

## THERAPIES

### HEADING : ATELIERS DE GIENS 2018 / Pharmacologie clinique

#### Des études mono-bras aux études de comparaison externe. Considérations méthodologiques et recommandations\*

Les études mono-bras

**Michel Cucherat<sup>a,\*</sup>, Silvy Laporte<sup>b</sup>, Olivier Delaitre<sup>c</sup>, Jehan-Michel Behier<sup>d</sup>, les participants à la table ronde « Recherche clinique » des Ateliers de Giens XXXV : Anne d'Andon<sup>e</sup>, Florence Binlich<sup>f</sup>, Serge Bureau<sup>g</sup>, Catherine Cornu<sup>h</sup>, Cécile Fouret<sup>i</sup>, Natalie Hoog Labouret<sup>j</sup>, Bruno Laviolle<sup>k</sup>, Houda Miadi-Fargier<sup>l</sup>, Xavier Paoletti<sup>m</sup>, Matthieu Roustit<sup>n</sup>, Tabassome Simon<sup>o</sup>, Nathalie Varoqueaux<sup>p</sup>, Eric Vicaut<sup>q</sup>, Jérémie Westerloppe<sup>r</sup>**

<sup>a</sup> *Service de pharmacotoxicologie, hospices civils de Lyon, 69003 Lyon, France ; Université Lyon 1, CNRS, UMR5558, laboratoire de biométrie et biologie évolutive, 69008 Lyon, France*

<sup>b</sup> *Unité de recherche clinique innovation pharmacologie, hopital Nord, batiment recherche, CHU Saint-Etienne ; UMR 1059 Sainbiose, Université Jean Monnet, 42022 Saint-Etienne, France*

<sup>c</sup> *Boehringer Ingelheim France, 75013 Paris, France*

<sup>d</sup> *Celgene, 92066 Paris La Défense cedex, France*

<sup>e</sup> *CEMKA, 92340 Bourg la Reine, France*

<sup>f</sup> *Servier, 92284 Suresnes, France*

<sup>g</sup> *AP-HP, 75010 Paris, France*

<sup>h</sup> *INSERM, CIC1407, hospices civils de Lyon, 69000, Lyon, France ; UMR 5558, université Claude Bernard Lyon 1, 69100 Lyon, France*

<sup>i</sup> *Medtronic France, 92100, Boulogne-Billancourt, France*

<sup>j</sup> *Institut national du cancer, 92513 Boulogne Billancourt, France*

<sup>k</sup> *Univ Rennes, CHU Rennes, service de Pharmacologie ; Inserm, CIC 1414, centre d'investigation clinique de Rennes, 35000 Rennes, France*

<sup>l</sup> *Bayer HealthCare, 59120 Loos, France*

<sup>m</sup> *Université de Versailles St-Quentin, INSERM U900, équipe de biostatistiques pour la médecine de précision et Institut Curie, 92210 Saint Cloud, France*

<sup>n</sup> *Univ Grenoble Alpes, HP2, Inserm, CHU Grenoble Alpes, 38000 Grenoble, France*

<sup>o</sup> *Service de pharmacologie clinique, plateforme de recherche clinique de l'Est parisien (URCEST-CRCEST-CRBHUEP-SU), Assistance Publique-Hôpitaux de Paris, hôpital Saint Antoine, Paris, France ; Sorbonne université, université Pierre et Marie Curie site St Antoine,*

75012 Paris, France ; FACT (French Alliance for CV Clinical Trials), réseau F-CRIN ;  
INSERM U1148, 75877 Paris, France

<sup>p</sup> Astra Zeneca, 92400 Courbevoie, France

<sup>q</sup> Unité de recherche clinique, hôpital Fernand Widal, AP-HP, 75010 Paris, France

<sup>r</sup> Bristol-Myers Squibb, 92500 Rueil-Malmaison, France

Texte reçu le 14 novembre 2019 ; accepté le 22 novembre 2019

\*Auteur correspondant. Service de pharmacotoxicologie, hospices civils de Lyon, 162 avenue Lacassagne, 69003 Lyon, France.

Adresse e-mail : michel.cucherat@univ-lyon1.fr (M. Cucherat)

\* Les articles, analyses et propositions issus des Ateliers de Giens sont ceux des auteurs et ne préjugent pas des propositions de leur organisation.

