



Société Française de
Pharmacologie et de Thérapeutique

Groupe de Travail Méthodologie

Livre blanc SFPT

De la nécessité de la méthodologie
dans l'évaluation des médicaments

Document compagnon

Dossier 14 – Les essais adaptatifs

14 février 2022

Groupe de rédaction / relecture (Par ordre alphabétique)

- Theodora Angoulvant
- Laurent Bertolotti
- Jean-Luc Cracowski
- Michel Cucherat
- Dominique Deplanque
- Guillaume Grenet
- François Gueyffier
- Behrouz Kassai
- Charles Khouri
- Silvy Laporte
- Bruno Laviolle
- Jean-Christophe Lega
- Clara Locher
- Florian Naudet
- Edouard Ollier
- Antoine Pariente
- Matthieu Roustit
- Tabassome Simon



[Licence Creative Commons](#)

Cette œuvre est mise à disposition selon les termes de la Licence Creative Commons Attribution 4.0 International

Vous êtes autorisé à :

- Partager — copier, distribuer et communiquer le matériel par tous moyens et sous tous formats
- Adapter — remixer, transformer et créer à partir du matériel pour toute utilisation, y compris commerciale.

Table des matières

1	Introduction.....	7
2	Problématiques méthodologiques	9
3	Etudes de cas.....	11
4	Méta-recherche.....	12
5	Conclusion	13

1 Introduction

Le terme de design adaptatif (ou essai adaptatif) recouvre un ensemble hétérogène de schémas expérimentaux ayant en commun, selon la définition de la FDA, la possibilité d'apporter des modifications planifiées de manière prospective à un ou plusieurs aspects de l'étude, sur la base de données accumulées au cours de l'essai [1]. Le caractère adaptatif d'un essai peut ainsi concerner des objectifs variés [1, 2, 3], parmi lesquels :

- Adapter le nombre total de patients à inclure dans l'essai : une vérification des hypothèses sur lesquelles se basait le calcul d'effectif initial après avoir inclus une partie des patients permet de réévaluer (en aveugle ou non) l'effectif en cours d'étude ou évaluer sa puissance conditionnelle, qui est la probabilité que le résultat final de l'étude soit statistiquement significatif, compte-tenu des données déjà observées [4].
- Prendre une décision sur la poursuite ou non de l'essai : les essais séquentiels en groupes (*group sequential designs*) prévoient différentes analyses intermédiaires ainsi que des critères d'arrêt définis *a priori*. Ces règles d'arrêt offrent la possibilité de conclure précocement à l'efficacité ou à la futilité (c'est-à-dire une faible probabilité que le traitement soit efficace à la fin de l'essai) du traitement.
- Adapter la population de l'essai : les stratégies dites « d'enrichissement » (*adaptive enrichment*) permettent de définir des sous-groupes de patients répondant le plus au traitement, sur la base d'une ou plusieurs analyses intermédiaires, et de focaliser sur la poursuite de l'essai sur ces sous-groupes.

Le terme « d'essai adaptatif » recouvre également d'autres approches, que nous ne détaillerons pas dans ce chapitre :

- Adapter le traitement ou la dose : la modification de la dose est de longue date réalisée dans les essais de phases précoces, selon une démarche algorithmique (par exemple 3+3). Une approche adaptative consiste à modéliser la probabilité de survenue d'un événement. C'est par exemple le cas des méthodes CRM (*Continual Reassessment Methods*) qui utilisent les données accumulées dans l'essai pour déterminer la dose administrée au patient suivant, selon une approche bayésienne.
- Introduire ou arrêter de nouveaux bras de traitement : dans certains cas, plusieurs doses sont testées en parallèles, et certaines peuvent être abandonnées en cours d'essais à la suite d'analyses intermédiaires, selon un design séquentiel en groupes. Les essais plateformes représentent un autre exemple d'essais adaptatifs : plusieurs traitements sont comparés à un même contrôle, qui représente le standard de prise en charge ; certains bras peuvent être introduits ou arrêtés, selon des règles prédéfinies.
- Adapter l'allocation des patients : c'est le cas par exemple des méthodes de minimisation, qui visent à équilibrer les groupes sur certaines caractéristiques lors de la randomisation, en se basant sur les caractéristiques des patients déjà inclus. D'autres approches, plus controversées, consistent à adapter la randomisation en fonction des résultats intermédiaires accumulés au cours de l'étude [5].

