



Synthèse du Congrès de l'ISPOR 2010

Mars 2011

COLLÈGE DES ÉCONOMISTES DE LA SANTÉ

NOVARTIS FRANCE

ÉDITORIAL

Face à l'intérêt suscité par les précédentes éditions, le Collège des Economistes de la Santé et les Laboratoires Novartis-France vous présentent, pour la troisième année consécutive une synthèse des principaux apports du dernier congrès européen de l'International Society for Pharmacoeconomics and Outcomes Research (ISPOR) qui s'est déroulé à Prague en Novembre 2010. Le rendez-vous annuel que constitue cette réunion a pris une place essentielle dans l'agenda des principaux acteurs intervenant dans le domaine de l'évaluation médico-économique des stratégies de santé (*Health Technology Assessment ou HTA*) et son succès ne s'est pas démenti avec plus de 2 900 participants.

Ce 13^{ème} congrès s'est inscrit dans un contexte associant en Europe, la sortie progressive d'une crise économique générale et la poursuite de la croissance rapide des dépenses de santé. L'évaluation des technologies de santé a été au cœur des sujets abordés durant les trois séances plénières d'un congrès à nouveau extrêmement riche en présentations, tables rondes et ateliers de travail. Cette thématique appelait naturellement un partage des expériences spécifiques des nombreux pays représentés d'une part et la recherche d'une harmonisation des approches actuellement menées dans les Nations les plus avancées tant sur un plan méthodologique que politique d'autre part. Cette évolution trouve par exemple une traduction concrète dans la mise en place du programme 2010-2012 de l'EUnetHTA, le réseau européen des Agences d'évaluation des technologies de santé, soutenu par la Commission Européenne et 34 pays, qui a été largement présenté et discuté.

Bien que les pays européens partagent généralement des principes communs dans le champ de la santé : universalité, accès à des soins de qualité, équité et solidarité, l'ambition de ce programme n'échappera pas aux acteurs du domaine compte tenu des différences notables existant entre les politiques actuellement mises en œuvre.

Comme chaque année, la synthèse du congrès a été rédigée par un groupe d'étudiants/jeunes diplômés : Louis Blech-Semin (Pharmacien, Université Paris Descartes et Master 2 Economie et Gestion de la santé, Université Paris Dauphine), Olivia Clément (MA in Economics, McGill University) et Maria Laura Silva (doctorante, GATE, UMR CNRS 5824, Lyon). L'intérêt de la collaboration entre le Collège des Economistes de la Santé et les Laboratoires Novartis se situe également dans cette mise en situation de jeunes professionnels appelés à travailler dans notre secteur de recherche et d'activités.

L'exercice n'est pas des plus simple compte tenu de la diversité des sujets traités. Nous espérons néanmoins qu'il répondra à nouveau à vos attentes.

Bruno DETOURNAY (CEMKA-EVAL)

AMELIORER L'EFFICIENCE DE L'EVALUATION DES TECHNOLOGIES DE SANTE EN EUROPE

Le titre un peu provocateur de la première conférence plénière était là pour rappeler quelques-uns des problèmes auxquels se heurte aujourd'hui l'évaluation des technologies de santé dans les pays européens. Parmi ces problèmes ont été évoqués, le caractère redondant des évaluations technologiques conduites successivement sur un même produit dans les différents pays voire régions, la variabilité des prérequis techniques ou réglementaires exigés pour les dépôts de dossiers nationaux ou encore l'absence de consensus méthodologique sur des questions aussi essentielles que la mesure des résultats cliniques.

Les orateurs ont également observés que les évaluations actuellement produites concernent essentiellement les produits pharmaceutiques et beaucoup moins d'autres technologies ou programmes et que les choix en termes de priorisation et d'affectation des ressources ne sont pas toujours clairs et cohérents. Lorsqu'elles sont menées les évaluations ne sont pas systématiquement prises en compte par les payeurs pour des raisons politiques ou plus simplement de calendrier décalé par rapport aux décisions (de remboursement par exemple). Enfin, les intervenants ont fait état de l'insuffisance des méthodes utilisées pour traiter des questions d'efficacité relative en vie réelle ou encore de ratios d'efficience.

