

# FORCES ET FAIBLESSES DES ESSAIS CLINIQUES

## Evolution en fonction de l'essor de la médecine personnalisée

PH. ERNEST (1), B. JANDRAIN (2), A.J. SCHEEN (1)

**RÉSUMÉ :** Les essais cliniques contrôlés représentent la pierre angulaire de la médecine factuelle. Construits selon les règles des bonnes pratiques cliniques, ils offrent de nombreuses forces, mais présentent également quelques faiblesses. La méthodologie rigoureuse utilisée permet, notamment, d'éviter les biais liés à des facteurs confondants (via un groupe contrôlé), les biais de sélection (via la randomisation) et les biais d'interprétation (via la technique du double aveugle). Cependant, les patients recrutés dans les essais cliniques et les conditions dans lesquelles ceux-ci sont menés diffèrent fondamentalement de la situation en vie réelle. De plus, les essais cliniques recrutent un mélange de patients bons et mauvais répondeurs, ce qui conduit à une réponse thérapeutique moyenne le plus souvent mitigée. Les essais cliniques devront s'adapter au développement de la médecine personnalisée pour devenir encore plus performants. Ils devront, à l'avenir, passer d'une analyse statistique sur de grands groupes à une analyse plus individualisée en fonction des caractéristiques phénotypiques et génotypiques des patients.

**MOTS-CLÉS :** Allocation aléatoire - Essai clinique - Médecine factuelle - Médecine personnalisée - Règles de bonne pratique

### INTRODUCTION

La médecine factuelle a conquis ses lettres de noblesse depuis le début des années 90, moment où elle a pris son essor (1). Ainsi, depuis lors, les médicaments reçoivent leur autorisation de mise sur le marché en fonction des preuves d'efficacité et de sécurité apportées dans des essais cliniques de phases II-III; les recommandations internationales se fondent sur les mêmes niveaux de preuve pour proposer des algorithmes de prise en charge des patients de manière optimisée; enfin, les médecins praticiens, eux-mêmes, devraient être ou sont imprégnés, parfois à leur insu, de ces concepts pour choisir ce qu'ils considèrent être le meilleur traitement pour leur patient.

Les essais cliniques représentent le fondement même de la médecine factuelle (2-4). Cependant les essais cliniques ont des forces et des faiblesses résumées dans le tableau I et qui seront

STRENGTHS AND WEAKNESSES OF RANDOMISED CLINICAL TRIALS :  
EVOLVING CHANGES ACCORDING TO PERSONALIZED MEDICINE

**SUMMARY :** Randomised Controlled Trials (RCTs) represent the cornerstone of Evidence-Based Medicine (EBM). Based upon the rules of Good Clinical Practice (GCP), they offer many strengths but also present some weaknesses. The rigorous methodology used allows avoid bias related to confounding factors (through a control group), selection bias (through randomisation) and interpretation bias (through double blinding). However, patients recruited in clinical trials and study experimental conditions markedly differ from the situation in real life. Furthermore, clinical trials recruit a mix of good and poor responders, so that the average therapeutic response is most often mitigated. Clinical trials must evolve according to the new concepts of personalized medicine to become even more performing. In a near future, they must progress from a statistical analysis on large cohorts of patients to a more individualized analysis guided by patient phenotype and genotype characteristics.

**KEYWORDS :** Clinical trial - Evidence-based medicine - Good Clinical Practice - Personalized medicine - Randomisation

développées ci-dessous. Au vu de l'évolution de la médecine et du nouveau paradigme dans le développement des médicaments décrit dans un autre article (5), il sera sans doute exigé, de plus en plus, que les essais cliniques s'adaptent dans la perspective de la médecine personnalisée. De plus, ils devront utiliser davantage des «critères de jugement durs» plutôt que des critères intermédiaires imparfaits du point de vue de la signification clinique (6).

### FORCES DES ESSAIS CLINIQUES

Les règles de bonne pratique clinique (GCP) ont imposé des contraintes de qualité pour accroître, autant que faire se peut, la fiabilité des résultats des études cliniques et les conclusions qui peuvent en être tirées. Pour ce faire, des stratégies ont été développées pour éviter des biais. Un biais peut se définir comme une erreur systématique résultant d'une faute méthodologique. Les essais cliniques contrôlés, conduits selon les directives GCP, permettent d'éviter les principaux types de biais (2-4).

