de celui observé dans le bras expérimental. Il faut bien avoir à l'esprit que ce type de plan expérimental n'est en aucun cas prévu pour comparer directement les effets des traitements entre les deux bras comme dans un essai de phase III, mais seulement pour renforcer la valeur des conclusions sur l'efficacité potentielle du traitement expérimental si, à l'issue de l'essai, le taux de réponse observé dans le bras de référence s'avère proche de celui attendu.

Le calcul du nombre de sujets est réalisé, pour le bras testé, de la même façon que pour un essai de phase II classique. Le nombre de patients à inclure dans le bras contrôle est fixé de manière que l'effectif des deux bras soit égal. On peut également envisager de mettre moins de patients dans le bras de référence (randomisation 1:2), qui sert seulement à vérifier l'hypothèse p_0 avec une précision correcte.

Exemple 3 (29) :

Il s'agit d'une étude randomisée de phase II de l'EORTC évaluant brostallicine versus doxorubicine en chimiothérapie de première ligne chez des patients atteints de sarcome des tissus mous avancé ou métastatique. Cent huit patients ont été inclus, avec une randomisation 2:1 : un tiers des patients dans le bras standard (doxorubicine) et deux tiers dans le bras expérimental (brostallicine). Le critère de jugement principal était le taux de non-progression à 6 mois. Le calcul du nombre de sujets a été défini par un plan de T.R. Fleming en 1 étape ($p_0 = 40\%$, $p_1 = 55\%$, $\alpha = \beta = 10\%$), ce qui devait permettre de conclure à l'efficacité de la brostallicine si 35 des 72 patients inclus se révélaient non progressifs. Le bras doxorubicine (36 patients) permettait de valider le taux de réponse obtenu avec la brostacilline ; aucun test statistique de comparaison formelle n'a été réalisé.

Essai avec prise en compte de la toxicité (30, 31)

La deuxième critique faite aux essais classiques de phase II est la non-prise en compte de la toxicité dans les règles de décision. Plusieurs plans ont été développés pour pallier ce problème. Le critère de jugement principal est défini afin de juger de l'efficacité, et un critère secondaire permet de juger de la toxicité. Des règles d'arrêt combinant taux de réponse et taux de toxicité sont ensuite déterminées.

Exemple 4 (30) :

Un essai de phase II basé sur un plan d'expérience à n étapes est conçu pour évaluer la chimiothérapie à forte dose dans le lymphome malin non hodgkinien. Le traitement est déclaré prometteur si son taux de réponse (R) est supérieur à celui du traitement standard (R_0) et que son taux de toxicité (T) est inférieur à celui du traitement standard (T_0).

Les hypothèses H_0 et H_1 peuvent alors être écrites sous la forme suivante :

$H_0 : PR \leq PR_0$ ou $PT \geq PT_0$; $H_1 : PR > PR_0$ et $PT < PT_0$.

Ces hypothèses peuvent être représentées comme sur la **figure 7**, avec les probabilités d'événements définies a priori suivantes :

- ✓ PR_0 : probabilité de réponse du traitement standard : 0,5 ;
- ✓ PT_0 : probabilité de toxicité du traitement standard : 0,3 ;
- ✓ $PT_0|PR_0$: probabilité de réponse en cas de toxicité : 0,35 ;
- ✓ PR : probabilité de réponse attendue pour le nouveau traitement : 0,75 ;
- ✓ PT : probabilité de toxicité attendue pour le nouveau traitement : 0,15.

Il faut également définir les risques de première et deuxième espèces :

- α_R : $(PR \leq PR_0 \text{ et } PT \geq PT_0) \leq 5\%$
- α_T : $(PR \leq PR_0 \text{ ou } PT \geq PT_0) \leq 30\%$
- $1 - \beta$: $(PR > PR_0 \text{ et } PT < PT_0) > 90\%$

Dans cet exemple, M.R. Conaway et G.R. Petroni utilisent un algorithme qui, à partir de toutes ces hypothèses, minimise la taille de l'essai et détermine les règles d'arrêt à chacune des étapes.