- Adapter la dose ou le choix du traitement qui sera évalué dans une phase ultérieure du développement, dans le cadre d'essais combinés (également appelés « sans couture », *seamless*). Ces essais peuvent regrouper des études de phase 1 et de phase 2, ou de phase 2 et de phase.

Les designs séquentiels en groupes et la réévaluation du nombre de sujets nécessaires en cours d'essai sont actuellement les formes les plus courantes d'essais adaptatifs. Ils présentent de nombreux avantages [1], que nous détaillerons dans ce chapitre :

- Sur le plan statistique, ils maximisent les chances de démontrer l'efficacité du traitement (donc la puissance), ou à puissance équivalente minimisent le nombre de patients inclus.
- Sur le plan éthique, ils permettent d'arrêter précocement un essai qui aura trop peu de chances de démontrer une supériorité du traitement (futilité), évitant ainsi d'exposer inutilement des patients à un risque d'effet indésirable. A l'inverse, un arrêt précoce pour supériorité permet de proposer à tous les patients le traitement le plus efficace.
- Leur flexibilité est un atout pour les promoteurs et les financeurs des essais, les arrêts précoces étant synonymes d'essais moins longs et moins coûteux.

2 Problématiques méthodologiques

La flexibilité des essais adaptatifs ne doit pas compromettre la validité et l'intégrité des résultats. En effet, les adaptations réalisées en cours d'étude ne peuvent pas être déterminées arbitrairement, mais en fonction de règles et de méthodes préétablies. Par définition, **un essai adaptatif prévoit la possibilité et les modalités d'adaptation *a priori*** ; il ne s'agit pas d'une modification substantielle du protocole apportée en cours d'étude. Il est donc essentiel de fournir suffisamment de détails sur le type d'adaptation et les méthodes utilisées (types d'analyses intermédiaires, définition des règles d'arrêt, méthodes statistiques utilisées) avant le début de l'étude.

Les essais adaptatifs s'appuyant sur l'utilisation de données accumulées au cours de l'étude elle-même, ils impliquent fréquemment des analyses intermédiaires. L'une des principales problématiques méthodologiques qui en découle est la **gestion de la multiplicité des comparaisons**, avec l'augmentation du risque de conclure à tort à l'intérêt du traitement (erreur de type I). Dans les essais séquentiels en groupes, qui comportent une ou plusieurs analyses intermédiaires, si le seuil de risque de 5% classiquement retenu est utilisé pour chaque analyse, le risque global dépassera ce seuil. Le seuil de significativité retenu pour chaque analyse doit ainsi être ajusté pour préserver un risque global acceptable. Différentes approches ont été proposées, plus ou moins conservatrices sur les probabilités d'arrêt précoce du traitement :

- La méthode de Pocock utilise un seuil constant à chaque analyse. Elle expose à un risque élevé d'arrêt prématuré.
- La méthode de O'Brien et Fleming est couramment utilisée car elle a un impact faible sur l'effectif global. En revanche la probabilité d'arrêt précoce est plus faible, ce qui dans certaines situations (comme les essais plateformes, où plusieurs traitements sont en concurrence), peut être jugé trop conservateur.
- La méthode de Lan et DeMets généralise cette approche en utilisant une fonction de consommation du risque alpha au cours du temps. Elle est plus flexible et permet des analyses « flottantes », basée sur une fraction d'information non définie *a priori*. Ainsi, ni le nombre d'analyses intermédiaires ni le moment de leur réalisation nécessitent d'être fixés à l'avance. Elle s'affranchit ainsi des contraintes logistiques liées au gel partiel de la base de données à un instant précis.

Notons que l'utilisation de ces seuils peut avoir pour finalité d'évaluer la supériorité et/ou la futilité. Le gain en termes de flexibilité (en l'occurrence la possibilité d'arrêter l'essai plus précocement) se fait au détriment d'une légère augmentation du nombre total de patients à inclure, qui s'explique par la multiplicité des comparaisons. Dans le cas des règles d'arrêt pour futilité, on distingue deux situations : en l'absence d'engagement sur l'arrêt de l'essai (*nonbinding stopping rules*) le comité de pilotage de l'essai peut décider de poursuivre l'étude même si les critères d'arrêt pour futilité sont satisfaits. Cette approche est plus flexible que les règles avec engagement d'arrêt (*binding*), au détriment là encore d'une légère augmentation de l'effectif total.

