



Avantages et inconvénients des phases de run-in dans les essais cliniques

Sont-elles toujours nécessaires et éthiquement justifiées?

Peter Kleist

PFC Pharma Focus AG, Volketswil

Quintessence

- Les phases dites de run-in peuvent renforcer l'efficacité des essais cliniques.
- L'exclusion de patients à l'issue d'une phase de run-in peut néanmoins compliquer l'interprétation des résultats de l'essai: l'efficacité et la sécurité du traitement testé peuvent en effet être surestimées, tout en limitant l'extrapolation des résultats à la pratique clinique quotidienne.
- Les phases de run-in sous placebo en simple aveugle sont controversées du point de vue de l'éthique.

Summary

Advantages and disadvantages of run-in phases for clinical trials. Are run-in phases always useful and ethically justifiable?

- *Run-in phases may enhance the efficiency of clinical trials.*
- *However, the exclusion of patients at the end of a run-in phase can complicate the interpretation of study results: the potential for overestimation of both the efficacy and safety of the investigational treatment limits the extent to which the trial results can be applied to everyday clinical practice.*
- *Single-blinded, placebo run-in phases are debatable from an ethical point of view.*

La phase de run-in: Quoi? Comment? Et pourquoi?

On entend en général par phase de run-in (ou délai de latence) une phase d'introduction de quelques semaines dans le cadre d'un essai clinique. Habituellement, elle est une partie préalable intégrante des études randomisées et contrôlées, c'est-à-dire une période qui précède la randomisation des patients dans les différents groupes de traitement. Les phases de run-in peuvent être de nature très variable. Il peut s'agir d'une phase ouverte, en simple ou même en double aveugle lorsque plusieurs traitements ou dosages sont testés. Au cours de cette période, les patients peuvent soit recevoir soit être privés d'un traitement actif. On utilise aussi souvent un placebo et l'administration de ce dernier se fera, le cas échéant, en simple aveugle, donc sans que le patient sache qu'il reçoit un placebo.

Les différentes formes des phases de run-in reflètent bien les problématiques particulières qu'elles adressent. Les phases de run-in poursuivent souvent simultanément plusieurs objectifs qui vont être brièvement décrits ci-dessous et qui seront abordés un à un par mesure de simplification.

Le but visé par les phases dites de wash-out est essentiellement d'obtenir des valeurs de référence qui ne soient pas influencées par le traitement antérieur; ceci permet de mieux standardiser les conditions initiales en prévision de la phase de test qui va suivre, de mieux contrôler la qualité du processus de randomisation et d'interpréter de manière plus fiable les résultats de chacun des traitements randomisés appliqués dans le cadre de l'étude [1].

L'administration d'un traitement actif au cours d'une phase de run-in, permet d'évaluer à l'avance les réponses, la tolérance et la fidélité au traitement (compliance ou observance) individuelles. L'observance peut également être testée à l'aide d'un placebo. Les phases de run-in sous placebo permettent de plus d'identifier les répondants au placebo.

Les examens réalisés durant une phase de run-in sur la réponse au traitement actif ou au placebo, sur la tolérance ou sur l'observance servent souvent à la sélection et à l'exclusion éventuelle de patients de l'étude proprement dite, autrement dit de la phase randomisée. Il s'agit en fait d'augmenter la puissance de discrimination (power) d'une étude, de diminuer le nombre de patients à inclure, de faciliter le déroulement de l'essai et d'augmenter la probabilité de succès de ce dernier [2].

Les phases de run-in offrent souvent plusieurs avantages et semblent de prime abord logiques; elles soulèvent cependant une série de questions, comme va le montrer le présent article. D'abord, l'exclusion prématurée de patients rend plus difficile l'interprétation des résultats et leur extrapolation dans la pratique quotidienne. Ensuite, les longues phases de wash-out sans aucun traitement et l'administration d'un placebo en simple aveugle sont contestables sur le plan éthique.