Journal Pre-proof

## Résumé

Les études mono-bras sont parfois utilisées comme études pivots mais elles présentent des limites méthodologiques qui ne leur permettent pas d'obtenir un niveau de fiabilité élevé comme peut l'être celui de l'essai contrôlé randomisé qui reste le standard de l'évaluation des nouveaux traitements. L'objectif de cette table ronde était de discuter des limites de ces études, d'analyser les solutions disponibles et acceptables afin de proposer des recommandations pour leur réalisation et leur évaluation. Les études mono-bras proprement dites sont intrinsèquement inadaptées pour démontrer le bénéfice d'un nouveau traitement car il est impossible de déduire ce bénéfice à partir d'une valeur sous traitement sans connaître ce qu'elle aurait été en l'absence du nouveau traitement. Une comparaison avec d'autres données est alors implicite. Cependant, cette comparaison est limitée par 1) le choix post hoc de la référence de la comparaison, 2) le biais de confusion qui nécessite impérativement une approche d'ajustement, et 3) les autres biais, de mesure et d'attrition entre autres. La prise en considération de ces limites devrait conduire, avant tout, à la réalisation d'essais à contrôle externe (« *externally controlled trial* ») à la place des essais mono-bras comme le propose la dernière version d'ICH E10. Par ailleurs, la comparaison externe doit être formalisée dans le protocole de l'étude avec un choix a priori de la référence de comparaison ainsi que la méthode de comparaison formelle : test par rapport à une norme, ajustement avec des données individuelles, groupe contrôle synthétique ou « *matching-adjusted indirect comparisons (MAIC)* ». Finalement, l'essai de comparaison externe doit être réservé aux situations où la randomisation est irréalisable. Pour être acceptable, ces études doivent pouvoir donner la garantie d'absence de biais de confusion résiduelle, ce qui n'est vraiment recevable que si l'effet observé est extrêmement important et l'évolution de la maladie fortement prévisible.

## MOTS CLÉS

Études mono-bras ; Comparaison indirecte ; Comparaisons indirectes ajustées ; Non-comparative ; Groupe contrôle ; Étude de contrôle externe

## Abréviations

AMM : autorisation de mise sur le marché

DM : dispositifs médicaux

FDA : *Food and Drug Administration*

ICH : *International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use*

MAIC : *matching-adjusted indirect comparisons*

NICE : National Institute for Health and Care Excellence

TTP : traitement ciblé sur le temps avant progression

TTP1 : TTP sous traitement standard

TTP2 : TTP sous thérapie ciblée

## Introduction

La méthodologie des essais contrôlés randomisés a été construite pour permettre une évaluation fiable, à l'abri des biais, de l'efficacité et de la sécurité propres des traitements. En l'absence d'alternative méthodologique permettant de donner le même niveau de garantie, l'essai contrôlé randomisé reste le standard de l'évaluation des nouveaux traitements.

Les études mono-bras (*single-arm studies*) intégrant une intervention thérapeutique consistent en une série prospective de patients recevant tous le traitement à évaluer. Ces études permettent une estimation de la valeur du critère de jugement sous traitement uniquement, « dans l'absolu » (% de décès à 12 mois par exemple). Sans comparatif, l'apport du nouveau traitement par rapport au placebo ou au traitement de référence n'est pas documenté. Ainsi, les études mono-bras ne permettent pas d'apporter directement la démonstration du bénéfice d'un nouveau traitement. Pourtant, de plus en plus d'études mono-bras sont proposées comme étude « pivot » pour l'enregistrement, le remboursement ou la modification des pratiques, que ce soit pour des médicaments ou des dispositifs médicaux (DM) [cf. infra].

L'objectif de cette table ronde était de discuter des limites de ces études, d'analyser les solutions disponibles et acceptables afin de proposer des recommandations pour leur réalisation et leur évaluation critique.

## Contexte et méta-épidémiologie

Bien qu'aucune revue méta-épidémiologique exhaustive n'ait été publiée, des premières données montrent que l'utilisation des études mono-bras connaît une nette augmentation au cours des dernières années. Entre 2000-2016, parmi les évaluations réalisées par le *National Institute for Health and Care*

*Excellence* (NICE), 22 sur 489 (4 %) étaient basées sur une étude mono-bras, dont la moitié en 2015-2016 ; 10/22 (45 %) de ces recommandations ont été positives [1].