#### 1) EVITER LES PRINCIPAUX BIAIS

##### a) Biais liés à des facteurs confondants

Les facteurs confondants peuvent jouer un rôle majeur dans l'interprétation des résultats

(1) Médecin Investigateur (2) Médecin interniste, Spécialiste en Endocrino-Diabétologie, Consultant.

(3) Professeur Ordinaire, Université de Liège. Chef de Service, Service de Diabétologie, Nutrition et Maladies métaboliques et Unité de Pharmacologie clinique, CHU de Liège.

TABLEAU I. FORCES ET FAIBLESSES DES ESSAIS CLINIQUES

Forces	Faiblesses
Impact des facteurs confondants minimisé par la comparaison à un groupe contrôle	Patients sélectionnés différents de la vraie vie
Biais de sélection évités par la randomisation des patients	Conditions d'étude différentes de la vie réelle
Biais d'interprétation évités par la technique du double aveugle	Prisonnier d'un protocole strict (sauf si design adaptatif)
Critères de jugement prédéfinis et hiérarchisés <i>a priori</i>	Réponses individuelles hétérogènes résultant en un effet thérapeutique global moyen
Méthodes statistiques pré-définies avec calcul de l'effectif et de la puissance statistique	Impossibilité de prévoir <i>a priori</i> les bons et mauvais réponders
Abandons et perdus de vue pris en compte dans l'analyse en intention de traiter	Difficulté d'extrapoler d'une réponse de groupe à une réponse individuelle
Possibilité d'analyses de sensibilité et par sous-groupes (si possible définies <i>a priori</i> )	Effectif et suivi limités en durée empêchant la détection de manifestations indésirables rares

des études épidémiologiques observationnelles (par exemple, les travaux analysant l'influence d'un facteur de risque sur l'apparition d'une maladie). Ils peuvent aussi avoir une influence non négligeable dans les études interventionnelles. Dans les essais thérapeutiques, un facteur confondant (appelé aussi, facteur de confusion ou variable confondante) est une variable liée à la maladie, à la population étudiée ou à un quelconque autre événement en relation avec les conditions de l'étude, qui est susceptible d'induire un biais dans l'analyse du lien entre le médicament administré et la réponse observée. L'interférence d'un facteur confondant peut aboutir à conclure à de fausses associations, en l'occurrence, attribuer une amélioration clinique à la prise du médicament alors qu'elle résulte d'un autre phénomène intercurrent (par exemple, tout simplement l'histoire naturelle de la maladie).

Le biais lié à d'éventuels facteurs confondants peut être évité en utilisant un groupe contrôle (témoin) dont les patients présentent les mêmes caractéristiques démographiques, les mêmes facteurs de gravité liés à la maladie proprement dite ou encore des comorbidités et/ou comédications semblables à celles présentes chez les patients du groupe d'intervention. Ces sujets témoins doivent, par ailleurs, être recrutés et suivis de façon synchrone, exactement dans les mêmes conditions d'étude que le groupe d'intervention.

### b) Biais de sélection

Idéalement, les patients des deux groupes doivent avoir des caractéristiques identiques, notamment en termes de pronostic (chance de guérison ou d'amélioration clinique), de façon à ce que la différence observée dans le critère d'évaluation clinique au cours de l'étude puisse être attribuée au seul médicament testé. Il faut absolument éviter que l'investigateur influence le choix d'inclure le patient recruté soit dans le groupe d'intervention, soit dans le groupe contrôle. Cela pourrait, en effet, facilement conduire à des distorsions entre les deux groupes (par exemple, si l'investigateur change le patient de groupe lorsqu'il sait que ce dernier n'a pas bien répondu préalablement à un médicament de la même classe thérapeutique que celle utilisée dans le groupe contrôle recevant le médicament actif de référence). Pour éviter ce piège, la répartition doit se faire au hasard («random» en anglais). C'est le fameux processus de randomisation. De plus en plus souvent, la randomisation se fait de façon centralisée pour éviter toute interférence de l'investigateur ou de son équipe dans ce processus capital. Dans les grandes études multicentriques, le système centralisé le plus utilisé est le système IVRS («Interactive Voice Respond System»).