Pour répondre à ces difficultés, l'action de la Commission européenne et des gouvernements d'une trentaine de pays (dont la plupart des pays membres de l'Union Européenne, mais aussi la Norvège, la Suisse et la Croatie) a abouti à la création d'un organisme, l'EUnetHTA, chargé de mettre en place une collaboration efficace et durable entre 34 agences nationales/régionales chargées de l'HTA. Les missions de cet organisme sont larges, qu'il s'agisse d'offrir un portail d'accès vers les différentes agences, de partage de connaissances et d'élaboration d'outils, d'assurance qualité, de soutien à la recherche ou de la mise en œuvre des évaluations. **A terme, est visée la création d'une véritable institution européenne chargée de l'évaluation des technologies de santé.**

On peut admettre que les stratégies de santé, les attentes des patients et les prix de plus en plus standardisés sont similaires entre les pays et qu'il existe de fait un marché unique à l'échelle européenne au moins pour le médicament et les dispositifs médicaux. Toutefois, l'action de l'EUnetHTA se situe, pour le moment, uniquement au niveau de l'information des payeurs nationaux en Europe qui restent seuls chargés de délivrer les autorisations. En ce sens, son rôle est finalement assez proche de celui de l'INAHTA (International Network of Agencies for Health Technology Assessment) qui regroupe des agences ou organismes de 26 pays dont les pays d'Amérique du Nord et Sud, l'Australie et quelques pays européens (dont la France) et d'Asie depuis quelques années.

La création de l'EUnetHTA n'est pas sans conséquences éventuelles sur le rôle des agences nationales d'évaluation, sur les orientations de l'EMA (European Medicines Agency) ou sur les relations entre les autorités de santé et les industriels. Ces derniers ont plaidé lors de la première session plénière pour que les travaux et décisions de l'EMA soient pleinement reconnus et acceptés et qu'aucune demande supplémentaire d'évaluation ne soit exigée à l'entrée sur le marché au niveau de chaque pays. Ces mêmes industriels ont souhaité que l'on continue à traiter séparément efficacité et intérêt thérapeutique du produit. Ils ont admis que cette dernière dimension pouvait comprendre la prise en compte de conséquences complémentaires dans une perspective sociétale ou dans la perspective d'agents particuliers dans les différents contextes nationaux mais uniquement si ces apports ne venaient pas contredire les résultats de l'évaluation clinique conduite au niveau européen.

1.1 Vers une harmonisation européenne de l'évaluation des technologies de santé

L'évaluation des technologies de santé est un processus multidisciplinaire, résumant toutes les informations disponibles concernant les apports médicaux, sociaux, économiques et éthiques relatifs à une stratégie de santé, de manière transparente, systématique, robuste et non biaisée. Le critère principal admis le plus souvent pour cette évaluation est la mesure de l'efficacité relative de la stratégie de santé étudiée. On cherche ainsi à déterminer la valeur ajoutée de l'innovation relativement aux approches diagnostiques ou thérapeutiques utilisées habituellement afin de prendre une décision quant au financement collectif du produit de santé.

Par ailleurs, cette valeur ajoutée « thérapeutique » est réévaluée régulièrement dans le cadre d'une utilisation du produit dans la pratique courante à travers des études observationnelles voire des essais pragmatiques.

Chaque agence en charge de l'HTA a sa propre vision de ce qui doit être pris en considération dans cette mesure et des méthodologies permettant de l'établir. Les différences sont notables en ce qui concerne par exemple la nature des résultats de santé qui sont considérés (refus ou acceptation des critères de substitution, des critères composites, des QALYs), l'intégration d'une dimension économique dans l'évaluation à travers des analyses coût-efficacité ou

coût-utilité, ou encore la prise en compte des « Patients Reported Outcomes » (PRO).

Les industriels sont alors confrontés au problème consistant à tenter de comprendre ce qui est attendu par chacune des 30 agences européennes d'HTA toutes indépendantes, avec des niveaux variables d'expertise et susceptibles de prendre des décisions contradictoires.