Plan bayésien (32-34)

L'approche séquentielle de type bayésien commence à être appliquée en cancérologie aux plans de phase II. Elle est basée sur le principe selon lequel l'hypothèse de départ n'est pas figée mais évolue en fonction des données observées. De multiples analyses sont réalisées en cours d'étude (tous les $n = 1, 2, 3, 4, 5 \dots$ patients), chacune faisant évoluer la loi de probabilité du critère principal à partir des informations concernant les premiers patients. En début d'essai, la réponse au traitement est considérée comme une variable aléatoire ayant une distribution de probabilité $P(\theta)$ [*prior*] peu discriminante. Les données observées au cours de l'essai permettent d'actualiser $P(\theta)$ pour obtenir une distribution de probabilité a posteriori (*posterior*) de la réponse au traitement, qui devient de plus en plus informative au fur et à mesure de l'avancement de l'essai.

Ces plans ont généralement l'avantage de réduire le nombre de patients à inclure, d'autant plus que le traitement est efficace.

Le nombre de sujets à inclure n'est pas déterminé de manière définitive avant le début de l'essai mais dépend des données recueillies au cours de l'essai.

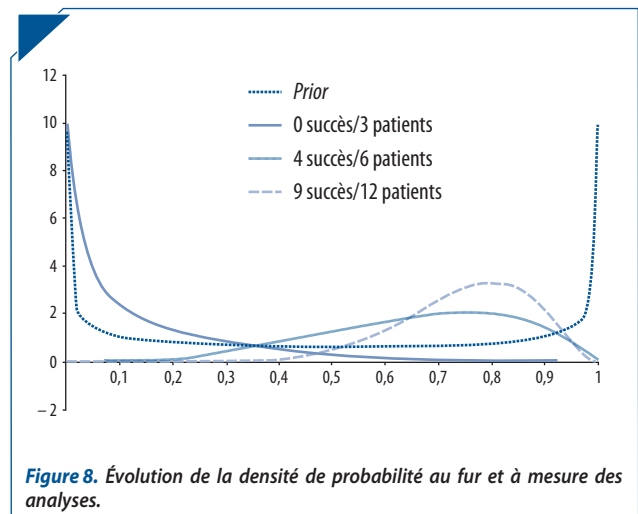
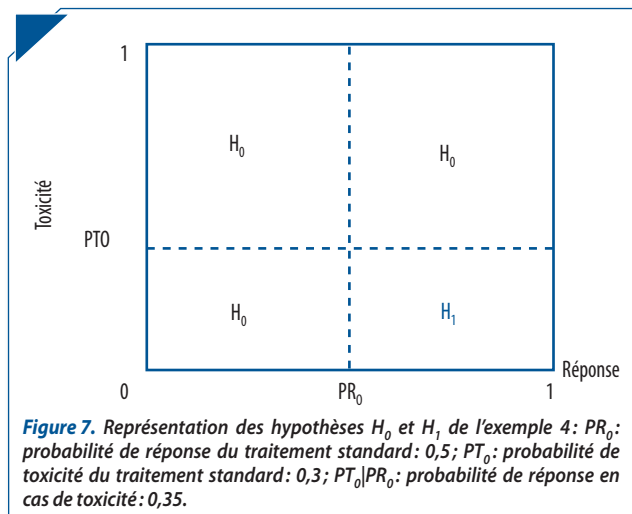
Exemple 5 (figure 8) :

La distribution de probabilité du taux de succès (par exemple, la réponse tumorale) évolue à chaque analyse. La courbe *prior* est non informative. Elle se traduit par une probabilité identique d'obtenir 0, 1, 2, ... ou n succès. Au fur et à mesure des analyses, cette densité de probabilité évolue vers une courbe de plus en plus concentrée vers la vraie valeur d'efficacité.