### c) Biais d'interprétation

Un autre biais majeur qu'il convient absolument d'éviter est le biais d'interprétation, de la part du médecin et/ou de la part du patient. En effet, l'un et l'autre peuvent être influencés dans l'interprétation des symptômes et signes observés s'ils sont informés du type de médicament administré. Le risque est d'autant plus important que les critères d'évaluation sont subjectifs et non objectifs. Pour éviter ce biais d'interprétation, la méthode la plus performante est celle faisant appel au double aveugle («double blind»). Ainsi, ni le médecin, ni le patient ne connaissent le médicament administré et l'interprétation de tous les événements enregistrés se fait donc en aveugle. Certes, cette situation n'est pas toujours très confortable ni pour le patient, ni pour le médecin, et cette procédure complique considérablement la préparation de l'étude. Cependant, c'est le prix à payer pour obtenir des résultats qui ne prêtent pas le flanc à la critique.

## 2) APPLIQUER UNE MÉTHODOLOGIE RIGOUREUSE

### a) Définir a priori les critères de jugement

Les critères de jugement («endpoints» en anglais) doivent être prédéfinis. S'ils sont plusieurs, ils doivent être hiérarchisés avec la définition, préalablement au début de l'essai, du critère de jugement primaire (principal), du ou des critères secondaires et du ou des critères tertiaires (éventuellement, exploratoires). Cette précaution évite que le choix des critères de jugement soit influencé par les résultats obtenus ou encore que le poids accordé à chaque critère soit modifié à la guise du promoteur de l'étude, une fois que les résultats sont disponibles. Pour augmenter le nombre d'événements pris en compte dans l'analyse statistique, on peut recourir à des critères composites, combinant plusieurs événements dans la même sphère d'intérêt. Par exemple, dans le domaine cardiovasculaire, un critère composite largement utilisé est le critère MACE («Major Adverse Cardiovascular Events») qui combine les décès d'origine cardiovasculaire, les infarctus non mortels et les accidents ischémiques cérébraux. Certaines études prennent en compte un critère MACE élargi, dans lequel sont également incluses les procédures de revascularisation et/ou les hospitalisations pour angor instable ou pour insuffisance cardiaque (7). Il est important que la médecine factuelle se base sur des critères de jugement forts plutôt que sur des critères de jugement intermédiaires («surrogate endpoints»), dont la pertinence clinique peut parfois être discutable (6). Pour qu'un critère intermédiaire puisse être utilisé comme critère de jugement, il doit avoir été validé, de façon aussi parfaite que possible, avec la démonstration d'une excellente corrélation entre le critère intermédiaire et la survenue d'événements cliniques (8).

### b) Calculer a priori la puissance statistique

Il est capital de calculer, lors de la conception même de l'essai, l'effectif de patients qu'il conviendra de recruter et d'analyser la puissance statistique de l'étude afin de pouvoir tirer des conclusions pertinentes, avec un degré de certitude suffisant. Cette précaution est valable pour tous les types d'études, de supériorité, de non-infériorité ou de bioéquivalence. Ce calcul se fait sur le critère de jugement principal. Il prend en compte l'amplitude de la différence que l'on veut mettre en évidence (considérée comme cliniquement significative), la variabilité intrinsèque du critère de jugement principal, le niveau de risque d'erreur alpha (risque de conclure qu'il y a une différence alors qu'il

n'y en a pas) et bêta (risque de conclure qu'il n'y a pas de différence alors qu'il y en a une) que l'on considère comme acceptable.

