

# Bénéfice médical réel estimé par comparaisons indirectes

Bruno Falissard<sup>1,2,3</sup>, Myriam Zylberman<sup>4</sup>, Michel Cucherat<sup>5</sup>, Valérie Izard<sup>6</sup>, François Meyer<sup>6</sup> et les participants à la table ronde n°5 Giens XXIV\*

1 INSERM, U669, Paris, France

2 Université Paris-Sud et Université Paris Descartes, UMR-S0669, Paris, France

3 AP-HP, Hôpital Paul Brousse, Département de Santé Publique, Villejuif, France

4 LILLY France, Suresnes, France

5 Faculté de Médecine Laënnec, EA 3736, Lyon, France

6 Haute Autorité de Santé (HAS), Saint-Denis La Plaine, France

Texte reçu le 15 novembre 2008 ; accepté le 20 novembre 2008

## Mots clés :

comparaisons indirectes ;  
comparaisons mixtes ;  
efficacité potentielle ;  
efficacité réelle

**Résumé** – Les données cliniques soumises au Comité des Médicaments à Usage Humain (CHMP) pour autorisation de mise sur le marché manquent fréquemment de comparaisons directes pertinentes des produits. Ces comparaisons peuvent permettre aux autorités nationales chargées de l'approbation du remboursement d'apprécier l'Amélioration du Service Médical Rendu (ASMR) d'une substance active nouvelle ou d'une entité pharmaceutique existante en cas d'extension d'indication. Les comparaisons indirectes et comparaisons mixtes sont des procédés issus du domaine de la méta-analyse et conçus pour aborder ce problème. Les comparaisons indirectes ajustées, les méta-régressions, les modèles mixtes, les analyses en réseaux bayésiens regroupent les résultats d'essais contrôlés randomisés permettant une synthèse quantitative des données. La procédure REAL (Relative Effectiveness Assessment of Listed drugs), récemment développée en France par la Haute Autorité de Santé (HAS), associe une comparaison mixte et un modèle d'effet basé sur des avis d'experts. Elle est destinée à traduire l'efficacité observée au cours des essais en une efficacité potentielle attendue en France dans la pratique clinique quotidienne.

## 1. Introduction

En France, l'efficacité potentielle de nouvelles entités pharmaceutiques en situation réelle de prescription est appréciée par la Commission de la Transparence de la Haute Autorité de Santé (HAS) avec l'aide d'un groupe important d'experts. Les 20 membres votants nommés notent la valeur thérapeutique ajoutée (Amélioration du Service Médical Rendu - ASMR) du médicament de 1 (amélioration majeure) à 5 (aucune amélioration). Cette notation est immédiatement utilisée par le Ministère de la Santé pour aider à déterminer le prix et le taux de remboursement de ce médicament en France.

En pratique, les sociétés pharmaceutiques doivent soumettre à la Commission de la Transparence un rapport standard de toutes

leurs données d'efficacité et de sécurité sur le produit. Dans la plupart des cas, ces données sont obtenues à partir d'essais contrôlés randomisés. Malheureusement, dans bon nombre de situations, il existe un manque d'essais comparatifs directs entre le produit à évaluer et les médicaments concurrents les plus importants.<sup>[1]</sup>

Pour résoudre ce problème, il a été souvent suggéré de combiner les données d'une série d'essais pour obtenir une *comparaison indirecte* ou une *comparaison mixte*.<sup>[2,3]</sup> Par exemple, les résultats des essais contrôlés randomisés comparant A à un placebo (P) et ceux comparant B à un placebo peuvent être regroupés pour obtenir, indirectement, une comparaison de A avec B.

L'objectif du présent article est de déterminer si de telles méthodologies peuvent être utilisées en routine pour évaluer l'efficacité potentielle réelle de nouveaux médicaments sur la base des données d'autorisation de mise sur le marché.

\* Pour la liste des participants, voir en fin d'article.