Une harmonisation est dès lors recherchée sur des thèmes tels que :

- l'adoption d'un vocabulaire commun ;
- la nature des informations attendues dans les dossiers soumis par les industriels ;
- les méthodologies de revue des données et des modèles produits lors de l'évaluation ;
- les perspectives à prendre en compte.

La première production de l'EUnetHTA a été un travail de définition de l'efficacité relative des traitements pharmaceutiques reprenant les recommandations du Forum Pharmaceutique de 2008 (organisme créé en 2005 par la Commission Européenne réunissant tous les 3 ans des responsables politiques, industriels ou d'associations de professionnels ou de patients).

Les principales définitions retenues sont les suivantes :

Efficacité (efficacy) : la mesure des effets bénéfiques d'une intervention par rapport à ses effets négatifs dans les circonstances idéales d'utilisation ;

Efficacité relative (relative efficacy): correspond à la même notion mais par comparaison à une ou plusieurs interventions alternatives ;

Efficacité en vie réelle/pratique courante (effectiveness) : la mesure des effets bénéfiques d'une intervention par rapport à ses effets négatifs dans les circonstances habituelles de la pratique médicale,

Efficacité relative en pratique courante (relative effectiveness) : même définition que précédemment mais dans le cadre d'une comparaison avec une ou plusieurs autres stratégies visant le même résultat.

S'ajoute à ces concepts, celui de la valeur ajoutée d'un produit qui repose sur la prise en compte des résultats d'efficacité relative en pratique courante mais également de l'ensemble de ses conséquences sociales, éthiques et économiques pour aboutir à une évaluation complète de la technologie de santé (*full HTA process*).

En complément de ces réflexions et à un niveau plus général une table ronde a été consacrée à la nécessaire clarification des problèmes de vocabulaire entre les notions d'EBM (Evidence Based Medicine), d'HTA et de CER (Cost-effectiveness Research) à partir d'un article publié récemment (Luce *et al.* 2010¹).

1.2 Les obstacles à l'harmonisation

Bien que souhaitée par tous, l'harmonisation européenne des évaluations des technologies de santé se heurte à un ensemble de difficultés. Si on peut admettre que l'efficacité d'un produit pharmaceutique présente un caractère universel,

- Des particularités épidémiologiques existent entre les pays européens ;
- Les pratiques médicales ne sont pas les mêmes dans tous les pays ni d'ailleurs la perception de la sévérité de certaines situations sanitaires ;
- Les modes de financement des soins sont différents ainsi que l'organisation des services de santé, ces deux champs relevant de politiques nationales.

En ce sens, la mesure de l'efficacité en vie réelle/pratique courante est en partie dépendante du contexte national ou régional dans laquelle elle est effectuée et il n'est pas nécessairement simple de conduire des études internationales dans ce domaine ou encore de transposer les résultats observés d'un pays à l'autre.

Les positions des intervenants du congrès dans ce domaine demeurent toutefois contrastées, certains estimant

que l'on exagère les différences entre les systèmes de santé et que, pour les produits ou dispositifs importants, les décisions sont finalement souvent cohérentes.

1.3 Quelle articulation avec les décisions de mise sur le marché ?

L'organisation européenne mise en place pour la délivrance des autorisations de mise sur le marché est évidemment beaucoup plus simple à appréhender pour les industriels que les évaluations destinées aux décisions de remboursement puisqu'un seul dossier doit être déposé pour l'ensemble des pays et que la décision obtenue s'applique dans tous ces derniers.

Les agences de régulation et en premier lieu l'agence européenne EMA fondent traditionnellement leur démarche sur une appréciation de la qualité, de la sécurité et de l'efficacité des produits pharmaceutiques par rapport à un traitement neutre (placebo) et le paradigme actuel veut qu'un produit soit autorisé quand les bénéfices dépassent les risques associés à son utilisation.