### c) Prendre en compte les arrêts d'étude et les perdus de vue

Dans tout essai clinique, *a fortiori* s'il est d'une certaine durée, il existe des patients qui interrompent le traitement ou qui sont perdus de vue. Les abandons peuvent s'expliquer par des motifs liés au traitement lui-même (manque d'efficacité, apparition de manifestations indésirables), au protocole d'étude (jugé trop astreignant) ou encore pour toute autre raison indépendante de l'essai. Analyser les résultats uniquement sur les patients qui ont terminé l'étude risquerait de fausser les conclusions concernant le médicament testé (par exemple, on pourrait ainsi ne tenir compte que des bons répondeurs qui ont, par ailleurs, bien toléré le produit, ce qui, à l'évidence, donnerait une vue trop optimiste de la réalité). Bien évidemment, les patients perdus de vue représentent un handicap majeur dans l'interprétation des études où le critère d'évaluation principal comprend les décès puisqu'on ne peut jamais exclure qu'ils soient décédés.

Pour éviter les pièges liés aux abandons, il est recommandé de privilégier une analyse des résultats dite «en intention de traiter («Intention-To-Treat» ou ITT en anglais) (9). Elle prend en compte tous les patients randomisés dans l'un ou l'autre bras, quelle qu'ait été leur adhésion au protocole d'étude, même s'ils n'ont pas terminé l'essai comme prévu. Cette analyse peut être complétée par une analyse «per protocol» ou une analyse limitée aux «completers». L'analyse «per protocol» ne prend en compte que les patients qui ont bien respecté le protocole prévu, en particulier la prise régulière du médicament à la posologie prescrite. L'analyse sur «completers» ne considère que les patients qui sont arrivés au bout de l'étude et exclut donc tous les abandons intercurrents, quelle qu'en ait été la raison. D'une façon générale, les résultats de l'analyse «per protocol» ou sur «completers» sont plus favorables que ceux de l'analyse en ITT, plus stricte (9).

### d) Prévoir des analyses de sensibilité ou par sous-groupes

Dans les grands essais cliniques sont régulièrement réalisées des analyses de sensibilité pour vérifier que la différence objectivée ne l'est pas uniquement par chance, mais correspond bien à une réalité clinique.

Des analyses par sous-groupes peuvent également être effectuées pour vérifier que le médicament est particulièrement efficace (ou, au contraire, s'avère inefficace) dans un sous-groupe de patients particuliers. Dans ce cas, il convient de bien faire la distinction entre des analyses par sous-groupes prévues *a priori* et dont les résultats ont alors force de preuve (mais, le revers est que cette stratégie requiert un plus grand nombre de patients à l'inclusion) ou des analyses par sous-groupes planifiées *a posteriori* et dont les résultats n'auront qu'une valeur indicative (à vérifier éventuellement ultérieurement dans un essai clinique spécifique).

## FAIBLESSES DES ESSAIS CLINIQUES

### 1) LIÉES À LA POPULATION SÉLECTIONNÉE

Les sujets recrutés pour des études cliniques répondent à des critères d'inclusion et d'exclusion relativement stricts. Les résultats obtenus dans ces essais ne peuvent être validés que pour une population de patients répondant aux mêmes caractéristiques et ne peuvent pas nécessairement être extrapolés aux patients non sélectionnés de la vraie vie.

Par ailleurs, les essais cliniques réalisés pour conforter la médecine factuelle (EBM) fondent leur analyse statistique sur des groupes de patients de telle sorte qu'il est difficile de prédire la réponse individuelle d'un patient particulier. De plus, comme déjà discuté dans un autre article (5), les patients recrutés, même s'ils sont judicieusement sélectionnés sur des critères d'inclusion et d'exclusion stricts, représentent tout de même un groupe hétérogène où chaque patient est particulier. Dès lors, on voit aisément la difficulté qui peut survenir lorsqu'on veut transférer les conclusions d'un essai répondant aux critères de la médecine factuelle à un patient individuel, avec ses caractéristiques phénotypiques et génotypiques propres, dans ce qu'il est convenu d'appeler la médecine personnalisée.