Remarques sur l'interprétation des résultats des études

Surestimation de la sécurité d'un traitement par des tests de tolérance préalables

Lorsqu'un médicament entraîne des effets indésirables, plus particulièrement en début de traitement, les patients qui ne supportent pas le médicament testé peuvent être identifiés à l'aide d'une phase de run-in. Cette démarche semble a priori sensée, mais les interruptions prématurées des traitements et l'exclusion active de certains patients à la suite de réactions d'intolérance prédéfinies, avant même la randomisation, sont susceptibles de compliquer l'interprétation des résultats de l'essai. Les patients restant dans le protocole ne sont en effet plus représentatifs du groupe global, puisqu'ils représentent une sélection de sujets au profil de tolérance favorable pour le médicament testé. Les effets indésirables rapportés durant la phase randomisée de l'étude tendront ainsi à sous-estimer leur incidence réelle dans le collectif non sélectionné. L'exemple suivant avec le bêtabloquant carvédilol, dont on sait qu'il peut aggraver initialement une insuffisance cardiaque préexistante, illustre bien ce phénomène.

1197 patients insuffisants cardiaques ont pris part à la phase ouverte d'un essai contrôlé par placebo [3]. Durant cette phase où tous les patients recevaient le carvédilol, 103 patients (8,6%) ont dû quitter l'étude, dont 67 (5,6%) en raison d'effets indésirables. Une péjoration de l'insuffisance cardiaque a été observée chez 17 patients (1,4%) et 7 patients (0,6%) sont décédés au cours de cette période. Pendant la phase randomisée qui a suivi avec les 1094 patients restants, le carvédilol a entraîné moins d'interruptions du traitement pour cause d'effets indésirables que le placebo (5,7 vs. 7,8%). Si on ajoute cependant à ce chiffre – dans l'idée d'un scénario catastrophe («worst case») sous la substance testée – les interruptions de traitement pour cause d'effets indésirables rapportés au cours de la phase de run-in, la différence serait exactement inverse [2].

Un autre problème inhérent aux phases de run-in avec un médicament actif tient au phénomène non contrôlé de «cumul» des effets indésirables. En l'absence de groupe de contrôle, la question de savoir si l'incidence des effets indésirables est due au médicament ou si elle est la conséquence de l'évolution naturelle de la maladie reste sans réponse [2].

Surestimation de l'utilité d'un traitement par une mesure préalable des taux de réponse

Réponse au médicament actif

On recourt aussi à des périodes de run-in pour sélectionner les patients qui répondent de manière mesurable au traitement ou du moins ceux dont l'état clinique ne se péjore pas. L'exclusion

des non-répondeurs revient toutefois encore une fois à une sélection positive de patients qui ne sont plus représentatifs d'une population non sélectionnée de patients souffrant d'une maladie donnée. Ceci peut évidemment donner lieu à une surestimation des effets du traitement par rapport aux conditions en pratique quotidienne. Voici un autre exemple:

Pour évaluer la réponse à un traitement et pour déterminer la dose optimale individuelle, les patients atteints de la maladie d'Alzheimer inclus dans une étude ont tout d'abord reçu en crossover, dans le cadre d'une phase de run-in en double aveugle de six semaines, deux doses différentes de tacrine, ainsi qu'un placebo [4]. Des 632 patients retenus initialement, 417 ont été exclus de la suite de l'essai car ils n'avaient pas répondu cliniquement au traitement ou parce qu'ils n'avaient pas supporté la tacrine. A peine un tiers du collectif initial a par conséquent été inclus dans la phase randomisée de l'essai destiné à tester l'efficacité de la tacrine versus placebo. Bien que le caractère préliminaire des résultats obtenus ait été clairement signalé dans la publication, les auteurs ne s'étendent pas sur la question discutable de l'extrapolation de ces derniers à la pratique quotidienne.