Les limites de ce type de plans sont parfois dues au fait que le *prior* est difficile à estimer ou au fait que le *posterior* devient trop vite informatif, et que les conclusions de l'étude sont tirées trop hâtivement (l'étude est arrêtée si le premier patient est en échec). Il peut donc être intéressant de définir un nombre minimal de patients à inclure. D'autre part, ces méthodes restent intéressantes si les analyses sont réalisées en temps réel, ce qui nécessite un monitoring serré des données d'autant plus difficile que l'étude est multicentrique, ainsi que des moyens de calcul plus complexes à mettre en œuvre que pour un plan de phase II classique, du fait de la répétition des analyses et des méthodes statistiques.

ESSAIS DE PHASE II-III (35-37)

Il se passe quelquefois un certain laps de temps entre la fin d'une étude de phase II ayant montré une efficacité antitumorale et le début d'une étude de phase III. Le principe des plans de phase II-III est justement de réduire ce temps entre essais en utilisant les données des patients inclus dans la phase II pour l'essai de phase III. Cette approche est possible en utilisant les critères d'arrêt précoce issus des essais de phase II dans des essais de phase III, ce qui pourrait être assimilable au principe d'une analyse intermédiaire. Un des intérêts de ce type de plan est de réaliser un essai de phase II avec n bras expérimentaux et de



sélectionner le plus efficace en tant que bras expérimental dans un essai comparatif de phase III.

Cette approche est d'autant plus intéressante si les critères d'efficacité utilisés pour les phases II et III sont fortement corrélés, voire identiques, ce qui n'est pas toujours le cas. Très souvent, en effet, la réponse tumorale utilisée comme critère d'efficacité en phase II est abandonnée au profit de la survie globale ou sans progression en phase III.

Conclusion

La phase II joue un rôle très important dans le cycle de développement d'une molécule, et les plans expérimentaux pouvant être mis en œuvre sont multiples. Le choix du meilleur plan devra tenir compte du potentiel de recrutement dans un temps donné. La définition des risques d'erreur α et β est importante, la priorité devant être donnée à la puissance si on ne souhaite pas passer à côté d'une molécule potentiellement efficace. Enfin, l'ajout d'une randomisation avec un bras contrôle de référence, très utile pour valider les résultats obtenus, ne permet en aucun cas de comparer directement les bras expérimentaux et contrôle avec une puissance suffisante.

ESSAIS DE PHASE III

Objectifs des essais de phase III

L'essai de phase III est l'étude comparative d'efficacité proprement dite. Son but est de tester l'efficacité d'un nouveau traitement par rapport au meilleur traitement connu, qui est le traitement de référence. C'est dans cette phase de développement que le rapport bénéfice/tolérance est estimé en situation réelle. Les études de phase III sont de grande envergure. Elles impliquent une centaine à plusieurs milliers de patients. Par rapport à la phase II, les durées d'exposition au médicament sont prolongées et une évaluation des interactions médicamenteuses peut être effectuée. C'est à l'issue de cette phase qu'une autorisation de mise sur le marché (AMM) pourra être donnée par les autorités de santé.

Si les objectifs d'un essai de phase III sont bien définis, la méthodologie reste basée sur la minimisation des biais (38). Cela conduit à réaliser des essais prospectifs, comparatifs, randomisés, si possible en double aveugle, avec un critère principal unique défini clairement, ainsi qu'à une analyse statistique réalisée en intention de traiter.

Randomisation

Si le principe de la randomisation n'est pas encore systématique en phase II, il est obligatoire dans un essai de phase III. En effet, le but de cette phase étant de comparer directement les traitements testés entre eux, il convient de s'assurer que la constitution des groupes est réalisée sans biais. La randomisation permet d'assurer cette attribution de manière aléatoire.

Différents niveaux de connaissance du traitement par le clinicien ou par le patient permettent également d'exclure totalement les biais au niveau de leur estimation de l'effet du traitement :

- ✓ essai en double aveugle: la nature du traitement alloué par randomisation est inconnue du médecin et du patient (essai à fort niveau de preuve);
- ✓ essai en simple aveugle: la nature du traitement alloué par randomisation est connue soit par le médecin, soit par le patient (essai à niveau de preuve intermédiaire);
- ✓ essai en ouvert: la nature du traitement alloué par randomisation est connue du médecin et du patient (niveau de preuve plus faible).