En France, l'analyse des avis de la commission de la transparence entre janvier 2017 et juin 2019, montre que, sur 75 avis en oncologie et hémato-oncologie, 16 (21 %) étaient basées, entre autres, sur des études mono-bras non comparatives. Dans le cadre des maladies rares, la proportion était de 22/67 (33 %).

Les essais mono-bras comme études pivots sont actuellement utilisés dans différents contextes. Dans le domaine du DM, les études mono-bras permettent l'évaluation des évolutions d'un dispositif dont l'intérêt a initialement été démontré par un essai clinique randomisé. De même dans les maladies rares, ces études sont parfois choisies comme seule étude pivot en raison du petit nombre de patients disponibles pour effectuer des essais cliniques randomisés.

Mais les études mono-bras peuvent aussi être à l'origine de certaines autorisations de mise sur le marché (AMM) alors qu'elles sont conduites dans des domaines où la réalisation d'essais contrôlés randomisés aurait pu être possible (ou en cours de réalisation en cas d'enregistrements accélérés). Elles posent alors le problème de la question de leur prise en compte, « la main forcée », au niveau de l'accès au marché et de l'élaboration des recommandations cliniques.

Les essais à contrôle externe sont utilisés pour l'évaluation des traitements dans les maladies rares, pour lesquelles le recrutement est difficile, du fait de leur hétérogénéité et des petits effectifs de patients. Quatre essais de phase 2 ont ainsi évalué l'efficacité de l'imatinib sur la fibrose cutanée de patients atteints de sclérodémie systémique. Parmi eux, trois étaient conduits en ouvert sans groupe contrôle. Les résultats montrent une diminution de la fibrose cutanée (évaluée par le score clinique de Rodnan) après 12 mois de traitement par imatinib (NCT00613171) [2]. En revanche, le quatrième essai de phase 2, contrôlé, randomisé et en double-aveugle, n'a pas permis de montrer une supériorité de l'imatinib par rapport au placebo sur la fibrose cutanée ; par contre, il y avait plus d'effets indésirables dans le groupe imatinib [3]. Face à ces résultats, l'imatinib n'a pas été testé en phase 3. Plus récemment, le tocilizumab a été évalué dans la même indication, mais l'essai de phase 2 était contrôlé et randomisé [4].

Il est cependant probable que les essais à contrôle externe connaissent dans les maladies rares un essor comparable à celui observé dans le domaine de la cancérologie. En effet, un document récent de la *Food and Drug Administration* (FDA) présente les spécificités et fournit des recommandations sur les études à contrôle externes (qualifiées ici de « *natural history studies* ») dans le contexte des maladies rares [5].

### **Les limites méthodologiques des études mono-bras**

En l'absence de comparaison avec un groupe contrôle, il est impossible de savoir si la valeur du critère de jugement observée sous traitement est meilleure que celle qui aurait été obtenue sans ce traitement dans la même étude : mêmes patients, même mesure du critère de jugement, même suivi, même contexte de soins.

La seule situation où il est raisonnablement possible de conclure sans comparaison est celle dite du « 100 % / 0 % », correspondant aux situations où l'on observe quelques évolutions favorables voire la suppression complète des évolutions péjoratives sous traitement dans une pathologie où

l'évolution, avec la prise en charge de référence, serait péjorative dans 100 % des cas. Cette recommandation est également celle du NCI américain pour les essais de combinaisons [6]. Les essais simples bras devraient ainsi être restreints aux cas où l'activité avec le traitement de référence est minimale, et la randomisation irréalisable. Ce type de situation est exceptionnel : même dans le cas de l'infection par le virus Ebola par exemple, le pronostic, bien que très péjoratif, n'est pas suffisamment sombre pour entrer dans la règle des 0 % / 100 % puisque certains patients survivent spontanément. Un essai randomisé a donc été réalisé pour évaluer l'efficacité clinique du ZMapp dans l'infection par le virus Ebola [7].

L'exploitation des études mono-bras pour démontrer le bénéfice d'un nouveau traitement nécessite donc une comparaison implicite qui fera obligatoirement appel à une référence de comparaison externe (comparaison historique par exemple). Plusieurs types de comparateurs externes sont envisageables (Tableau 1), ainsi que des méta-analyses lorsque plusieurs études identiques sont disponibles ([8] pour un exemple de méta-analyse réalisée dans ce but).

Les comparaisons externes présentent plusieurs limites importantes sur le plan méthodologique.