### 2) LIÉES AUX CONDITIONS D'ÉTUDE

Un protocole d'étude clinique est très strict et s'éloigne sensiblement des conditions de la vraie vie. Tout d'abord, les patients sont suivis par des médecins investigateurs, relativement spécialisés, et des collaborateurs formés aux essais cliniques. Ensuite, le suivi du patient est généralement beaucoup plus intensif que dans la vie réelle, ne fût-ce que par le nombre de visites dans le décours de l'étude. A cela, il faut ajouter des contrôles téléphoniques, cliniques et biologiques réguliers. Les mesures hygiéno-dié-

tétiques sont habituellement renforcées. Enfin, l'observance thérapeutique est stimulée et strictement vérifiée (10). Il en résulte ce qu'il est convenu d'appeler un «effet étude», qui explique que les patients sous placebo, dans un essai clinique, voient leurs paramètres sensiblement améliorés. Le revers, bien évidemment, est qu'il n'est pas toujours aisé de prédire ce que deviendra la réponse au traitement dans des conditions de vie réelle. Si les études de phase III ont été bien conçues, on peut raisonnablement penser que les effets observés dans les conditions d'étude seront reproduits dans les études d'observation de la vraie vie. Une analyse récente dans le domaine de la diabétologie rapporte des résultats encourageants à cet égard (11).

### 3) LIÉES AUX LIMITATIONS DE L'EFFECTIF ET DE LA DURÉE D'OBSERVATION

Les essais cliniques, même s'ils sont devenus de plus en plus volumineux en termes de nombre de patients enrôlés et de durée d'observation, seront toujours limités par rapport aux conditions de la vie réelle où le médicament sera prescrit à un nombre beaucoup plus important de malades qui prendront le médicament de façon nettement plus prolongée (en tout cas, pour ce qui concerne le traitement des maladies chroniques). Ainsi, des manifestations indésirables très rares pourront ne pas être détectées dans les essais cliniques et devront être repérées après la commercialisation du médicament par le système de la pharmacovigilance. Si l'on pouvait mieux comprendre quels sont les sujets à risque de développer pareilles manifestations indésirables, ils pourraient faire l'objet d'études cliniques spécifiques sous haute surveillance qui permettraient de préciser l'ampleur du danger et, ultérieurement, de ne pas retirer du marché un médicament potentiellement utile à une majorité de patients dépourvus du risque de présenter semblables effets indésirables (5).

## ÉVOLUTION EN FONCTION DE LA MÉDECINE PERSONNALISÉE

L'industrie pharmaceutique doit faire face à divers défis (5). Le premier est sans doute d'assurer la pérennité de ses bénéfices, pour qu'ils puissent être réinjectés dans la recherche et le développement de nouveaux médicaments. En effet, comme discuté par ailleurs (5), les essais cliniques coûtent de plus en plus cher et les chances de mise sur le marché et d'obtention du remboursement s'amenuisent. De nouvelles stratégies sont envisagées pour améliorer l'efficacité et réduire les coûts des essais cliniques.

Une des plus avancées est celle qui a recours à des protocoles d'études cliniques au design adaptatif (12, 13). En fonction des données préliminaires récoltées dans l'essai clinique en cours (lors d'une analyse intermédiaire pré-spécifiée), les instances réglementaires donnent l'opportunité de modifier certains points du protocole, par exemple en termes de critères de recrutement de patients. Le but de ce type d'approche est généralement de pouvoir réduire le nombre de patients à inclure ou la durée d'exposition au médicament. Cependant, il faut être prudent car cette modification en cours d'étude conduit à un essai clinique moins pur, dont l'interprétation des résultats peut s'avérer plus délicate (13).

Une autre avancée majeure peut être attendue de l'avènement de la médecine personnalisée ou de la médecine de précision (14). En effet, les résultats des essais cliniques sont pénalisés par le fait que sont enrôlés des patients dont le profil de réponse au traitement testé n'est pas ou est mal connu. De ce fait, les protocoles d'essais cliniques, malgré les critères d'inclusion et d'exclusion prédéfinis utilisés, englobent des patients bons répondeurs et mauvais répondeurs. Il en résulte que le résultat final n'est pas spectaculaire, ce qui rend critique la procédure conduisant à l'autorisation de mise sur le marché et, davantage encore, celle sensée aboutir à un éventuel remboursement à un prix acceptable (5). Le développement de la médecine personnalisée devrait permettre de mieux prédire la réponse au traitement grâce à des biomarqueurs définissant le phénotype et, davantage encore, le génotype du patient (15). La possibilité de sélectionner uniquement un sous-groupe de bons répondeurs potentiels dans un essai clinique permettrait, non seulement de réduire considérablement le nombre de patients à recruter, mais aussi d'amplifier la réponse thérapeutique observée. Le développement de tests de diagnostic «compagnons» est une avancée importante dans ce sens, même s'il reste encore d'importants progrès à faire (5).