Dans ce type de phases de run-in, qui soumettent à une présélection le traitement actif ultérieurement contrôlé par placebo, il convient de ne pas non plus perdre de vue deux autres aspects: d'abord, des effets de type biais de rémanence (carry-over) du traitement actif peuvent survenir chez les patients du groupe placebo, ce qui conduit chez eux à une surestimation des valeurs initiales mesurées avant la randomisation. Ensuite, si des phénomènes d'interruption jouent encore un rôle, il faudra en outre compter avec une péjoration de l'état clinique sous placebo. Ces deux facteurs peuvent donc entraîner une distorsion systématique supplémentaire des résultats de la phase randomisée en faveur du traitement actif, à moins que la phase de run-in n'ait été complétée par une phase de wash-out [2].

Réponse au placebo

De nombreuses maladies se distinguent par un fort taux de répondeurs au placebo, susceptible de rendre plus difficile la démonstration de l'efficacité d'un traitement actif. Les phases de run-in sous placebo en simple aveugle sont donc fréquemment utilisées dans les essais sur la douleur, ainsi que sur les états dépressifs ou anxieux, dans le but d'identifier et ensuite d'exclure les répondeurs au placebo. Cette méthode est destinée à augmenter la probabilité qu'un traitement est efficace.

Il va sans dire qu'on ne traite pas les répondeurs au placebo de cette façon sur le terrain. Il n'est d'ailleurs pas évident d'admettre que les patients exclus de l'étude auraient peut-être aussi ré-

pondu au traitement actif [5]. Mais hormis l'exclusion des non répondeurs au traitement actif, celle des répondeurs au placebo peut potentiellement induire une surestimation des effets thérapeutiques par rapport à des conditions normales en pratique quotidienne, comme va le montrer l'exemple suivant:

Un essai randomisé évaluant l'efficacité antidépressive de la fluoxétine avec ou sans adjonction de pindolol fut précédé d'une phase de run-in sous placebo d'une durée de sept jours [6]. Des 132 patients du début, 19 ont été exclus de l'étude en raison d'une nette amélioration du score de la dépression sur l'échelle de Hamilton. A l'issue de la phase randomisée de l'essai, qui a porté sur 111 patients, on a mis en évidence un nombre significativement plus élevé de répondeurs sous traitement combiné (41/55 patients = 75%) que sous fluoxétine en monothérapie (33/56 patients = 59%; $p = 0,04$).

Supposons maintenant que les 19 répondeurs au placebo n'aient pas été exclus de l'étude, mais qu'ils aient été répartis en nombre égal dans les deux groupes de traitement (merci de fermer les yeux sur le fait qu'un patient devrait être «partagé en deux» au cours du processus). Même en admettant que tous les répondeurs au placebo auraient également répondu au traitement actif, il ne resterait plus qu'une légère tendance non significative en faveur du traitement combiné (50,5/64,5 [78%] vs. 42,5/65,5 patients [65%]; $p = 0,09$) [2].

Il est néanmoins fort probable que le scénario décrit ci-dessus constitue plutôt une exception: quatre méta-analyses, dont trois sur les antidépresseurs et une dans les troubles anxieux généralisés, n'ont pas réussi à établir la preuve que l'exclusion des répondeurs au placebo au cours d'une phase de run-in influence réellement l'effet corrigé pour le placebo des antidépresseurs ou des anxiolytiques [7-10]. Les auteurs remettent par conséquent en doute l'intérêt d'une phase de run-in sous placebo en simple aveugle dans l'optique d'une amélioration de l'efficacité d'une étude.

Si l'exclusion des répondeurs au placebo a véritablement une influence sur les résultats de l'étude, comme c'est le cas dans l'exemple qui précède, cette méthode va favoriser le traitement actif et rendre plus difficile l'extrapolation à la pratique sur le terrain. Si en revanche cette «manipulation» n'a pas d'effets, l'élaboration d'une procédure de sélection (screening), l'allongement de la durée de l'étude et le supplément de ressources nécessaires doivent être franchement remis en question. Quoiqu'il en soit, il convient de se demander dans les deux cas si les phases de run-in sous placebo en simple aveugle sont vraiment justifiées.