- 1) Le choix de la référence de comparaison est souvent effectué de manière *post hoc* : de ce fait, rien ne permet de garantir que le choix du comparateur n'a pas été délibérément fait pour favoriser le nouveau traitement. Ces études ne doivent plus être conçues comme des essais mono-bras descriptifs, mais comme de véritables essais comparatifs avec un groupe contrôle externe prédéfini : l'objectif explicite est bien la démonstration de l'intérêt clinique du nouveau traitement et non la simple description de l'évolution des patients avec le traitement. Après recensement de toutes les cohortes existantes, incluant une revue systématique de la littérature, le choix de la référence de comparaison est alors fixé dès la conception du protocole, tout comme la méthode d'analyse, empêchant ainsi tout choix *post hoc*. Le choix de la ou des cohortes de comparaison doit être argumenté. Cette approche est d'ailleurs avalisée dans la nouvelle version de ICH E10 qui propose le terme d'essai à contrôle externe (*externally controlled trial*) [9]. De plus, la détermination a priori de la référence de comparaison au moment de la construction du protocole de l'étude permet également une meilleure prévention des biais en utilisant, par exemple, le même critère de jugement et la même méthode de mesure (cf. infra). L'idéal est de disposer des données individuelles des patients pour permettre une meilleure prise en compte des facteurs de confusion (scores de propension, analyses multivariées ...) et ce point devra être pris en compte dans le choix des comparateurs externes.
- 2) L'autre limite majeure des comparaisons externes est le biais de confusion : rien ne garantit que les groupes ou les résultats comparés sont comparables en termes de patients. Aussi, plusieurs méthodes de formalisation des comparaisons externes (aussi appelées comparaisons indirectes [*unanchored indirect comparisons*]) ont été proposées afin de tenter de fiabiliser les conclusions tirées des études à comparateur externe [10-13]. Ces approches ont toutes pour objectif de

suppléer l'absence de contrôle par design des facteurs de confusion en les prenant en compte au niveau de l'analyse. Les méthodes permettent, **en théorie**, de supprimer le biais de confusion à condition qu'absolument **tous** les facteurs de confusion puissent être pris en compte (facteurs conditionnant le critère de jugement et de distribution différente entre la cohorte traitée et son comparateur, y compris les modificateurs de l'effet du traitement [14]). Pour cela, les facteurs de confusion potentiels de chaque critère de jugement doivent d'abord être identifiés par une revue systématique des études pronostiques. L'ajustement réalisé (quelle que soit la méthode) doit prendre en compte tous ces facteurs qui doivent être prédéfinis au protocole également. Afin de confirmer que ces ajustements aient permis d'atteindre le but escompté, l'importance du biais de confusion résiduel peut être appréciée par l'utilisation de contrôles négatifs comme variables de falsification et par une « *bias analysis* » [15]. Compte tenu des difficultés rencontrées pour apporter ces garanties, les résultats ne seront convaincants qu'en cas de résultats de grande ampleur, ne pouvant être seulement expliqués par les biais et la confusion résiduelle. Le choix a priori de la référence de comparaison au moment de la construction du protocole contribuera aussi à la minimisation du biais de confusion, en permettant d'aligner les critères d'éligibilité avec ceux utilisés pour la construction du contrôle externe (approche par restriction). Cette limite du biais de confusion fait qu'au niveau réglementaire, les recommandations de l'*International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use* (ICH) E10 n'envisagent le recours aux essais à contrôles externes que dans des situations très particulières où l'effet du traitement est extrêmement important et le cours de la maladie hautement prévisible (« *Use of the external control design is restricted to situations in which the effect of treatment is dramatic and the usual course of the disease highly predictable* ») [16].

- 3) D'autres limites existent en termes de biais de mesure, de sélection et d'attrition. Le Tableau 2 résume ces limites et les solutions envisageables. Ces biais sont inhérents au design et ne pourront pas être corrigés par des analyses statistiques. L'évaluation du risque de biais à l'aide de la grille de Robin-I devrait permettre de donner un degré de confiance que l'on peut avoir dans les résultats.

Tout comme les études contrôlées randomisées, une étude de comparaison externe doit aussi assurer la pertinence clinique de la comparaison effectuée : critères de jugement principaux pertinents, comparateur loyal représentant le « standard of care » valide au moment de la prise de décision, durée de suivi, évaluation de la sécurité. En d'autres termes, sur tous ces points, **l'étude doit être comparable à ce qu'aurait été un essai randomisé pivot**.

