

Critique des critères d'évaluation intermédiaires en oncologie

L'avantage clinique de nombreux médicaments antitumoraux n'est pas clairement établi. L'une des raisons à la base de cette incertitude est l'utilisation parfois inadéquate de critères d'évaluation intermédiaires dans les études. Les commentateurs plaident pour une utilisation plus rationnelle des critères d'évaluation intermédiaires et pour un relèvement du seuil pour l'autorisation de ces médicaments. Cet article traite de critères d'évaluation intermédiaires fréquemment utilisés en oncologie, en s'attardant sur un certain nombre de problèmes tels que les faibles corrélations avec les critères d'évaluation qui importent réellement : la survie et la qualité de vie.

Introduction

Les médicaments antitumoraux constituent depuis quelques années la plus grande catégorie de nouveaux médicaments approuvés par l'Agence européenne des médicaments (EMA). Un certain nombre de publications soulignent que l'avantage clinique d'une part importante de ces médicaments n'est pas clairement établi. Dans le présent article, nous abordons une importante limitation méthodologique dont souffrent les études effectuées, notamment l'utilisation parfois inadéquate de critères d'évaluation intermédiaires.

Les critères d'évaluation intermédiaires sont utilisés dans certaines études en remplacement de critères d'évaluation cliniquement pertinents (ou forts). Les critères d'évaluation intermédiaires permettent de limiter l'étendue, la durée, et donc le coût d'une étude. Lorsqu'il existe une solide corrélation entre les deux types de critères, les études basées sur des critères d'évaluation intermédiaires peuvent fournir des résultats fiables. Mais cette solide association fait souvent défaut et les études qui ont un critère d'évaluation intermédiaire comme critère d'évaluation primaire doivent être interprétées avec prudence.

En oncologie, la **survie globale** et la **qualité de vie** sont les critères d'évaluation les plus importants, c.-à-d. les critères d'évaluation forts. Pour les tumeurs solides, les investigateurs utilisent le plus souvent des critères d'évaluation intermédiaires sur la base de mesures des tumeurs par imagerie médicale. Cette méthode permet de contrôler, dès le début de l'étude, si la tumeur grandit ou régresse pendant le traitement. Les nouveaux médicaments antitumoraux sont souvent enregistrés sur la base d'études dont les seules données reposent sur des critères d'évaluation intermédiaires, donc sans informations relatives à la survie ou à la qualité de vie. Il semble logique que la régression d'une tumeur résultera en une survie plus longue et/ou de meilleure qualité, mais la réalité n'est pas aussi linéaire, ce qui est à l'origine du débat. Peut-on considérer une réduction du volume tumoral comme un bon critère de substitution pour la survie globale et/ou la qualité de vie ?

Critères d'évaluation intermédiaires utilisés en oncologie

Cet article aborde 3 critères d'évaluation intermédiaires qui sont souvent utilisés en oncologie à savoir « *objective response rate* » (ORR), « *progression-free survival* » (PFS), et « *disease-free survival* » (DFS). Au moment de concevoir une étude, le choix d'un critère d'évaluation dépend de facteurs tels que le pronostic de la maladie et la phase du traitement (p. ex. avant ou après la chirurgie, en traitement de fond, en contexte palliatif).

Tous les critères d'évaluation (intermédiaires) possibles ont des avantages et des inconvénients¹. Même la survie globale, qui fait office de référence absolue, a des inconvénients. Ce critère d'évaluation est influencé par la stratégie de prise en charge (p. ex. traitements d'appoint) après l'arrêt du traitement à l'étude, si bien qu'il reste difficile de déterminer l'effet du médicament étudié en termes de survie. Les études qui autorisent les crossover sont confrontées à un problème similaire. Dans le cadre d'un crossover, les patients du bras témoin peuvent, en cas d'échec du traitement (souvent traitement standard + placebo), passer dans le bras expérimental (cf. plus loin).