En cancérologie, les essais en double aveugle sont rares. En effet, les traitements testés ont très souvent des effets secondaires caractéristiques, ce qui rend le double aveugle impossible. Cette limite peut être compensée par l'utilisation d'un critère de jugement non subjectif (mortalité globale).

Critère principal

C'est le paramètre unique sur lequel on va baser la conclusion de l'essai. Ce critère doit donc être défini très précisément et doit correspondre au critère clinique le plus pertinent vis-à-vis de l'objectif de l'étude. C'est sur ce critère que le nombre de patients à inclure dans l'essai est calculé et c'est donc lui seul qui garantit la puissance requise pour pouvoir conclure. Un effort particulier doit donc être fait pour la qualité du recueil des données et pour ne pas perdre de données sur ce critère. Contrairement aux essais de phase II, dans lesquels le critère de jugement est très souvent un critère binaire rapidement observable, les essais de phase III utilisent plutôt un critère clinique. En cancérologie, le critère principal est le plus souvent défini comme la survie globale ou la survie sans progression.

À côté du critère de jugement principal, d'autres facteurs d'intérêt peuvent être étudiés: ce sont les critères secondaires. La puissance statistique n'est pas garantie sur ces critères et peut donc être faible.

Types d'essais

L'essai de phase III est mis en œuvre pour répondre à une question précise. L'objectif peut alors être de trois ordres:

- ✓ montrer la supériorité du nouveau traitement par rapport au traitement standard;
- ✓ montrer l'équivalence du nouveau traitement par rapport au traitement de référence;
- ✓ montrer la non-infériorité du nouveau traitement par rapport au traitement de référence.

Chacun de ces trois types d'essais a une méthodologie qui lui est propre, et un essai conçu pour montrer la supériorité du nouveau traitement par rapport au traitement de référence ne peut finalement conclure à une non-infériorité, et inversement.

● Essai de supériorité ou d'efficacité

Cela concerne la majorité des essais. Les hypothèses sont définies comme suit:

- ✓ H_0 : hypothèse nulle, à rejeter: le nouveau traitement est inférieur ou égal au traitement de référence;
- ✓ H_1 : hypothèse alternative: le nouveau traitement est supérieur au traitement de référence.

Lorsque l'hypothèse H_0 ne peut pas être rejetée, l'essai est non concluant. L'absence de supériorité ne permet pas de conclure à une équivalence ou à une non-infériorité.

● Essai d'équivalence ou de non-infériorité

Ce type d'essai est de plus en plus fréquent. Il peut servir à comparer deux formes galéniques, deux rythmes d'administration, ou encore deux molécules appartenant à une même classe thérapeutique. On peut également ne pas être intéressé par une efficacité supérieure mais par une meilleure tolérance, une plus grande facilité d'utilisation ou un coût plus faible.

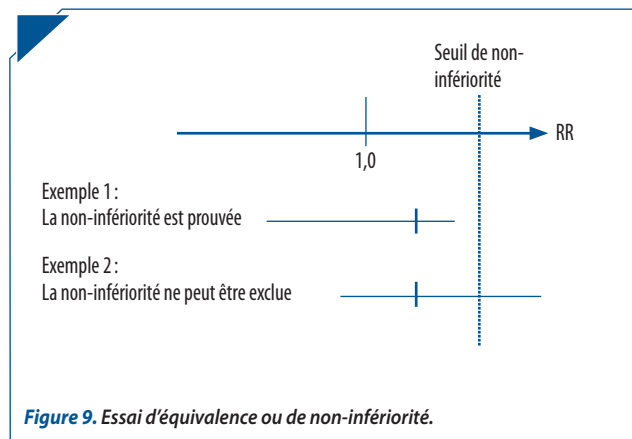
La méthodologie statistique mise en œuvre nécessite de déterminer a priori une zone d'équivalence. La valeur-seuil de cette zone correspond à la perte d'efficacité maximale à laquelle on consent avec le nouveau traitement. Le choix de ce seuil est donc crucial dans la mise en œuvre d'un tel essai. Le nombre de patients à inclure est d'autant plus grand que cette zone d'équivalence est réduite.