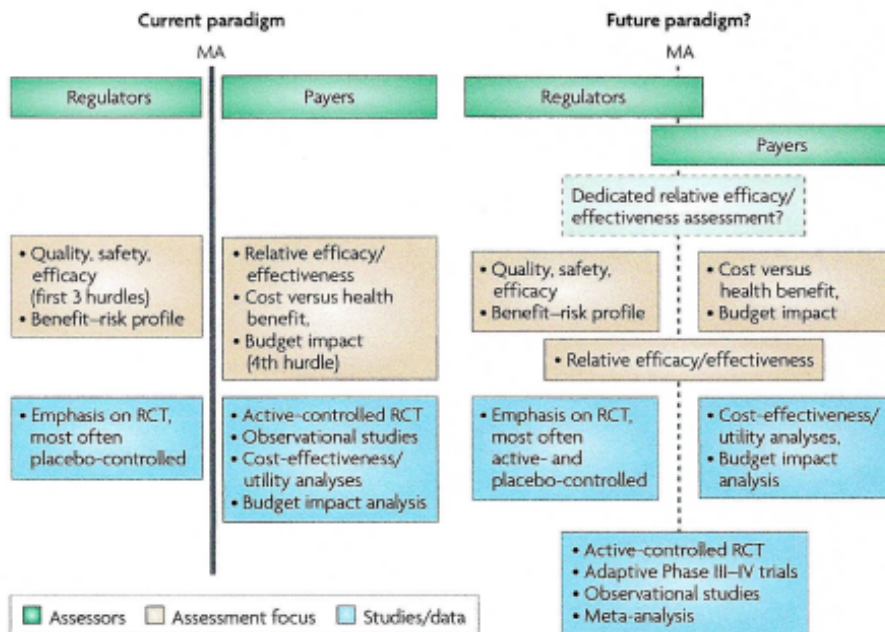
Ce paradigme est toutefois en train de se modifier et il tend à rapprocher les exigences des agences de régulation de celles des agences d'HTA. Ainsi, l'EMA souhaite de plus en plus disposer de résultats d'efficacité relative versus comparateurs actifs et non plus seulement versus placebo. L'EMA s'intéresse également plus souvent aux résultats des études observationnelles particulièrement lorsque ces dernières sont réalisées à partir de données collectées en pratique courante dans les systèmes publics ou privés de gestion électronique des dossiers patients. Par ailleurs, à travers le développement des autorisations de mise sur le marché conditionnelles et des plans de gestion des risques, l'EMA se préoccupe désormais des résultats obtenus par les produits après la commercialisation.

Les besoins des agences de régulation et des agences d'HTA (censées orienter les décisions des payeurs) ont donc tendance à se rapprocher et la plupart des industriels comprennent désormais que les évaluations de leurs produits devront satisfaire dans l'avenir simultanément les attentes des agences de régulation et celles des payeurs. Ce rapprochement a été illustré dans un article publié en 2010 et cité lors du congrès (²).

¹ Luce BR, Drummond M, Jönsson B, Neumann PJ, Schwartz JS, Siebert U, Sullivan SD. EBM, HTA, and CER: clearing the confusion. *Milbank Q.* 2010 Jun;88(2):256-7

² Eichler HG, Bloechl-Daum B, Abadie E, Barnett D, König F, Pearson S Relative efficacy of drugs: an emerging issue between regulatory agencies and third-party payers. *Nat Rev Drug Discov.* 2010 Apr;9(4):277-91.

Figure 1 Anticipation de l'évolution de l'évaluation de l'efficacité relative chez les payeurs et dans les agences de régulation en Europe (d'après 2 Eichler et al. 2010)



2. La MESURE DE L'EFFICACITE RELATIVE EN VIE REELLE

La mesure de l'efficacité relative entre deux ou plusieurs traitements soulève des problèmes particuliers surtout lorsque l'on souhaite étudier cette dernière dans le contexte de la pratique courante.

Lorsqu'il existe des alternatives, on s'aperçoit souvent que la plupart des interventions envisageables n'ont pas été directement comparées dans des essais cliniques randomisés (RCT) de taille suffisante et pour de bonnes raisons. Ainsi, durant la période 1971-2002, plus de 100 protocoles de chimiothérapie différents ont été étudiés dans le cancer du poumon (non à petites cellules). Cette abondance interdit de fait les comparaisons deux à deux de l'ensemble de ces schémas thérapeutiques. Par ailleurs, parmi les 254 essais réalisés dans ce domaine seuls deux d'entre eux ont inclus plus de 1 000 individus car ces essais sont difficiles à mettre en œuvre et extrêmement coûteux.