Pour les designs d'enrichissement, qui peuvent se baser sur l'utilisation d'un biomarqueur, il est impératif que celui-ci soit déterminé *a priori*, ou si c'est en cours d'essai que son utilisation ne soit pas motivée par les données de l'essai [3].

Notons que certaines méthodes adaptatives sont non comparatives, elles n'auront donc pas d'impact sur le risque d'erreur de type I. Il s'agit par exemple de la ré-estimation du calcul du nombre de sujets

basée sur la variance du critère de jugement principal recueilli dans l'étude, sans levée d'aveugle, ou encore la valeur pronostique d'un biomarqueur pour les designs d'enrichissement.

Un autre aspect essentiel des essais adaptatifs est de **maintenir l'intégrité de l'essai, alors que l'accès à une partie de l'information est nécessaire** pour effectuer les adaptations prévues. Aux problématiques méthodologiques énoncées ci-dessus s'ajoutent donc des contraintes logistiques visant à garantir :

- L'obtention de données de bonne qualité pour la réalisation des analyses intermédiaires, au moment où celles-ci sont prévues.
- Une restriction de l'accès aux résultats intermédiaires comparatifs (particulièrement aux personnes directement impliquées dans la conduite de l'essai).

Les moyens mis en œuvre peuvent parfois être complexes et ajouter des coûts à l'essai.

3 Etudes de cas

L'essai dit PARADIGM HF [6] est une étude contrôlée, randomisée, en double-aveugle, comparant l'association sacubitril + valsartan à l'enalapril chez des patients ayant une insuffisance cardiaque à fraction d'éjection diminuée. Le critère de jugement principale était un composite de décès de cause cardiovasculaire ou d'hospitalisations pour insuffisance cardiaque. Le protocole prévoyait initialement deux analyses intermédiaires :

The O'Brien-Fleming type of boundary with Lan-DeMets alpha spending function will be used for the interim efficacy analyses to assess superiority. As currently planned, two interim efficacy analyses are to be expected approximately at 1/3 and 2/3 of information time (i.e. approximately 803 and 1607 patients, respectively, with a primary events of CV mortality or HF hospitalization). The interim efficacy analysis with the boundary will spend approximately an alpha of 0.0001 (one-sided) at the first interim analysis and 0.00605 (one-sided) at the second interim analysis. The actual alpha to be spent for the interim efficacy analyses will be precisely determined based on the Lan-DeMets alpha spending function using the actual number of patients who have experienced a primary events at the interim efficacy analyses.

Finalement une troisième analyse intermédiaire a été ajoutée à la demande du DMSB, en adaptant les règles d'arrêt (ce qui est permis en utilisant la méthode flexible de Lan et DeMets). Une p-value nominale <0.001 ayant été obtenue lors de la 3^{ème} analyse intermédiaire, l'essai a été arrêté précocement pour supériorité, bien que le recrutement ait été terminé au moment de cette analyse [7].

Mi-2020 un essai contrôlé, randomisé, en double-aveugle, contre placebo, a évalué l'efficacité de l'hydroxychloroquine par rapport à un placebo dans la prévention de la COVID-19, chez 821 adultes asymptomatiques ayant été en contact avec une personne contaminée. Le protocole prévoyait au total trois analyses intermédiaires. L'incidence d'une maladie symptomatique suite à une exposition étant encore peu connue au début de l'essai, une analyse intermédiaire ayant pour objectif un re-calcule d'effectif était prévue au protocole [8] :

Because the estimates for both incident symptomatic COVID-19 after an exposure and loss to follow-up were relatively unknown in early March 2020, the protocol prespecified a sample-size reestimation at the second interim analysis. This reestimation, which used the incidence of new infections in the placebo group and the observed percentage of participants lost to follow-up, was aimed at maintaining the ability to detect an effect size of a 50% relative reduction in new symptomatic infections.

Cette analyse intermédiaire a permis de réduire l'effectif total de l'étude, compte-tenu d'une survenue d'infections symptomatiques dans le groupe contrôle plus importante qu'initialement prévue.

Par ailleurs, la puissance conditionnelle, qui est la probabilité que le résultat final de l'étude soit statistiquement significatif, compte tenu des données observées jusqu'alors [4], était estimée à chacune des analyses intermédiaires. Cette approche a permis l'arrêt prématuré de l'essai pour futilité à la troisième analyse, la puissance conditionnelle étant alors un inférieure à 1% [8].