La comparaison est ensuite basée sur le test d'hypothèse :

- ✓ En situation bilatérale : équivalence
 - H_0 : les traitements ne sont pas équivalents ;
 - H_1 : les traitements sont équivalents.
- ✓ En situation unilatérale : non-infériorité
 - H_0 : le nouveau traitement est inférieur au traitement de référence ;
 - H_1 : le nouveau traitement est égal ou supérieur au traitement de référence.

En pratique, l'utilisation d'un test bilatéral n'a que peu d'intérêt puisqu'on court le risque de rejeter l'hypothèse d'équivalence si le nouveau traitement est supérieur au traitement de référence. Le plus souvent, c'est donc la non-infériorité qui intéresse le clinicien, c'est-à-dire que le nouveau traitement soit au moins aussi efficace que le traitement de référence.

La méthodologie statistique mise en œuvre dans ce type d'essai consiste à construire un intervalle de confiance autour du risque relatif et à comparer la borne supérieure de cet intervalle avec la limite de non-infériorité choisie, ce qui est illustré sur la **figure 9**.



Plan d'expérience

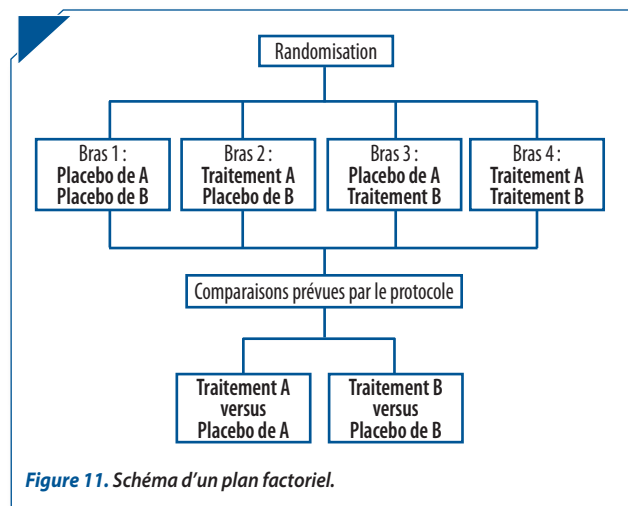
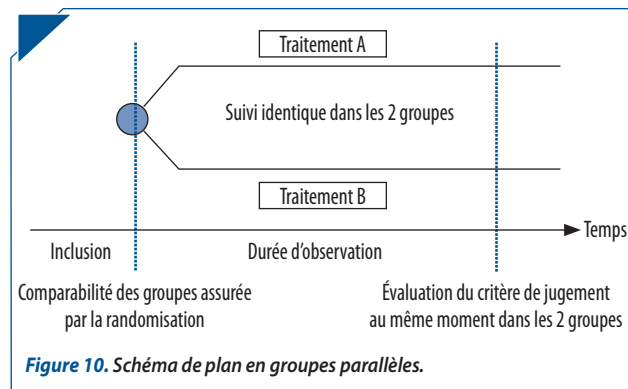
Les plans d'expérience utilisés en phase III sont construits de manière à limiter l'exposition aux biais. Trois grands types de plans expérimentaux sont utilisés : les plans en groupes parallèles, les plans factoriels et les plans croisés.

● Plan en groupes parallèles

C'est le plan le plus utilisé. Le traitement testé est comparé au traitement de référence en utilisant deux groupes de patients rendus comparables avant traitement par randomisation. Les patients étant suivis et évalués de manière identique, la différence observée en fin d'essai ne peut être due qu'au seul effet du traitement. Ce schéma d'étude est facilement extrapolable à plus de deux groupes de traitement, sous réserve de tenir compte des problèmes statistiques posés par les comparaisons multiples. Selon le bénéfice que l'on souhaite montrer, ce plan peut nécessiter d'inclure jusqu'à plusieurs milliers de patients. La **figure 10** illustre ce plan d'expérience.