On ne dispose donc pas, la plupart du temps, de résultats d'études comparant les stratégies proposées et de résultats suffisamment valides (les études étant souvent sous-dimensionnées pour répondre à l'ensemble des questions d'intérêt qui se posent simultanément).

Lorsque les essais de comparaison directe n'apportent pas un niveau de preuve suffisant, il est possible, en théorie, d'utiliser les résultats d'essais complémentaires mettant en jeu les produits étudiés pour évaluer l'efficacité relative des différents traitements ce qui limite en partie les besoins

d'études. Il existe dès lors un fort intérêt actuel pour ces techniques tant du côté des industriels que des agences d'évaluation.

Bien que des avancées théoriques importantes aient été obtenues ces dernières années sur les méthodes à mettre en œuvre en ce sens, la pertinence de ces approches reste discutée. Il n'est pas certain que l'EMA par exemple acceptera des dossiers de demande d'AMM fondés sur des comparaisons indirectes. Il est encore moins certain que ces résultats seront pris en compte par tous les organismes d'HTA.

2.1 Les comparaisons indirectes

Les techniques de comparaisons indirectes (Indirect Treatment Comparison – ITC) offrent la possibilité de comparer ce qui n'a pas pu être comparé par des essais spécifiques: les différentes classes thérapeutiques, les médicaments différents dans la même classe, les différentes doses du même médicament en sont des exemples.

Relèvent de ces approches, des méthodologies assez diverses qui ne sont pas toutes reconnues comme valides. De nombreux travaux ayant un objectif de comparaison indirecte utilisent encore des méthodes inadaptées telles que la comparaison naïve des estimations ponctuelles ou de bras actifs issus d'essais contrôlés distincts. Ces approches sont biaisées car elles ne prennent pas en compte la non-

transitivité de la plupart des tests statistiques (le médicament A peut être supérieur au médicament B, lui-même supérieur au médicament C alors que C est supérieur à A). Par ailleurs, une telle comparaison ne peut avoir de sens que s'il existe une similarité (similarity) entre les études. La similarité doit porter sur les aspects cliniques (caractéristiques similaires des patients, des interventions, des contextes d'étude, de la durée de suivi, des critères de jugement, etc.) mais aussi sur les conditions de mise en œuvre de l'étude (procédures, monitoring, plan d'analyse, etc.).

Les méthodes acceptables de comparaison indirecte de traitement deux à deux incluent les comparaisons ajustées, et les approches mixtes associant comparaison directes et indirectes. Les comparaisons indirectes peuvent en effet être utiles même lorsque des essais de comparaisons directes sont disponibles mais que ces derniers sont insuffisants pour répondre aux questions posées.

Les évaluations simultanées de plusieurs traitements concurrents modélisent quant à elles le réseau des comparaisons possibles entre tous les traitements concurrents et elles constituent une extension des approches mixtes deux à deux.

Des comparaisons entre ces différentes approches ont été présentées tout au long du congrès.

2.2 Les méthodes de comparaison mixtes en réseau

La plupart des présentations du congrès portaient sur les comparaisons mixtes en réseau (Mixed Treatment Comparison - MTC) combinant comparaisons directes et indirectes.

L'idée générale du réseau est d'utiliser l'ensemble de l'information disponible à travers les essais et de présenter cette dernière sous la forme d'une structure spatiale où les sommets correspondent aux traitements (ou combinaisons de traitements) comparés et les liens entre ces sommets correspondent aux essais disponibles. La supériorité d'un traitement par rapport aux autres est établie de manière quantitative à partir de l'estimation des tailles d'effet.

Plusieurs méthodes statistiques pour l'estimation des paramètres de ces modèles sont utilisées :

- Des méthodes bayésiennes reposant sur des techniques d'intégration par Monte Carlo des chaînes de Markov. Cette méthode repose sur un processus d'estimation fondé sur l'ensemble des données disponibles et la génération successive de distributions de nombres aléatoires jusqu'à ce que la distribution de ces derniers converge vers la distribution réelle du paramètre à estimer. Ces méthodes nécessitent des traitements informatiques lourds et des logiciels adaptés.