4 Méta-recherche

Une étude de méta-recherche conduite dans le domaine de la sclérose latérale amyotrophique montre que parmi l'ensemble des essais contrôlés randomisés, peu adoptent un schéma séquentiel en groupes. Pourtant, cette approche permet fréquemment de réduire la durée des études [9].

En revanche, une étude de coûts assez récente confirme que les différents designs adaptatifs sont plus consommateurs de ressources que les designs fixes traditionnels [10].

5 Conclusion

Par définition, un essai adaptatif prévoit *a priori*, dans le protocole, la possibilité de modifier certains aspects de l'essai selon les données accumulées. La méthodologie est ainsi adaptée pour prendre en compte les éventuels risques associés à ces modifications. Les essais adaptatifs regroupent un ensemble hétérogène de designs. Les designs séquentiels en groupes et la réévaluation du nombre de sujets nécessaires en cours d'essai sont des approches robustes, désormais courantes, **bien acceptées par les autorités et les agences**. Toutefois nous notons trois points essentiels :

Le type d'adaptation et les méthodes utilisées doivent être définis avant le début de l'étude, ou si s'ils sont ajoutés en cours d'étude, que ce soit indépendamment des résultats obtenus.

Sur le plan méthodologique, les analyses multiples augmentent le risque de conclure à tort à l'efficacité d'un nouveau traitement ; elles nécessitent donc la **mise en œuvre de méthodes adaptées pour prendre en compte cette inflation du risque d'erreur de type I**.

L'accès aux données accumulées lors de l'essai pour réaliser ces analyses ajoute des contraintes logistiques, avec une augmentation des coûts associés. Par ailleurs, pour respecter l'intégrité et la validité de l'essai, **un DSMB avec des membres expérimentés sur la gestion de ces designs est indispensable** [3].

Références

- 1 U.S. Department of Health, Human Services - Food, Drug Administration. Adaptive Designs for Clinical Trials of Drugs and Biologics - Guidance for Industry 2019 ;
- 2 Bhatt DL, Mehta C, Drazen JM, et al. Adaptive Designs for Clinical Trials. *New Engl J Med* 2016;375:65–74 doi:10.1056/NEJMra1510061;
- 3 Porcher R, Lecocq B, Vray M, et al. Les méthodes adaptatives quand et comment les utiliser dans les essais cliniques ? *Thérapie* 2011;66:309–17 doi:10.2515/therapie/2011042;
- 4 Lachin JM. A review of methods for futility stopping based on conditional power. *Stat Med* 2005;24:2747–64 doi:10.1002/sim.2151;
- 5 Thall P, Fox P, Wathen J. Statistical controversies in clinical research: scientific and ethical problems with adaptive randomization in comparative clinical trials. *Annals of Oncology* 2015;26:1621–28 doi:10.1093/annonc/mdv238;
- 6 McMurray JJV, Packer M, Desai AS, et al. Angiotensin–Neprilysin Inhibition versus Enalapril in Heart Failure. *New Engl J Med* 2014 doi:10.1056/NEJMoa1409077;
- 7 McMurray JJV, Ostergren J, Swedberg K, et al. Effects of candesartan in patients with chronic heart failure and reduced left-ventricular systolic function taking angiotensin-converting-enzyme inhibitors: the CHARM-Added trial. *The Lancet* 2003;362:767–71 doi:10.1016/S0140-6736(03)14283-3; PMID:13678869;
- 8 Boulware DR, Pullen MF, Bangdiwala AS, et al. A Randomized Trial of Hydroxychloroquine as Postexposure Prophylaxis for Covid-19. *New Engl J Med* 2020 doi:10.1056/NEJMoa2016638;
- 9 van Eijk RPA, Nikolakopoulos S, Ferguson TA, et al. Increasing the efficiency of clinical trials in neurodegenerative disorders using group sequential trial designs. *Journal of Clinical Epidemiology* 2018;98:80–88 doi:10.1016/j.jclinepi.2018.02.013;
- 10 Wilson N, Biggs K, Bowden S, et al. Costs and staffing resource requirements for adaptive clinical trials: quantitative and qualitative results from the Costing Adaptive Trials project. *BMC Med* 2021;19:1–17 doi:10.1186/s12916-021-02124-z;