Pour comprendre les critères d'évaluation intermédiaires dans les études sur des tumeurs solides, il faut d'abord définir la notion de **réponse tumorale**. Dans ces études, les critères d'évaluation intermédiaires sont souvent évalués à l'aide de mesures par CT-scanner (ou tomодensitométrie TDM), les lésions tumorales étant mesurées et suivies conformément à des critères standard (voir « plus d'infos »). De manière globale et simplifiée, la réponse de la tumeur au traitement peut être catégorisée comme suit :

- **Réponse partielle (*partial response, PR*)** : une réduction de > 30 % des dimensions des lésions tumorales
- **Maladie progressive (*progressive disease, PD*)** : une croissance de > 20 % des dimensions des lésions tumorales
- **Maladie stable (*stable disease, SD*)** : ni PD, ni PR
- **Réponse complète (*complete response, CR*)** : plus aucune preuve de maladie objectivable

Les mesures effectuées sur les clichés CT sont une source de critique à l'encontre des critères d'évaluation intermédiaires. Il existe une importante variabilité intra- et inter-observateurs dans l'évaluation de la réponse. Un deuxième CT-scanner doit dès lors toujours venir confirmer une réponse pour limiter ce problème. De même, tous les clichés doivent de préférence être analysés par un évaluateur indépendant, qui ne connaît pas le traitement. Par ailleurs, les valeurs limites qui sont utilisées pour définir une réponse partielle (30 % de régression tumorale) ou une maladie progressive (20 % de croissance tumorale) sont arbitraires. En ce sens qu'une tumeur qui grandit de 19 % ou de 21 % ne donnera vraisemblablement pas de différence en termes de symptômes, et pourtant la valeur limite utilisée pour définir la progression de la maladie dans les études est fixée à 20 %. Dans le même ordre d'idées, la réponse tumorale d'un patient dont la tumeur régresse de 29 % ou grandit de 19 % sera classée dans la catégorie des maladies stables (SD), alors que le résultat clinique attendu pourra être très différent entre les deux situations.

Taux de réponse

Le taux de réponse objective ou *objective response rate* (ORR) est le pourcentage total de patients qui obtiennent une réponse partielle (PR) ou une réponse complète (CR) dans une étude.

Le **taux de réponse objective (*objective response rate, ORR*)** est utilisé dans de nombreuses études comme critère d'évaluation en vue d'évaluer l'efficacité de médicaments antitumoraux. Ce critère d'évaluation est aussi souvent utilisé dans les études de phase II, parfois sans bras témoin lorsqu'il n'existe pas d'autres options thérapeutiques, afin d'accélérer l'octroi de l'autorisation de mise sur le marché (voir Folia de septembre 2018). Il ne faut pas perdre de vue 2 choses. D'une part, la régression d'une tumeur ne correspond pas nécessairement en soi à un avantage clinique pour le patient. En d'autres termes, ce critère d'évaluation illustre uniquement ce que le traitement fait à la tumeur, et non l'avantage qu'il représente pour le patient. Par exemple, quel bénéfice un patient dyspnéique tire-t-il d'une diminution radiologique du volume des métastases pulmonaires s'il reste dépendant d'une oxygénothérapie ? En outre, une tumeur peut par exemple effectivement régresser mais elle pourrait, à terme, développer une résistance au traitement étudié et adopter une forme plus agressive et difficile à traiter, ce qui reviendrait à un effet net nul, voire négatif, sur la survie. D'autre part, certains traitements ont un effet de stabilisation (SD) plutôt que de réduction tumorale, ce qui n'est pas mesuré par le paramètre ORR. D'autres critères d'évaluation en tiennent compte, mais ils ne sont pas abordés dans le présent article. Souvent, les investigateurs évaluent aussi la **durée de la réponse (*duration of response, DOR*)**, à savoir l'intervalle de temps entre une réponse (PR ou CR) et une progression de la maladie (PD). Ce paramètre est probablement plus important que le pourcentage de réduction du volume tumoral.