## CONCLUSION

La médecine factuelle repose sur les résultats d'essais cliniques contrôlés, randomisés, menés en double aveugle. Ils se fondent sur une analyse statistique rigoureuse réalisée sur de grands groupes de patients, forcément hétérogènes quant à la réponse thérapeutique ou à la survenue de manifestations indésirables. Les progrès de la médecine personnalisée, en particulier les possibilités de détecter les patients

bons répondeurs grâce à des tests de diagnostic «compagnons», amèneront à revisiter les protocoles d'essais cliniques de façon à les rendre plus performants à l'avenir et à pouvoir plus aisément extrapoler les résultats à un patient individuel.

## BIBLIOGRAPHIE

1. Pasleau F.— Fondements de la médecine factuelle. *Rev Med Liège*, 2015, **70**, 225-231.
2. Scheen AJ.— Evidence-based medicine. Apport des essais cliniques contrôlés. *Rev Med Liège*, 2000, **55**, 216-219.
3. Umscheid CA, Margolis DJ, Grossman CE.— Key concepts of clinical trials: a narrative review. *Postgrad Med*, 2011, **123**, 194-204.
4. Jaillon P.— L'essai clinique contrôlé randomisé. *Bull Acad Natl Med*, 2007, **191**, 739-756; discussion 756-738.
5. Scheen AJ.— L'industrie pharmaceutique face à la médecine personnalisée : changement de paradigme dans le développement des nouveaux médicaments. *Rev Med Liège*, 2015, **70**, 237-241.
6. Naudet F, Falissard B, Boussageon R, et al.— Has evidence-based medicine left quackery behind? *Intern Emerg Med*, 2015, Apr 1. [Epub ahead of print].
7. Scheen AJ.— Comment j'explore ... Les critères de jugement dans les essais cliniques : réflexions à propos d'études récentes de prévention cardio-vasculaire. *Rev Med Liège*, 2006, **61**, 260-266.
8. Lassere MN.— The Biomarker-Surrogacy Evaluation Schema: a review of the biomarker-surrogate literature and a proposal for a criterion-based, quantitative, multi-dimensional hierarchical levels of evidence schema for evaluating the status of biomarkers as surrogate endpoints. *Stat Methods Med Res*, 2008, **17**, 303-340.
9. Shrier I, Steele RJ, Verhagen E, et al.— Beyond intention to treat : what is the right question? *Clin Trials*, 2014, **11**, 28-37.
10. Jandrain BJ, Ernest P, Radermecker RP, et al.— Stratégies pour éviter l'inertie et la non-observance dans les essais cliniques. *Rev Med Liège*, 2010, **65**, 246-249.
11. Flory JH, Mushlin AI.— Observational and clinical trial findings on the comparative effectiveness of diabetes drugs showed agreement. *J Clin Epidemiol*, 2015, **68**, 200-210.
12. Bowalekar S.— Adaptive designs in clinical trials. *Perspect Clin Res*, 2011, **2**, 23-27.
13. Chow SC.— Adaptive clinical trial design. *Ann Rev Med*, 2014, **65**, 405-415.
14. Marquet P, Longerey PH, Barlesi F, et al.— Translational research : precision medicine, personalized medicine, targeted therapies: marketing or science? *Thérapie*, 2015, **70**, 11-19.
15. Cavalier E.— Approche globale et personnalisée des biomarqueurs. *Rev Med Liège*, 2015, **70**, 257-261.

Les demandes de tirés à part sont à adresser au Pr A.J. Scheen, Département de Médecine, CHU de Liège, Belgique.