Dans la première partie, nous allons détailler les procédés statistiques actuellement disponibles. Les données sont le plus souvent obtenues à partir d'essais contrôlés randomisés de phase III et l'évaluation va se focaliser sur l'efficacité observée au cours de ces essais (« le médicament agit-il dans des conditions optimales ? ») plutôt que sur l'efficacité potentielle réelle (« le médicament agit-il lors des soins de routine ? »).<sup>[4]</sup>

Dans la deuxième partie, nous allons aborder la question de la comparaison indirecte de l'efficacité réelle qui est centrale dans le contexte de l'évaluation de nouveaux produits.

## 2. Comparaisons indirectes et comparaisons mixtes

En l'absence de comparaisons directes, il peut être tentant de comparer qualitativement les différentes quantités d'effet des médicaments en lice et de regarder si les intervalles de confiance des effets de différents traitements se superposent ou non, ce qui n'est vraiment pas recommandé.<sup>[5]</sup>

L'approche des comparaisons indirectes « non ajustées » consiste à collecter tous les essais contrôlés randomisés impliquant le traitement A et tous les essais contrôlés randomisés impliquant le traitement B, et à comparer l'amélioration moyenne des patients traités par A à l'amélioration moyenne des patients traités par B. Cette approche présente l'inconvénient majeur de nier la nature randomisée de chaque essai contrôlé individuel et doit aussi être écartée.<sup>[5]</sup>

Les comparaisons indirectes « ajustées » prennent en compte la nature randomisée de chaque essai. Elles ne considèrent que les essais contrôlés randomisés comparant par exemple, A et P et les essais contrôlés randomisés comparant B et P. On calcule alors la moyenne des différences d'amélioration des patients traités par A et des patients traités par P, moyenne que l'on compare à celle des différences d'amélioration des patients traités par B et des patients traités par P. La méthodologie est donc proche de celle utilisée dans les méta-analyses, de sorte que, comme c'est le cas dans ce dernier domaine, la question de l'hétérogénéité devient cruciale. En fait, les différents essais contrôlés randomisés inclus dans l'analyse sont réalisés à divers moments, au sein de populations distinctes, avec des dosages parfois différents. Cette hétérogénéité peut conduire à de possibles biais. Par exemple, si tous les essais A-Placebo sont effectués sur des patients très atteints (de sorte que la différence entre A et le placebo pourra s'avérer élevée) et si tous les essais B-Placebo sont réalisés sur des patients modérément atteints (de sorte que la différence entre B et le placebo sera plus faible), les comparaisons A-B de la comparaison indirecte peuvent être inappropriées.

Le concept de méta-régression<sup>[5]</sup> a justement été développé pour traiter cet inconvénient : les facteurs confondants (la sévé-

rité de l'atteinte des patients dans l'exemple ci-dessus) peuvent être ajoutés au modèle, de sorte que certaines sources possibles de biais peuvent être prises en compte.

De la même manière, des modèles mixtes ont été proposés. Dans ces cas, une variable « étude » est considérée comme un effet randomisé qui peut aider à la généralisation des résultats.

Enfin, une approche plus sophistiquée, dite méta-analyse en réseaux, a été développée récemment.<sup>[2]</sup> Basée sur les modèles bayésiens, elle peut traiter une large variété de situations : plusieurs produits peuvent être comparés indirectement sans qu'un seul comparateur commun doive être disponible.

Une question importante au sujet des comparaisons indirectes et des comparaisons mixtes concerne leur validité. Song *et al.*<sup>[6]</sup> ont comparé les résultats obtenus par comparaisons directes avec les résultats obtenus par comparaisons indirectes ajustées. Parmi les quarante-quatre situations envisagées (comparaison de deux traitements pour une pathologie donnée), il n'y a eu incohérence entre les deux approches que dans trois cas seulement.