Survie sans progression

La survie sans progression (*progression-free survival, PFS*) peut être définie comme l'intervalle de temps entre la randomisation et la progression de la maladie (PD) confirmée à l'imagerie ou le décès toutes causes confondues.

La **survie sans progression (*progression-free survival, PFS*)** est probablement le critère d'évaluation le plus utilisé en oncologie au cours des dernières années chez les patients atteints d'un cancer à un **stade**

avancé. Les nouveaux médicaments antitumoraux sont souvent étudiés en premier lieu à ce stade de la maladie. Dans les études randomisées contrôlées (*randomized controlled trials*, RCT), la PFS remplace de plus en plus souvent la survie globale comme critère d'évaluation primaire. Ce critère d'évaluation composite peut être défini comme l'intervalle de temps entre la randomisation et la progression de la maladie (PD) confirmée à l'imagerie ou le décès toutes causes confondues. Contrairement au critère d'évaluation 'survie globale', le critère d'évaluation 'PFS' n'est pas influencé par les différents schémas de traitement après l'échec du médicament étudié ou par le *crossover* à l'atteinte du critère d'évaluation dès le premier signe probant de progression de la maladie (PD). La PFS (plus précisément la progression de la maladie) étant atteinte plus tôt que le décès, les études qui utilisent comme critère d'évaluation primaire la PFS plutôt que la survie globale sont de plus courte durée. Les médicaments innovants pourraient de ce fait être plus rapidement à la disposition des patients. Ces avantages expliquent le recours fréquent à la PFS comme critère d'évaluation primaire.

Néanmoins, l'utilisation de la PFS fait également l'objet de critiques. L'utilisation croissante du critère d'évaluation intermédiaire 'PFS' ne s'explique pas par de solides corrélations avec la survie globale ou la qualité de vie (voir plus loin). Dans les études, un gain de PFS ne se traduit pas toujours par un gain de survie. Nous avons déjà évoqué les possibles biais dus aux problèmes relatifs aux mesures des tumeurs. En l'absence de gain de survie en dépit d'un gain de PFS, l'explication parfois avancée, à tort ou à raison, tient aux différences de traitement après la progression de la maladie entre le groupe interventionnel et le groupe témoin. Voir « plus d'infos » pour de plus amples précisions sur les permutations dans ce contexte.

Une autre critique formulée à l'encontre de l'utilisation de la PFS en lieu et place du critère d'évaluation 'survie globale' dans le processus d'autorisation de mise sur le marché d'un nouveau médicament. La PFS étant utilisée dans des études auprès de patients atteints de cancer avancé, ce gain de temps n'est en effet pas toujours énorme. Une analyse américaine (relative, donc, à des autorisations octroyées par la FDA) a estimé, sur la base de données datant de 2006 à 2018, qu'il serait question d'un gain de temps moyen de 11 mois (IC 95 % : 5-17 mois) sur la durée d'une étude⁴. Il faut savoir qu'une autre étude américaine a estimé l'intervalle de temps entre les premières études cliniques et l'autorisation d'un médicament à 7,3 ans en moyenne (intervalle : de 5,8 à 15,2 ans)⁵. Le gain de temps limité obtenu grâce à l'utilisation de la PFS plutôt que de la survie globale doit donc être mis en balance avec la plus grande incertitude quant à l'efficacité du médicament.

Notons qu'une analyse récente montre que les nouveaux médicaments antitumoraux sont approuvés beaucoup plus rapidement par la FDA que par l'EMA en Europe, ce dont il faut tenir compte lors de l'interprétation des chiffres américains mentionnés ci-dessus⁶. Au niveau européen, il semblerait donc que l'utilisation de la PFS au lieu de la survie globale présente encore moins un avantage en termes de gain de temps.

Survie sans maladie

La survie sans maladie (*disease-free survival*, DFS) est atteinte en cas de récurrence du cancer ou de décès toutes causes confondues.