Surestimation de l'efficacité d'un traitement par des tests d'observance préalables

Certaines études utilisent une phase de run-in pour tester l'observance des sujets en vue de leur inclusion dans un essai clinique et dans le but

d'exclure les patients défaillants en termes de prise des médicaments (actifs ou placebo). Si un traitement est efficace, cela augmentera évidemment la probabilité de pouvoir le démontrer durant la phase randomisée.

L'inclusion exclusive de patients très observants constitue néanmoins une situation idéale qui ne reflète pas la réalité quotidienne. Il se peut aussi que ces patients se distinguent de ceux qui ont été exclus sur la base d'autres critères, par exemple un mode de vie plus sain. On peut s'attendre d'une manière générale à ce que les effets thérapeutiques observés durant un essai dans un collectif de patients sélectionnés soient plus importants qu'ils ne le seraient en moyenne dans un collectif de patients non sélectionnés de la pratique quotidienne. Ceci est bien illustré par ces deux essais de prévention primaire avec l'acide acétylsalicylique (AAS) dans la maladie coronarienne:

Dans la «Physicians' Health Study» [11], la phase contrôlée par placebo était précédée d'une phase de run-in de 18 semaines. Les patients insuffisamment observants – ils représentaient 34% des participants – ont été exclus avant la randomisation. Au cours de l'étude de cinq ans, l'observance des patients retenus était de 90% en moyenne. A la fin de l'étude, l'AAS a entraîné une réduction de 44% des infarctus du myocarde par rapport au placebo.

Une étude semblable, la «British Physicians' Study» [12], a par contre été réalisée sans phase de run-in. Les patients ont donc été randomisés en deux groupes et ont soit reçu l'acide acétylsalicylique, soit été informés de s'abstenir de prendre un quelconque produit contenant de l'AAS. L'observance pour la prise d'AAS était en moyenne de 70%. L'essai n'a finalement pas mis en évidence de différence significative entre les deux groupes.

La différence d'observance entre les deux études citées ci-dessus peut être considérée comme l'un des principaux facteurs permettant d'expliquer cette contradiction dans les résultats observés.

Considérations éthiques

Il n'y a pas de réponse simple à la question de savoir dans quels cas des phases sans traitement sont justifiées au début d'une étude. La réflexion devra intégrer plusieurs facteurs, notamment le degré de sévérité de la maladie de base, le risque immédiat qu'impliquerait un arrêt du traitement, la durée de la phase de run-in, ainsi que le tableau clinique. Les patients avec des valeurs tensionnelles fortement augmentées et simultanément porteurs de facteurs de risque de complications cardiovasculaires ne devraient par exemple être privés que de manière très transitoire de leur traitement antihypertenseur. Une autre question a trait à la durée pendant laquelle

on peut se contenter d'une diète seule chez des patients diabétiques de type 2 traités auparavant par des antidiabétiques, dans le but d'atteindre des valeurs initiales d'HbA1c élevées.

Les phases de run-in avec wash-out posent fondamentalement des problèmes comparables à ceux des essais contrôlés par placebo. Selon la «Note of Clarification» publiée en 2002 par l'Association mondiale des médecins à propos du § 29 de la «Déclaration d'Helsinki» [13], la justification de l'utilisation du placebo dans une étude dépend largement de la nécessité méthodologique. De même, on doit pouvoir répondre par l'affirmative à la question de savoir si une phase de run-in est indispensable à la puissance de démonstration de l'étude, au-delà de son aspect médical proprement dit.