## 3. Généralisation des résultats de comparaisons mixtes aux pratiques réelles

L'appréciation par la Commission de la Transparence du bénéfice médical réel ne peut pas uniquement reposer sur des résultats d'essais contrôlés randomisés de phase III, conçus pour obtenir une approbation de la Food and Drug Administration (FDA) ou de l'European Medicines Agency (EMA). En fait, ces essais contrôlés randomisés sont généralement effectués dans des pays pourvus de systèmes de santé spécifiques ; les patients randomisés doivent vérifier des critères spécifiques d'inclusion ; en outre, le suivi et les médecins impliqués dans ces essais ne sont peut-être pas représentatifs de la pratique quotidienne.

Idéalement, les essais contrôlés randomisés réalisés en vie réelle ou les études d'observation devraient être effectués dans chacun des pays où le nouveau médicament est disponible. Actuellement, ces études sont l'exception. Pour traiter ce problème, la HAS a proposé récemment une méthodologie spécifique d'évaluation de l'efficacité réelle relative des produits pharmaceutiques dite méthode REAL (Relative Effectiveness Assessment of Listed drugs).<sup>[7]</sup>

La méthode REAL est une procédure qui se déroule en trois étapes :

1. Première étape : la collecte de données issues d'essais contrôlés randomisés. Une liste d'essais est généralement soumise par l'industriel dans un rapport standard. D'autres essais obtenus à partir d'une revue systématique de la littérature ou d'avis d'experts peuvent être ajoutés à cette liste par la Commission ;

2. Deuxième étape : la collecte de données sur le caractère approprié d'une généralisation et d'une transposition des résultats d'essais contrôlés randomisés sélectionnés à l'étape 1. Cette étape est basée sur un modèle d'effet élémentaire.<sup>[8]</sup> Plus précisément, on postule qu'il existe une relation linéaire entre l'efficacité relative de A par rapport à B, telle qu'établie par un essai contrôlé randomisé [cette efficacité relative est exprimée dans ce cas, par un log du rapport des cotes, noté « log(ORRCT) »] et l'efficacité réelle relative attendue de A par rapport à B estimée idéalement dans une étude réelle [à savoir, log(ORRL)]. On pose :

$$E[\log(\text{ORRL})] = a + E[\log(\text{ORRCT})] \quad (1)$$

et

$$\text{Var}[\log(\text{ORRL})] = b \times \text{Var}[\log(\text{ORRCT})] \quad (b \geq 1) \quad (2)$$

Si  $a < 0$ , alors l'efficacité réelle sera inférieure à l'efficacité observée dans l'essai. Si  $a > 0$ , l'inverse est vrai. La constante  $b$  représente un manque de confiance dans l'estimation de log(ORRL) par comparaison à la confiance disponible pour l'estimation de log(ORRCT).

Les constantes  $a$  et  $b$  seront obtenues à partir des réponses des 20 experts de la Commission de la Transparence à quatre questions.

Ces réponses seront graduées en utilisant des échelles visuelles analogiques :

Question 1 : Cochez le segment ci-dessous en fonction de votre manque de confiance dans l'efficacité relative du produit (Re) dû aux faiblesses méthodologiques potentielles de l'essai ;

Question 2 : Cochez le segment ci-dessous en fonction de votre manque de confiance dans l'efficacité réelle relative du produit (Rr) en conditions réelles, dû à des problèmes de transposabilité des résultats des essais à la vie réelle ;

Question 3 : Cochez le segment ci-dessous en fonction de votre perception de la relation entre l'efficacité relative du produit (Re) observée dans l'essai et l'efficacité réelle relative (Rr) observée en conditions réelles après une administration courte (quelques mois, quelques années). Basez votre réponse sur votre expérience clinique et votre connaissance des essais cliniques ;

Question 4 : Au vu de votre expérience clinique, pensez-vous que l'efficacité potentielle relative de A et de B peut significativement changer à long terme : oui/non.

Si oui, cochez le segment ci-dessous en fonction du changement que vous envisageriez au niveau de l'efficacité relative Rr des deux produits à long terme (espérance de vie du patient) dans la vie réelle.