La **survie sans maladie (*disease-free survival*, DFS)** est comparable au critère d'évaluation 'survie sans progression', mais est utilisée dans des études menées chez des patients qui sont encore traités dans une optique curative. Plus précisément dans des études qui évaluent un traitement post-chirurgie ou post-radiothérapie ou, autrement dit, en **contexte adjuvant**. Ce critère d'évaluation composite est atteint en cas de récurrence du cancer ou de décès toutes causes confondues. En général, la survie sans maladie possède une meilleure valeur prédictive de la survie globale que la survie sans progression. Une rechute du cancer peut en outre être déterminée de manière plus objective qu'une réponse (PD ou PR) sur une tumeur existante. Une plus-value clairement établie de médicaments en contexte adjuvant revêt une importance supplémentaire étant donné que, dans ce cas, ce sont des patients sans signe probant de cancer qui sont exposés à des médicaments associés à des effets indésirables potentiellement graves.

Limitations en termes de validité et de pouvoir prédictif

Pour évaluer le pouvoir prédictif d'un critère d'évaluation intermédiaire en matière de survie globale, on détermine les corrélations entre les deux⁷. Ces études de validation doivent systématiquement être répétées pour chaque indication et pour chaque intervention. Mais il arrive souvent que ces études ne soient pas réalisées et, lorsqu'elles le sont, les corrélations sont souvent faibles ou modérées (voir « plus d'infos »).

Même si l'association entre la PFS et la survie globale est faible, en présence d'un gain de PFS, on pourrait s'attendre à une qualité de vie améliorée, étant donné que la progression de la maladie (à l'imagerie) est retardée. Mais là encore, les études révèlent souvent de faibles corrélations entre les critères intermédiaires et le critère 'qualité de vie'^{9,10}. Si l'on veut évaluer la qualité de vie dans une étude (ce qui est encore trop rare), il est préférable d'évaluer directement la qualité de vie pendant l'étude.

Bien que les critères intermédiaires ne soient souvent pas en mesure de prédire des critères cliniques en oncologie, ils servent de base à l'autorisation de mise sur le marché (conditionnelle) de nouveaux médicaments⁷. Ce point serait moins problématique si les études post-autorisation de mise sur le marché (AMM) démontraient de nets avantages au niveau de critères d'évaluation cliniquement pertinents, mais ce n'est généralement pas le cas. Dans un certain nombre de cas, des médicaments n'ont pas été retirés du marché alors même qu'aucun gain de survie n'a pu être établi^{7,10}. Les commentateurs plaident pour une utilisation plus rationnelle des critères intermédiaires et pour un relèvement du seuil pour l'autorisation de ces médicaments¹¹⁻¹⁴.

Autorisations de mise sur le marché (AMM) conditionnelles, critères d'évaluation intermédiaires et transparence

Dans les Folia de septembre 2018, nous évoquons les AMM conditionnelles (*conditional market authorisations*) qui permettent, depuis 2006, d'accélérer la disponibilité de médicaments innovants répondant à un besoin médical non rencontré (*unmet medical need*) sur la base de données réduites. Ces AMM conditionnelles sont souvent, mais pas toujours, basées sur des études qui utilisent des critères d'évaluation intermédiaires. Le cas échéant, la société pharmaceutique doit fournir - après l'octroi de l'AMM conditionnelle - des données complémentaires qui doivent confirmer l'avantage clinique. Nous retrouvons dans la littérature une critique récurrente concernant ces études de confirmation, qui ne sont pas réalisées, qui sont conduites beaucoup plus tard ou qui utilisent aussi des critères d'évaluation intermédiaires comme critère d'évaluation primaire. Par ailleurs, le critère « besoin médical non rencontré » lié à cette procédure accélérée est parfois remis en question, car il ne s'agit souvent pas des premiers médicaments de leur classe dont on peut attendre qu'ils apportent d'importantes différences par rapport à d'autres médicaments de la même classe qui, eux, sont déjà approuvés ou qu'il existe d'autres médicaments dotés d'un autre mécanisme d'action.