### **Méthodes de formalisation des comparaisons externes**

## Comparaison à une norme

La forme la plus simple de formalisation de la comparaison externe est la comparaison statistique à une valeur de référence, prédéfinie au protocole et censée représenter la valeur attendue sans le nouveau traitement. Cette approche est limitée par l'impossibilité d'ajuster la valeur de référence en fonction des caractéristiques des patients effectivement inclus et par la possibilité que la valeur choisie sous-estime la performance du traitement contrôle. Cette valeur doit donc être solidement justifiée au protocole, en particulier par une revue systématique de la littérature.

## Ajustement avec des données agrégées ou *matching-adjusted indirect comparisons* (MAIC)

La méthode de « *matching-adjusted indirect comparisons* » (MAIC) [11, 12, 14] permet d'utiliser comme contrôle externe par exemple les données agrégées d'une publication (cohorte historique, essai clinique de validation du traitement de référence, etc.). Elle permet de réaliser un ajustement par pondération des patients de l'étude mono-bras (pour laquelle on dispose des données individuelles) pour les rendre comparables, sur les variables disponibles, à la population de la publication. Le résultat ajusté de la cohorte traitée devient ainsi comparable, dans une certaine mesure, à celui rapporté dans la publication de référence et permet d'estimer l'apport du nouveau traitement.

## Ajustement avec des données individuelles

Lorsque les données individuelles du comparateur externe sont disponibles (cohortes, groupes contrôle de ou des essais du traitement de référence, etc.), la comparaison peut utiliser les mêmes techniques d'ajustement que les études observationnelles de pharmaco-épidémiologie par exemple. L'ajustement peut être fait par *matching*, stratification, régression ou pondération en utilisant comme covariables les facteurs de confusion potentiels eux-mêmes ou un score de propension. Cette approche est certainement la plus propice à faire disparaître le biais de confusion si tous les facteurs de confusion ont été pris en compte.

Dans ce cadre, l'approche des « *synthetic control groups*<sup>TM</sup> »<sup>1</sup> est largement employée. De manière générique, elle consiste à extraire d'une cohorte de référence (par exemple bases de données médico-administratives, entrepôts de données des hôpitaux ...) un échantillon de patients appariés aux patients de la cohorte mono-bras traitée avec le nouveau traitement.

Cependant, l'utilisation même d'une cohorte de référence constituée et suivie en même temps que l'étude mono-bras remet en question l'impossibilité de randomiser et rend difficilement acceptable l'évaluation basée sur un essai mono-bras.

## Autres méthodes

D'autres approches sont parfois utilisées comme des comparaisons avant/après ou des simulations du groupe contrôle manquant (*simulated treatment comparison*) [12].

Les comparaisons avant/après peuvent être particulièrement biaisées, notamment par des biais de sélection, car l'inclusion dans la cohorte traitée se fait souvent conditionnellement à l'évolution précédente du patient (échappement ou non-réponse au traitement précédent par exemple). En oncologie par exemple, Von Hoff a proposé d'évaluer l'impact d'un nouveau traitement ciblé sur le temps avant progression (TTP) chez des patients avec des maladies réfractaires en utilisant le patient

<sup>1</sup> Le terme « *synthetic control arm* » est une marque déposée de Medidata Solutions, Inc.



comme son propre témoin [17]. Ces patients reçoivent la thérapie ciblée après progression sous un traitement standard. Le TTP sous thérapie ciblée (TTP2) est comparé au TTP sous traitement standard (TTP1). C'est une sorte de plan en cross-over sans randomisation de la séquence pour contrôler l'effet temps. Sous l'hypothèse que (i) le cancer est de plus en plus agressif avec le temps et (ii) que des résistances multiples se développent, on s'attendrait à ce que  $TTP2 < TTP1$  (ou le ratio  $TTP2 / TTP1 < 1$ ). Von Hoff a suggéré qu'un ratio  $> 1,3$  était indicateur d'un bénéfice individuel au patient. Toutefois, ces hypothèses d'une histoire naturelle qui irait en s'accéléralant pour tous les patients sont contredites par des études récentes [18, 19]. La TTP1 prédit assez mal la TTP2 et donc ce ratio est peu corrélé à la survie globale.