Les considérations éthiques relatives aux phases de run-in concernent plus particulièrement l'information des patients et la déclaration de consentement. Cette dernière suppose en effet que les patients ont été clairement informés des objectifs et du déroulement précis de l'essai. Dans la «Physicians' Health Study» citée ci-dessus, les participants n'avaient pas connaissance de l'existence d'une phase de run-in. Ils pensaient en fait que l'étude avait déjà commencé. C'est sous un prétexte quelconque que le traitement a été modifié après un certain temps, afin de pouvoir démarrer la phase randomisée de l'étude [14].

Les phases de run-in sous placebo en simple aveugle doivent être abordées dans un esprit critique. Si les patients savaient qu'ils reçoivent un placebo, le recours à ce dernier serait évidemment absurde. On ne peut par conséquent pas fournir toute l'information et les patients ne donnent donc pas véritablement leur consentement à cette partie de l'essai. Le professeur Stephen Senn pose ainsi la question suivante: «Si les médecins investigateurs ne sont pas complètement honnêtes avec leurs patients, comment pourraient-ils attendre une entière franchise de leur part?» [15].

L'administration d'un placebo au cours de la phase de run-in donne l'impression d'un traitement actif, bien qu'aucun des patients n'ait une quelconque chance d'en bénéficier [1, 15]. C'est là en effet que réside la principale différence avec les phases randomisées des essais contrôlés par placebo: chaque patient a les mêmes chances, dans le cadre de la randomisation, de recevoir le médicament actif ou le placebo. Le médecin investigateur peut donc tenir vis-à-vis de ses patients un discours parfaitement ouvert; il peut même leur remettre le protocole de l'étude sans arrière-pensée puisqu'il n'a rien à leur cacher.

Un autre problème des phases de run-in sous placebo en simple aveugle concerne l'attitude à adopter face aux répondants au placebo. L'information contient-elle la mention d'une éventuelle exclusion de l'étude, alors que le traitement semble un succès et que le patient pourrait ne pas

être d'accord avec l'interruption? Et que peut-on dire à ces patients après qu'ils aient bien répondu au «traitement» [16]?

Conclusions et recommandations

Il est incontestable qu'une phase de run-in améliore l'efficacité d'un essai clinique. La garantie d'une observance suffisante des participants à une étude peut par exemple constituer une condition sine qua non pour la reconnaissance de la vraie valeur d'une intervention et pour l'élimination d'un biais significatif dans l'évaluation des effets [17]. Une publication séparée sur la «Physicians' Health Study» a très bien montré l'influence de la phase de run-in sur la puissance du test: l'observance a été augmentée de près de 40%, la dimension de l'échantillon a chuté de 34% et la durée du suivi a été raccourcie de 7% [18].

Une phase de run-in précédant la randomisation et aboutissant à l'exclusion de certains patients définis ne limite pas automatiquement sa validité interne, c'est-à-dire sa validité scientifique. Les difficultés relèvent plutôt d'une limitation de la validité externe, autrement dit de l'impossibilité de pouvoir généraliser les résultats en raison d'un collectif sélectionné [19]. Les résultats ne peuvent ainsi plus être extrapolés que de façon conditionnelle à l'usage thérapeutique courant. Par rapport aux conditions de la pratique quotidienne, les études surestiment avec leurs phases de run-in et l'exclusion subséquente de certains patients aussi bien l'efficacité que la sécurité et la tolérance des traitements testés. Cette surestimation peut ensuite se reporter sur les méta-analyses ultérieures et induire une distorsion des recommandations thérapeutiques qui en découlent, si l'on n'a pas tenu compte de l'éventuel biais de sélection.

La valeur scientifique des phases de run-in sous placebo en simple aveugle pour «éliminer» les répondants au placebo est controversée, comme nous l'avons déjà dit. Du point de vue éthique, la réalisation préalable de phases de run-in sous placebo en simple aveugle mérite une certaine prudence, car l'information dispensée aux patients ne peut répondre en elle-même aux standards en la matière. Il est clair qu'on veillera aussi dans les autres formes de phases de run-in à donner une information adéquate aux patients.