Quand l'EMA approuve un médicament, elle publie une information détaillée au sujet dudit médicament dans un rapport EPAR (*European Public Assessment Report*). Une étude a montré que ces rapports d'évaluation de médicaments autorisés entre 2011 et 2018 n'indiquaient pas systématiquement si les critères d'évaluation intermédiaires utilisés étaient validés ou, en d'autres termes, s'ils possédaient une valeur prédictive à l'égard de critères d'évaluation cliniquement pertinents¹⁵. Ce point n'était d'ailleurs pas non plus mentionné dans les notices ou dans les Résumés des Caractéristiques du Produit (RCP). Les rapports EPAR contiennent aussi les mesures ou les obligations auxquelles les sociétés pharmaceutiques doivent se plier après l'octroi d'une AMM conditionnelle. Là aussi, les rapports étudiés n'indiquaient pas toujours clairement si ces mesures devaient confirmer un avantage clinique plus tard. Pas plus qu'ils ne mentionnaient systématiquement le critère d'évaluation à utiliser dans les études de confirmation. Les auteurs de cette étude ont conclu à la nécessité d'une meilleure documentation, plus uniforme, de toutes ces informations dans les rapports EPAR, dans les notices et dans les RCP afin que les médecins et les patients aient une meilleure vue d'ensemble¹⁵. L'objectif visé est double : d'une part, ne pas surestimer l'efficacité de médicaments antitumoraux, et d'autre part, ne pas sous-estimer leurs effets indésirables.

L'EMA doit établir des critères plus clairs et contraignants à l'attention de l'industrie pharmaceutique avant l'octroi d'une AMM conditionnelle, de sorte que l'autorisation puisse être révoquée, si nécessaire, dans l'éventualité où les critères ne seraient pas rencontrés¹⁴⁻¹⁶.

Conclusion et commentaire

- Les autorités chargées de la réglementation des médicaments autorisent de plus en plus souvent des médicaments antitumoraux sur la base d'études qui utilisent des critères d'évaluation intermédiaires souvent peu à même de prédire des critères d'évaluation cliniquement pertinents. Les commentateurs plaident pour une utilisation plus rationnelle des critères d'évaluation intermédiaires dans les études et pour un relèvement du seuil pour l'autorisation de ces médicaments.
- Le débat sur l'efficacité incertaine de nombreux médicaments antitumoraux va bien au-delà de l'utilisation parfois inadéquate de critères d'évaluation intermédiaires dans les études. D'autres limitations méthodologiques^{14,17}, le gain de survie marginal dans certaines études ainsi que les prix exorbitants et le manque de transparence qui les entoure font, eux aussi, débat. Ce sujet est approfondi dans un rapport (cliquez [ici](#)) publié l'an dernier par le Centre fédéral d'expertise des soins de santé (KCE)¹⁶. Le rapport se termine (à partir de la page 40) sur 19 recommandations qui devraient, à l'avenir, permettre aux médecins et aux patients de mieux estimer la valeur ajoutée de ces médicaments. Le CBIP soutient ces recommandations et y ajoute qu'il ne faut pas confondre limitation de l'accès au marché pour des médicaments (innovants) sans plus-value établie et limitation de l'innovation.