L'approche de « *simulated treatment comparison* » fait quant à elle appel à une modélisation du critère de jugement à partir des données de la cohorte traitée, modélisation dont la performance prédictive n'est pas vérifiable en soi.

## Recommandations

- L'essai contrôlé randomisé reste l'approche de référence pour l'évaluation des nouveaux traitements, les comparaisons externes formalisées ne permettant pas de fournir les mêmes garanties de fiabilité des résultats, ni même le même niveau de preuve.
- Le recours éventuel à une étude de comparaison externe doit être solidement justifié par des arguments indubitables.

Deux situations ont été distinguées : la mise en place d'une nouvelle étude et l'exploitation « post-hoc » d'une étude déjà réalisée ; mais une première recommandation est commune à ces 2 situations.

## Recommandations pour la conception d'une nouvelle étude

- Concevoir l'étude comme une étude de comparaison externe (et non comme une étude mono-bras descriptive) : la comparaison externe dans le but de démontrer le bénéfice du nouveau traitement est l'objectif initial de l'étude.
- Garantir la pertinence clinique des critères de jugement, des traitements comparateurs, des patients inclus comme dans une étude pivot d'enregistrement, de remboursement ou de changement des pratiques.
- Déterminer et justifier a priori le contrôle externe, avec une revue systématique de la littérature pour rechercher toutes les cohortes publiées, et le recensement des cohortes existantes.
- Privilégier les comparateurs externes qui permettent des analyses ajustées : c'est à dire dont on a accès aux données individuelles disponibles, ou cohortes publiées les mieux documentées.
- Documenter a priori les facteurs de confusion potentiels par une revue systématique.
- Définir a priori la stratégie d'évaluation du biais de confusion résiduel.

- Prévoir a priori un contrôle et une discussion des autres biais (mesure, sélection, attrition)
- Définir a priori le plan d'analyse statistique (choix de la méthode détaillée, analyses de sensibilité).
- Tous ces éléments doivent figurer dans un protocole daté et enregistré dans une base de données telle [clinicalTrials.gov](http://clinicaltrials.gov)

### **Recommandations pour l'exploitation post hoc d'une étude déjà réalisée**

Cette situation correspond à une situation de dégradation méthodologique très importante qu'il convient d'éviter. Dans de nombreux cas, il sera impossible de relever le niveau de preuve de l'étude, car tous les choix de comparateurs et de méthodes s'effectuent, par définition, de manière post hoc, et il est impossible d'apporter les garanties permettant d'exclure des choix influencés par les résultats. Pour tenter de lever au maximum les écueils intrinsèques du post hoc, la table ronde recommande de :

- Documenter tous les contrôles envisageables pour cette étude : revue systématique des données agrégées publiées (cohortes de patients, bras contrôle d'essais, etc.) et lister toutes les cohortes de patients existantes pouvant servir potentiellement de contrôles externes.

Tester la robustesse des résultats en montrant que la vibration des résultats conduit à la même décision, quel que soit le contrôle utilisé.

### **Conclusion**

Les études mono-bras sont intrinsèquement inadaptées pour démontrer le bénéfice d'un nouveau traitement, car il est impossible de déduire ce bénéfice à partir d'une valeur sous traitement sans connaître ce qu'elle aurait été en l'absence du nouveau traitement.

Pour remplir cet objectif, il est impératif d'effectuer cette comparaison : si la réalisation d'une étude comparative à groupe contrôle interne n'est pas envisagée, la seule alternative possible est d'effectuer cette comparaison à l'aide d'un contrôle externe. Dans ce contexte, les études ne doivent plus être de simples études mono-bras purement descriptives, mais doivent être conçues comme de véritables études contrôlées, à groupe contrôle externe. Ces études doivent être construites de façon à éviter toute comparaison post hoc, remettant fortement en question la validité des conclusions produites. Même à cette condition, les résultats produits sont fortement compromis par la présence de biais de confusion liée à la non-comparabilité des patients entre la cohorte traitée et le contrôle externe. Pour être acceptables, ces études doivent pouvoir donner la garantie d'absence de biais de confusion résiduelle. Il doit être aussi possible d'écarter la présence des autres biais (mesure, sélection et attrition). Enfin, l'effet doit être suffisamment important pour qu'il ne puisse pas être la conséquence uniquement des facteurs de confusion résiduels ou des biais.

### **Déclaration de liens d'intérêts**

MC Interventions ponctuelles : activités de conseil pour BMS, Bayer, AstraZeneca, Sanofi.