● Plan factoriel

Ce plan, schématisé sur la **figure 11**, n'est pas forcément plus compliqué à mettre en œuvre qu'un plan parallèle et permet



de répondre à deux questions avec un même essai, donc avec un effectif plus faible. Chaque participant à l'essai entre dans l'estimation des deux comparaisons d'intérêt. En pratique, la randomisation répartit les patients en quatre groupes :

- ✓ bras 1 : les patients reçoivent le placebo de A et le placebo de B;
- ✓ bras 2 : les patients reçoivent le traitement A et le placebo de B;
- ✓ bras 3 : les patients reçoivent le placebo de A et le traitement B;
- ✓ bras 4 : les patients reçoivent le traitement A et le traitement B.

Deux comparaisons sont réalisées au moment de l'analyse statistique. La première sert à évaluer les résultats chez les patients recevant le traitement A (bras 2 + 4) par rapport aux patients ne le recevant pas (bras 1 + 3); la deuxième permet d'évaluer les patients recevant le traitement B (bras 3 + 4) par rapport à ceux qui n'en reçoivent pas (bras 1 + 2). Il n'est en aucun cas prévu de comparer les quatre bras de traitement entre eux.

Ce plan d'expérience n'est valide que s'il n'existe pas d'interaction entre les traitements A et B, c'est-à-dire que A et B agissent de manière indépendante. La première étape de l'analyse statistique consiste donc à vérifier cette hypothèse. En cas d'interaction avérée, le regroupement des bras n'est plus envisageable. L'analyse doit alors porter sur trois comparaisons: bras 2 versus bras 1, bras 3 versus bras 1 et bras 4 versus bras 1. Malheureusement, l'effectif n'étant pas prévu pour ces comparaisons, la puissance de l'étude est très faible.

● Plan croisé

L'avantage d'un tel plan réside dans le fait que les patients reçoivent chaque bras de traitement dans un ordre aléatoire. Chaque sujet peut donc être pris comme son propre témoin. L'essai est divisé en périodes et une séquence de traitement, A suivi de B ou B suivi de A, est attribuée par randomisation à chaque patient (figure 12). Le critère de jugement est mesuré à la fin de chaque période, ce qui permet de comparer l'effet des traitements chez un même patient. Une période dite "de lavage" doit être prévue entre 2 périodes de traitement successives afin qu'aucun effet rémanent du traitement de la première période ne subsiste

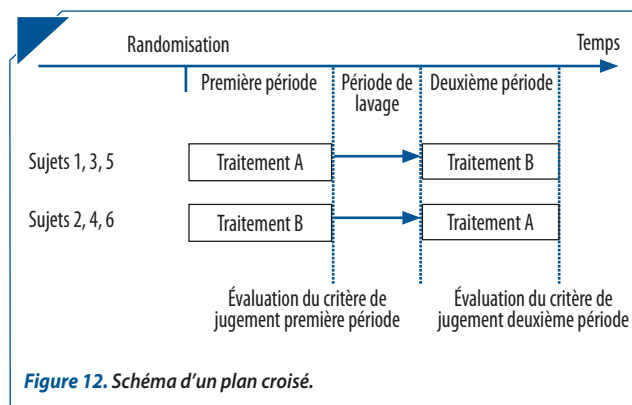


Figure 12. Schéma d'un plan croisé.

lorsque la deuxième période débute. L'analyse statistique porte sur la différence observée pour chaque patient entre les deux périodes. En l'absence d'effet du traitement, cette différence fluctue autour de 0; sinon, la distribution de ce paramètre est centrée sur la vraie valeur de l'effet du traitement.

Les avantages d'un tel plan d'expérience résident dans la forte comparabilité des groupes de traitement, chaque patient appartenant à chaque groupe, ce qui, en d'autres termes, rend ce type d'essai beaucoup plus puissant qu'un essai en groupes parallèles. En revanche, certaines limites rendent ce plan difficile à mettre en œuvre en cancérologie. En effet, il a une durée de réalisation importante puisque chaque patient est soumis à plusieurs périodes de traitement, ce qui nécessite que l'état du patient ne se dégrade pas trop vite afin que les périodes soient comparables. De plus, le critère de jugement doit être réversible afin d'être mesuré à la fin de chaque période, ce qui exclut totalement de travailler sur la survie globale.