Sources

- 1 Wilson MK, Karakasis K, Oza AM. Outcomes and endpoints in trials of cancer treatment: the past, present, and future. *Lancet Oncol.* 2015 Jan;16(1):e32-42.
- 2 https://imaging.cancer.gov/clinical_trials/imaging_response_criteria.htm
- 3 Haslam A, Prasad V. When is crossover desirable in cancer drug trials and when is it problematic?. *Ann Oncol.* 2018;29(5):1079-1081.
- 4 Chen EY, Joshi SK, Tran A, et al. Estimation of Study Time Reduction Using Surrogate End Points Rather Than Overall Survival in Oncology Clinical Trials. *JAMA Intern Med.* 2019 May 1;179(5):642-647.
- 5 Prasad V, Mailankody S. Research and Development Spending to Bring a Single Cancer Drug to Market and Revenues After Approval. *JAMA Intern Med.* 2017 Nov 1;177(11):1569-1575.
- 6 Lythgoe MP, Desai A, Gyawali B, et al. Cancer Therapy Approval Timings, Review Speed, and Publication of Pivotal Registration Trials in the US and Europe, 2010-2019. *JAMA Netw Open.* 2022;5(6):e2216183.
- 7 Kemp R, Prasad V. Surrogate endpoints in oncology: when are they acceptable for regulatory and clinical decisions, and are they currently overused? *BMC Med.* 2017 Jul 21;15(1):134.
- 8 Haslam A, Hey SP, Gill J, Prasad V. A systematic review of trial-level meta-analyses measuring the strength of association between surrogate end-points and overall survival in oncology. *Eur J Cancer.* 2019 Jan;106:196-211.
- 9 Kovic B, Jin X, Kennedy SA, Hylands M, et al. Evaluating Progression-Free Survival as a Surrogate Outcome for Health-Related Quality of Life in Oncology: A Systematic Review and Quantitative Analysis. *JAMA Intern Med.* 2018 Dec 1;178(12):1586-1596.
- 10 Gyawali B, Rome B N, Kesselheim A S. Regulatory and clinical consequences of negative confirmatory trials of accelerated approval cancer drugs: retrospective observational study *BMJ* 2021; 374 :n1959
- 11 Booth CM, Eisenhauer EA. Progression-free survival: meaningful or simply measurable? *J Clin Oncol.* 2012 Apr 1;30(10):1030-3.
- 12 Mintzes B, Vitry A. Flawed evidence underpins approval of new cancer drugs *BMJ* 2019; 366 :l5399
- 13 Dawoud D, Naci H, Ciani O, Bujkiewicz S. Raising the bar for using surrogate endpoints in drug regulation and health technology assessment *BMJ* 2021; Sep 15;374:n2191.
- 14 Schnog, J.B., Samson, M.J., Gans, R.O.B. et al. An urgent call to raise the bar in oncology. *Br J Cancer* 125, 1477–1485 (2021).
- 15 Schuster Bruce C, Brhlikova P, Heath J, et al. The use of validated and nonvalidated surrogate endpoints in two European Medicines Agency expedited approval pathways: A cross-sectional study of products authorised 2011-2018. *PLoS Med.* 2019 Sep 10;16(9):e1002873.
- 16 Neyt M, Devos C, Thiry N, et al. Benefits and costs of innovative oncology drugs in Belgium (2004-2017). *Health Technology*

Assessment (HTA) Brussels: Belgian Health Care Knowledge Centre (KCE). 2021. KCE Reports 343. D/2021/10.273/23

17 Naci H, Davis C, Savović J, et al. Design characteristics, risk of bias, and reporting of randomised controlled trials supporting approvals of cancer drugs by European Medicines Agency, 2014-16: cross sectional analysis. *BMJ*. 2019 Sep 18;366:l5221.

Colophon

Les *Folia Pharmacotherapeutica* sont publiés sous l'égide et la responsabilité du *Centre Belge d'Information Pharmacothérapeutique* (Belgisch Centrum voor Farmacotherapeutische Informatie) a.s.b.l. agréée par l'Agence Fédérale des Médicaments et des Produits de Santé (AFMPS).

Les informations publiées dans les *Folia Pharmacotherapeutica* ne peuvent pas être reprises ou diffusées sans mention de la source, et elles ne peuvent en aucun cas servir à des fins commerciales ou publicitaires.

Rédacteurs en chef: (redaction@cbip.be)

T. Christiaens (Universiteit Gent) et
J.M. Maloteaux (Université Catholique de Louvain).

Éditeur responsable:

T. Christiaens - Nekkersberglaan 31 - 9000 Gent.