Si  $b_1$  et  $b_2$  représentent les réponses aux questions 1 et 2, et que  $a_1$  représente la réponse à la question 3, et si la réponse à la question 4 est « non », alors  $a = \log(a_1)$  et  $b = 1/(b_1 \times b_2)$ , et

l'efficacité réelle relative log(ORRL) et sa confiance peuvent être calculées en utilisant les équations (1) et (2). Si la réponse à la question 4 est « oui » et qu'elle est représentée par  $a_2$ , alors  $a = \log(a_1 \times a_2)$  et  $b = 1/(b_1 \times b_2)$ .

3. Troisième étape : regroupement de log(ORRL), analyse de sensibilité. Le regroupement est effectué en utilisant une échelle monodimensionnelle,<sup>[9]</sup> l'analyse de sensibilité en utilisant un bootstrap.

*Exemple* : un médicament antidépresseur A est proposé pour le traitement du trouble anxieux généralisé. Quatre essais sont disponibles. Le premier compare A (75 mg par jour) à un placebo chez 160 + 168 patients. Le deuxième essai compare également A (75 mg par jour) au placebo chez 160 + 168 patients. Le troisième essai compare A (75 mg par jour), A (150 mg par jour) et B (40 mg par jour) au placebo chez 157 + 158 + 160 + 154 patients. Le quatrième essai compare également A (75 mg par jour), A (150 mg par jour) et B (40 mg par jour) au placebo chez 156 + 160 + 151 + 157 patients. Comme présenté ci-dessus, le modèle d'effet est obtenu à partir des cotations des experts, les moyennes de  $a_1$ ,  $a_2$ ,  $b_1$  et  $b_2$  pour tous les experts et tous les bras des essais sont respectivement de 1,13, 0,94, 0,88 et 0,69. Il convient de noter que  $a_1 \times a_2$  est supérieur à 1. En d'autres termes, il est suggéré par les experts que les efficacités réelles relatives attendues de A et de B sont, en moyenne, légèrement supérieures aux efficacités relatives observées dans les essais. Une interprétation possible qui pourrait expliquer ce résultat apparemment inattendu, est qu'il est probable que les patients du bras placebo d'un essai contrôlé randomisé de phase III dans le trouble anxieux généralisé, reçoivent *de facto*, un traitement non pharmacologique efficace. Un tel suivi minutieux est cependant rare dans la pratique quotidienne (au moins en France) de sorte qu'il est probable que l'efficacité réelle relative (en conditions réelles) du traitement A ou B par rapport à l'absence de médication soit supérieure à l'efficacité relative de A ou B par rapport au placebo.

Les résultats de la procédure REAL appliquée à ces essais sont présentés figure 1. En termes d'efficacité réelle, le traitement A semble être supérieur au placebo, avec un effet qui est comparable au traitement B. Les efficacités relatives observées dans les essais sont également représentées (pour les estimations obtenues lorsque  $a = 0$  et  $b = 1$ ) afin de visualiser l'influence des cotations des experts sur les distributions. Il existe une augmentation attendue de la variation de l'efficacité réelle relative par comparaison à l'efficacité relative observée dans les essais ; cette augmentation de la variation est due à l'incertitude de la généralisation et de la transposition des résultats des essais contrôlés randomisés aux conditions réelles. Il existe également un léger déplacement entre l'efficacité réelle relative et l'efficacité relative observée dans les essais ; ceci est cohérent avec les remarques exposées ci-dessus.

## 4. Conclusion

Les comparaisons indirectes et les comparaisons mixtes exigent une méthodologie rigoureuse passant par une revue minutieuse de la littérature et l'utilisation de modèles statistiques appropriés. Des analyses de sensibilité doivent être effectuées pour évaluer la robustesse des résultats. Il y a de nombreuses limites. L'hétérogénéité des essais contrôlés randomisés incorporés dans l'analyse est l'une des principales. Le nombre minimal d'essais disponibles requis pour utiliser des procédés de comparaison sophistiqués est une autre limite et un minimum de quatre essais par comparaison de traitements a été proposé.<sup>[9]</sup>

En pratique, on suggère que les comparaisons indirectes et les comparaisons mixtes soient effectuées quand aucune comparaison directe n'est disponible, ou bien lorsque les comparaisons directes concernent un comparateur non optimal, ou encore lorsque les essais contrôlés randomisés présentent des faiblesses méthodologiques.