Conclusion

Les essais de phase III offrent une grande variété en termes de plan d'expérience, ce qui permet de choisir le schéma d'étude le mieux adapté pour répondre à la question posée. Une importante méthodologie est disponible afin de conduire à une conclusion non biaisée de l'effet du traitement. Un budget important pouvant être nécessaire pour mener à bien un essai de phase III, il convient que cet investissement permette de conclure de manière fiable à l'efficacité d'un nouveau traitement.

CONCLUSION GÉNÉRALE

La méthodologie des essais lors du développement des médicaments a beaucoup évolué ces dernières années, notamment avec l'application des méthodes bayésiennes. Les phases de développement qui ont le plus bénéficié de cet apport sont en premier lieu les essais de phase I, puis viennent les essais de phases II et II-III, les essais de phase III restant bien codifiés. Il est à noter que le passage d'une phase à l'autre aurait tout à gagner au travers de simulations réalisées à partir des résultats de la phase précédente pour en optimiser le dessin. En effet, lors de la détermination d'un plan d'expérience, il serait judicieux de déterminer non seulement la meilleure dose à administrer mais aussi le meilleur schéma d'administration. Étant donné les nombreuses possibilités dans le choix du schéma, il serait judicieux de déterminer celui-ci ou l'ensemble de schémas à tester sur des bases scientifiques et non empiriques.

RÉFÉRENCES BIBLIOGRAPHIQUES

1. Sheiner LB. Learning versus confirming in clinical drug development. *Clin Pharmacol Ther* 1997;61:275-91.
2. Weiner LM, Padavic-Shaller K, Kitson J et al. Phase I evaluation of combination therapy with interleukin 2 and gamma-interferon. *Cancer Res* 1991;51:3910-8.
3. Sandler A. Clinical experience with the HER1/EGFR tyrosine kinase inhibitor erlotinib. *Oncology* 2003;17:17-22.