SL L'auteur déclare ne pas avoir de liens d'intérêts

OD L'auteur déclare ne pas avoir de liens d'intérêts

JMB L'auteur déclare ne pas avoir de liens d'intérêts

AA L'auteur déclare ne pas avoir de liens d'intérêts

FB L'auteur déclare ne pas avoir de liens d'intérêts

SB L'auteur déclare ne pas avoir de liens d'intérêts

CC L'auteur déclare ne pas avoir de liens d'intérêts

CF L'auteur déclare ne pas avoir de liens d'intérêts

NHL L'auteur déclare ne pas avoir de liens d'intérêts

BL L'auteur déclare ne pas avoir de liens d'intérêts

HMF L'auteur déclare ne pas avoir de liens d'intérêts

XP L'auteur déclare ne pas avoir de liens d'intérêts

MR L'auteur déclare ne pas avoir de liens d'intérêts

TS L'auteur déclare ne pas avoir de liens d'intérêts

NV L'auteur déclare ne pas avoir de liens d'intérêts

EV L'auteur déclare ne pas avoir de liens d'intérêts

JW L'auteur déclare ne pas avoir de liens d'intérêts

TS Activité de conseil et d'expertise, présence au board : Astrazeneca, Novartis, Sanofi, MSD, Astellas. Grants obtenus pour des recherches : Astrazeneca, Boehringer, Daiichi-Sankyo, Eli Lilly, GSK, MSD, Novartis, Sanofi

## Références

[

[1] Anderson M, Naci H, Morrison D, Osipenko L, Mossialos E. A review of NICE appraisals of pharmaceuticals 2000-2016 found variation in establishing comparative clinical effectiveness. *J Clin Epidemiol* 2019;105:50-9.

[2] Spiera RF, Gordon JK, Mersten JN, Magro CM, Mehta M, Wildman HF, et al. Imatinib mesylate (Gleevec) in the treatment of diffuse cutaneous systemic sclerosis: results of a 1-year, phase IIa, single-arm, open-label clinical trial. *Ann Rheum Dis* 2011;70(6):1003-9.

[3] Prey S, Ezzedine K, Doussau A, Grandoulier AS, Barcat D, Chatelus E, et al. Imatinib mesylate in scleroderma-associated diffuse skin fibrosis: a phase II multicentre randomized double-blinded controlled trial. *Br J Dermatol* 2012;167(5):1138-44.

[4] Khanna D, Denton CP, Jhrees A, van Laar JM, Frech TM, Anderson ME, et al. Safety and efficacy of subcutaneous tocilizumab in adults with systemic sclerosis (faSScinate): a phase 2, randomised, controlled trial. *Lancet* 2016;387(10038):2630-40.

[5] Food and Drug Administration. Rare diseases: natural history studies for drug development. Guidance for industry. 2019. <https://www.fda.gov/media/122425/download>. [Consulté le 22 novembre 2019 (22 pp.)].

[6] Foster JC, Freidlin B, Kunos CA, Korn EL. Single-arm phase ii trials of combination therapies: a review of the CTEP experience 2008-2017. *J Natl Cancer Inst* 2019 Sep 23 pii: djz193.

[7] Group PIW, Multi-National PIIST, Davey RT, Jr., Dodd L, Proschan MA, Neaton J, et al. A randomized, controlled trial of ZMapp for Ebola virus infection. *N Engl J Med* 2016;375(15):1448-56.

[8] Korn EL, Liu PY, Lee SJ, Chapman JA, Niedzwiecki D, Suman VJ, et al. Meta-analysis of phase II cooperative group trials in metastatic stage IV melanoma to determine progression-free and overall survival benchmarks for future phase II trials. *J Clin Oncol* 2008;26(4):527-34.

[9] Choice of control group and related issues in clinical trials E10. ICH harmonised tripartite guideline: International Conference on Harmonisation of technical requirements for registration of pharmaceuticals for human use. 2000. <https://www.ich.org/>. [Consulté le 22 novembre 2019].