Concernant la transposition et la généralisation d'une synthèse quantitative des résultats des essais contrôlés randomisés, la HAS développe actuellement la procédure REAL en France. Cette procédure combine des comparaisons mixtes avec un modèle d'effet. Ce modèle d'effet est obtenu à partir d'avis d'experts. Il est destiné à traduire l'efficacité observée dans les essais en efficacité réelle attendue en pratique clinique quotidienne. Les distributions des efficacités réelles obtenues de cette manière sont alors utilisées par la Commission de la Transparence pour une évaluation plus fiable du bénéfice médical réel. La procédure REAL est au stade expérimental et devrait être utilisée en routine dans un futur proche. Elle est encore sujette à améliorations, par exemple le recours à des risques absolus à la place de « odds-ratios » ; de plus, dans la mesure où REAL est basée en partie sur des avis subjectifs d'experts, elle devra être évaluée après une phase d'essai.

*Remerciements.* Les auteurs remercient tous les membres de la « table ronde 5 : valeur ajoutée relative basée sur les comparaisons indirectes » pour leurs discussions stimulantes et productives.

## Participants

Alexandre Barna (HAS), Michel Cucherat (Hôpital Laënnec, Lyon), Frédéric De Bels (HAS), Rima de Sahb (Laboratoires

Roche), Jean Delonca (Laboratoire Bristol Myers Squibb), Valérie Izard (HAS), Pascale Jolliet (CHU Nantes), Jean-Michel Joubert (UCB Pharma), Robert Launois (Réseau Évaluation en Économie de la Santé), Claire Le Jeunne (Service de Médecine Interne, Hôtel Dieu, Paris), Michel Lièvre (Faculté de Médecine, Lyon), Florence Mahlberg Gaudin (Laboratoires Servier), Aurélie Moser (Laboratoire Roche), Patricia Perles (Laboratoire Sanofi Aventis), Martine Pigeon (CNAMTS), Cécile Rey Coquais (Laboratoire Wyeth Pharmaceutical), Nadine Weisslinger (Laboratoire GlaxoSmithKline), Bertrand Xerri (HAS).

## Références

1. Ioannidis JP. Indirect comparisons: the mesh and mess of clinical trials. *Lancet* 2006; 368: 1470-2
2. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; 20: 3105-24
3. Cucherat M. Les comparaisons indirectes : méthodes et validité. 2009, Haute Autorité de Santé - Service évaluation des médicaments: Paris. p. 66
4. Bombardier C, Maetzel A. Pharmacoeconomic evaluation of new treatments: efficacy *versus* effectiveness studies? *Ann Rheum Dis* 1999; 58 Suppl 1: 182-5
5. Gartlehner G, Moore CG. Direct *versus* indirect comparisons: a summary of the evidence. *International Journal of Technology Assessment in Health Care* 2008; 24(2): 170-7
6. Song F, Altman DG, Glenny AM, *et al.* Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ* 2003; 326 (7387): 472 [Letter]
7. Falissard B, Izard V, Xerri B, *et al.* Relative Effectiveness Assessment of Listed Drugs (REAL): a new method for an early comparison of the effectiveness of approved health technologies. *International Journal of Technology Assessment in Health Care*, 2009. in press
8. Boissel JP, Collet JP, Lievre M, *et al.* An effect model for the assessment of drug benefit: example of antiarrhythmic drugs in postmyocardial infarction patients. *J Cardiovasc Pharmacol* 1993; 22(3): 356-63
9. Glenny AM, Altman DG, Song F, *et al.* International stroke trial collaborative group: indirect comparisons of competing interventions. *Health Technol Assess* 2005; 9(26): 1-134

Correspondance et offprints : Bruno Falissard, INSERM U669, Maison de Solenn, 97 boulevard de Port Royal, 75679 Paris Cedex 14, France.  
E-mail : falissard\_b@wanadoo.fr