4. Simon R, Freidlin B, Rubinstein L et al. Accelerated titration designs for phase I clinical trials in oncology. *J Natl Cancer Inst* 1997;89:1138-47.
5. Plummer R, Ghilmini M, Calvert P et al. Phase I and pharmacokinetic study of the new taxane analog BMS-184476 given weekly in patients with advanced malignancies. *Clin Cancer Res* 2002;8:2788-97.
6. O'Quigley J, Pepe M, Fisher L. Continual reassessment method: a practical design for phase I clinical trials in cancer. *Biometrics* 1990;46:33-48.
7. Goodman SN, Zahurak ML, Piantadosi S. Some practical improvements in the continual reassessment method for phase I studies. *Stat Med* 1995;14:1149-61.
8. Cheung YK, Chappell R. Sequential designs for phase I clinical trials with late-onset toxicities. *Biometrics* 2000;56:1177-82.
9. Thall PF, Estey EH, Sung HG. A new statistical method for dose-finding based on efficacy and toxicity in early phase clinical trials. *Invest New Drugs* 1999;17:155-67.
10. Braun TM. The bivariate continual reassessment method; extending the CRM to phase I trials of two competing outcomes. *Control Clin Trials* 2002;23:240-56.
11. Plaxe SC, Christen RD, O'Quigley J et al. Phase I and pharmacokinetic study of intraperitoneal topotecan. *Invest New Drugs* 1998;16:147-53.
12. Eisenhauer EA, O'Dwyer PJ, Christian M, Humphrey JS. Phase I clinical trial design in cancer drug development. *J Clin Oncol* 2000;18:684-92.
13. Collins JM, Zaharko DS, Dedrick RL, Chabner BA. Potential roles for preclinical pharmacology in phase I clinical trials. *Cancer Treat Rep* 1986;70:73-80.
14. Ploin DY, Tranchand B, Guastalla JP et al. Pharmacokinetically guided dosing for intravenous melphalan: a pilot study in patients with advanced ovarian adenocarcinoma. *Eur J Cancer* 1992;28A:1311-5.
15. Dees EC, Whitfield LR, Grove WR et al. A phase I and pharmacologic evaluation of the DNA intercalator CI-958 in patients with advanced solid tumors. *Clin Cancer Res* 2000;6:3885-94.
16. Gianni L. Learning from CI-941 about pharmacokinetically guided dose escalation. *Eur J Cancer* 1992;28A:1302-4.
17. Van Kesteren C, Mathot RA, Beijnen JH, Schellens JH. Pharmacokinetic-pharmacodynamic guided trial design in oncology. *Invest New Drugs* 2003;21:225-41.
18. Korn EL, Simon R. Using the tolerable-dose diagram in the design of phase I combination chemotherapy trials. *J Clin Oncol* 1993;11:794-801.
19. Morita S, Toi M, Kobayashi T et al. Application of a continual reassessment method to a phase I clinical trial of capecitabine in combination with cyclophosphamide and epirubicin (CEX) for inoperable or recurrent breast cancer. *Jpn J Clin Oncol* 2004;34:104-6.
20. Thall PF, Millikan RE, Mueller P, Lee SJ. Dose-finding with two agents in Phase I oncology trials. *Biometrics* 2003;59:487-96.
21. Kramar A, Potvin D, Hill C. [Research design for patient selection in clinical trials phase II]. *Rev Epidemiol Sante Publique* 1996;44:364-71.
22. Fleming TR. One-sample multiple testing procedure for phase II clinical trials. *Biometrics* 1982;38:143-51.
23. Gehan EA. The determination of the number of patients required in a preliminary and a follow-up trial of a new chemotherapeutic agent. *J Chronic Dis* 1961;13:346-53.
24. Simon R. Optimal two-stage designs for phase II clinical trials. *Control Clin Trials* 1989;10:1-10.
25. Ensign LG, Gehan EA, Kamen DS, Thall PF. An optimal three-stage design for phase II clinical trials. *Stat Med* 1994;13:1727-36.
26. Dahut WL, Gulley JL, Arlen PM et al. Randomized phase II trial of docetaxel plus thalidomide in androgen-independent prostate cancer. *J Clin Oncol* 2004;22:2532-9.
27. Lee JJ, Feng L. Randomized phase II designs in cancer clinical trials: current status and future directions. *J Clin Oncol* 2005;23:4450-7.
28. Wieand HS. Randomized phase II trials: what does randomization gain? *J Clin Oncol* 2005;23:1794-5.
29. Leahy M, Ray-Coquard I, Verweij J et al. Brostallicin, an agent with potential activity in metastatic soft tissue sarcoma: a phase II study from the European Organisation for Research and Treatment of Cancer Soft Tissue and Bone Sarcoma Group. *Eur J Cancer* 2007;43:308-15.
30. Conaway MR, Petroni GR. Bivariate sequential designs for phase II trials. *Biometrics* 1995;51:656-64.
31. Bryant J, Day R. Incorporating toxicity considerations into the design of two-stage phase II clinical trials. *Biometrics* 1995;51:1372-83.
32. Estey EH, Thall PF. New designs for phase 2 clinical trials. *Blood* 2003;102:442-8.
33. Berry DA. Bayesian clinical trials. *Nature Reviews Drug Discovery* 2006;5:27-36.
34. Thall PF, Wathen JK. Practical Bayesian adaptive randomisation in clinical trials. *Eur J Cancer* 2007;43:859-66.
35. Inoue LY, Thall PF, Berry DA. Seamlessly expanding a randomized phase II trial to phase III. *Biometrics* 2002;58:823-31.
36. Schaid DJ, Ingle JN, Wieand S, Ahmann DL. A design for phase II testing of anticancer agents within a phase III clinical trial. *Control Clin Trials* 1988;9:107-18.
37. Storer BE. A sequential phase II/III trial for binary outcomes. *Stat Med* 1990;9:229-35.
38. Cucherat M, Lièvre M, Leizorovicz A, Boissel JP. Lecture critique et interprétation des résultats des essais cliniques pour la pratique médicale. *Flammarion médecine-sciences* 2004.