Pour faciliter l'interprétation des essais comportant une phase de run-in, les publications devraient fournir des indications totalement transparentes sur l'existence d'une phase de run-in, sur les conséquences que celle-ci a eues sur la partie randomisée de l'étude et sur la manière d'interpréter les résultats, compte tenu de la méthode utilisée. Si un traitement actif a été administré au cours de la phase de run-in, le type et la fréquence des effets indésirables observés devront être clairement mentionnés.

Références

- 1 Evans M. Justified deception? The single blind placebo in drug research. *J Med Ethics*. 2000;26:188–93.
- 2 Pablos-Méndez A, Barr RG, Shea S. Run-in periods in randomized trials. Implications for the application of results in clinical practice. *JAMA*. 1998;279:222–5.
- 3 Packer M, Bristow MR, Cohn JN, et al. The effect of carvedilol on morbidity and mortality in patients with chronic heart failure. *N Engl J Med*. 1996;334:1349–55.
- 4 Davis KL, Thal LJ, Gamzu ER, et al. A double-blind, placebo-controlled multicenter study of tacrine for Alzheimer's disease. *N Engl J Med*. 1992;327:1253–9.
- 5 Van Marwijk HWJ, Ader HJ. Placebo washouts inflate antidepressant effects in general practice. *BMJ*. 2005;330:420.
- 6 Pérez V, Gilaberte I, Faries D, et al. Randomised, double-blind, placebo-controlled trial of pindolol in combination with fluoxetine antidepressant treatment. *Lancet*. 1997;349:1594–7.
- 7 Greenberg RP, Fisher S, Riter JA. Placebo wash out is not a meaningful part of antidepressant drug trials. *Percept Mot Skills*. 1995;81:688–90.
- 8 Lee S, Walker JR, Jakul L, et al. Does elimination of placebo responders in a placebo run-in increase the treatment effect in randomized clinical trials? A meta-analytic evaluation. *Depress Anxiety*. 2004;19:10–9.
- 9 Trivedi MH, Rush J. Does a placebo run-in or a placebo treatment cell affect the efficacy of antidepressant medications? *Neuropsychopharmacology*. 1994;11:33–4.
- 10 Mitte K, Noack P, Steil R, Hautzinger M. A meta-analytic review of the efficacy of drug treatment in generalized anxiety disorder. *J Clin Psychopharmacol*. 2005;25:141–50.
- 11 Steering Committee of the Physicians' Health Study Research Group. Final report on the aspirin component of the ongoing Physicians' Health Study. *N Engl J Med*. 1989;321:129–35.
- 12 Peto R, Gray R, Collins R, et al. Randomised trial of prophylactic daily aspirin in British male doctors. *BMJ*. 1988;296:313–6.
- 13 World Medical Association. Note of clarification on paragraph 29 of the WMA Declaration of Helsinki. Available from: www.wma.net/e/policy/b3.htm.
- 14 Lang JM. The use of a run-in to enhance compliance. *Stat Med*. 1990;9:87–93.
- 15 Senn S. Are placebo run ins justified? *BMJ*. 1997;1191–3.
- 16 Kaptchuk TJ. Concerns about run-in periods in randomized trials. Letter to the editor. *JAMA*. 1998;279:1526–7.
- 17 Glynn RJ, Buring JE, Hennekens CH. Concerns about run-in periods in randomized trials. Letter to the editor. *JAMA*. 1998;279:1526.
- 18 Lang JM, Buring JE, Rosner B, et al. Estimating the effect of the run-in on the power of the Physicians' Health Study. *Stat Med*. 1991;10:1585–93.
- 19 Rothwell PM. External validity of randomized controlled trials: "To whom do the results of this trial apply?" *Lancet*. 2005;365:82–93.

Correspondance:
Dr Peter Kleist
PFC Pharma Focus AG
Chriesbaumstrasse 2
CH-8604 Volketswil
peter.kleist@pfc.ch