- [10] Phillippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. Methods for population-adjusted indirect comparisons in health technology appraisal. *Med Decis Making* 2018 Feb;38(2):200-211.
- [11] Signorovitch JE, Sikirica V, Erder MH, Xie J, Lu M, Hodgkins PS, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. *Value Health* 2012;15(6):940-7.
- [12] Signorovitch JE, Wu EQ, Yu AP, Gerrits CM, Kantor E, Bao Y, et al. Comparative effectiveness without head-to-head trials: a method for matching-adjusted indirect comparisons applied to psoriasis treatment with adalimumab or etanercept. *Pharmacoeconomics* 2010;28(10):935-45.
- [13] Caro JJ, Ishak KJ. No head-to-head trial? simulate the missing arms. *Pharmacoeconomics* 2010;28(10):957-67.
- [14] Phillippo DM, Dias S, Palmer S, Abrams KR, Welton NJ. NICE DSU technical support document 18: methods for population-adjusted indirect comparisons in submissions to NICE. NICE DSU technical support document NICE. Decembre 2016. <http://nicedsu.org.uk/wp-content/uploads/2018/08/Population-adjustment-TSD-FINAL-ref-rerun.pdf>. [Consulté le 22 novembre 2019 (81 pp.)].
- [15] Greenland S, Lash T. Bais analysis. In: Rothman K, Greenland S, Lash T, (Eds.), *Modern epidemiology*. 3rd ed. Philadelphia: Lippincott Williams and Wilkins 2008:345-80.
- [16] ICH. Choice of control group and related issues in clinical trials E10. International Conference on Harmonisation of technical requirements for registration of pharmaceuticals for human use. 2019. <https://www.ich.org/>. [Consulté le 22 novembre 2019].
- [17] Von Hoff DD, Stephenson JJ, Jr., Rosen P, Loesch DM, Borad MJ, Anthony S, et al. Pilot study using molecular profiling of patients' tumors to find potential targets and select treatments for their refractory cancers. *J Clin Oncol* 2010;28(33):4877-83.
- [18] Watson S, Menis J, Baldini C, Martin-Romano P, Michot JM, Hollebecque A, et al. Time to progression ratio in cancer patients enrolled in early phase clinical trials: time for new guidelines? *Br J Cancer* 2018;119(8):937-9.

- [19] Buyse M, Quinaux E, Hendlisz A, Golfinopoulos V, Tournigand C, Mick R. Progression-free survival ratio as end point for phase II trials in advanced solid tumors. *J Clin Oncol* 2011;29(15):e451-2.

Journal Pre-proof

**Tableau 1. Types de comparateurs externes possibles**

- Norme, valeur de référence qui peut être soit fixée par une « exigence réglementaire », soit issue de la littérature (étude de cohorte par exemple) ou d'une conférence de consensus
- Cohorte disponible sous la forme d'une publication (données individuelles inaccessibles) ou publication d'un essai randomisé dont un des groupes peut servir de comparaison externe qui pourra servir de base à une MAIC
- Cohorte avec accès aux données individuelles qui pourra servir à une approche type étude observationnelle
- Essai randomisé avec accès aux données individuelles qui pourra servir à une approche type étude observationnelle

MAIC : *matching-adjusted indirect comparisons*

**Tableau 2. Limites méthodologiques des comparaisons externes et solutions envisageables**

<b>Limites méthodologiques</b>	<b>Solutions envisageables</b>
Absence de groupe contrôle, impossibilité de déterminer l'efficacité du traitement évalué	Réaliser à la place des essais contrôlés à contrôle externe
Choix post hoc de la référence de comparaison	Intégrer le choix de la référence de comparaison a priori dans un protocole dès la conception de l'étude
Biais de confusion des comparaisons externes	Méthode d'ajustement prévue a priori, dès la conception de l'étude, assurant la prise en compte de tous les facteurs de confusion potentiels identifiés à partir d'une revue systématique des études pronostiques. Les autres variables conditionnant l'évolution des patients (contexte, variables socioéconomiques, pratique de soins) doivent aussi être considérées comme potentiels facteurs de confusion
Biais de mesure	Aligner le protocole de l'étude à contrôle externe sur celui de l'étude servant de comparateur (même définition du critère, même méthode de mesure, etc.)
Biais de sélection	Éviter les comparaisons avant/après
Biais d'attrition	Utiliser un contrôle externe de qualité, ayant assuré un suivi exhaustif des patients (faible réelle attrition)
Biais de confusion résiduelle	Prévoir des contrôles négatifs comme variable de falsification et une évaluation de la robustesse du résultat vis-à-vis des facteurs de confusion non pris en compte et des autres biais (« <i>bias analysis</i> »)



Journal Pre-proof