- Des modèles linéaires mixtes (appelés aussi modèles à effets aléatoires) qui permettent de prendre en compte le fait que les effets d'un traitement peuvent varier de manière aléatoire et qui intègrent de ce fait deux niveaux de variabilité : la variabilité au niveau de chaque essai et la variabilité entre tous les essais.
- Des méta-régressions qui modélisent l'effet observé dans les essais à l'aide de covariables définies a priori (dose utilisées, caractéristiques des patients inclus, caractéristiques des essais, etc.).

Dans tous les cas, l'estimation est globale, utilisant les comparaisons directes disponibles mais aussi les comparaisons indirectes réalisables deux à deux.

Pour être valide, cette approche doit concerner un réseau d'études similaires mais, en plus, cohérentes (consistency) dans leurs résultats. Une incohérence des données peut être observée (Test I^2 de Higgins) si une différence existe entre les résultats des comparaisons directes et ceux des comparaisons indirectes.

2.3 Exemples d'application

Plusieurs illustrations de l'utilisation des méthodes de comparaison mixtes pour évaluer l'efficacité relative des traitements ont été présentées lors du congrès.

Adams et al. ont cherché à déterminer les conséquences sur la prise de décision que peut avoir le choix de la méthode utilisée pour réaliser une MTC. Quatre modélisations différentes ont été utilisées pour estimer l'efficacité relative d'un traitement et les résultats intégrés dans une analyse coût-efficacité. Les résultats d'efficacité relative obtenus variaient légèrement selon les modèles utilisés et donc, selon les résultats en termes de ratio coût-efficacité. Cependant, le positionnement relatif des estimations et le sens général des résultats obtenus n'étaient pas affectés, les auteurs en concluent qu'il est peu probable que le choix du modèle utilisé dans les MTC ait un impact fort sur la décision globale.

Dans le cadre des évaluations conduites par le NICE, une revue exhaustive de la littérature et une MTC ont été récemment réalisées pour le traitement de première ligne chez les femmes ménopausées atteintes de cancer du sein métastatique avec récepteurs hormonaux positifs et surexpression tumorale de HER2 (patients co-positifs). Cette étude avait pour objectif de déterminer s'il existe un effet classe entre les différents inhibiteurs de l'aromatase (IA) disponibles sur le marché et l'efficacité relative des différents IA dans le cas contraire en association avec deux traitements trastuzumab et lapatinib ciblant les récepteurs HER2.

Tous les essais étudiant l'effet des différents IA ou des associations trastuzumab + IA ou lapatinib + IA et les comparateurs éventuels ont été identifiés et analysés. La cohérence des articles trouvés dans la revue de la littérature a été

discutée avec des experts cliniques et également évaluée statistiquement. Une MTC a été conduite selon une méthode bayésienne en l'absence de comparaison directe entre les différents traitements ou combinaisons de traitements étudiés.

Les critères d'efficacité retenus étaient la survie sans progression et la survie totale. Aucune différence n'a été mise en évidence entre les 3 IA comparés. Dans toutes les comparaisons la combinaison trastuzumab + IA a toujours eu la plus grande probabilité de meilleure efficacité. Cependant très peu des essais utilisés dans l'analyse portaient sur la population cible co-positive visée et des analyses de sensibilité ont été planifiées. Bien que la probabilité d'être la meilleure combinaison thérapeutique variait dans les analyses de sensibilité, la combinaison trastuzumab + IA a continué d'être systématiquement associée avec la plus grande probabilité de meilleure efficacité.

Cette étude constitue un exemple d'application opérationnelle des méthodes de comparaison indirecte. Les résultats obtenus ont été ensuite intégrés dans une étude coût-efficacité et la prise en compte des coûts a permis de montrer que deux traitements s'inscrivaient sur la frontière d'efficacité : la combinaison trastuzumab/anastrozole et la monothérapie par anastrozole (cette dernière étant moins efficace que la combinaison précédente).

Une équipe hollandaise a conduit une MTC pour comparer l'efficacité de différents médicaments antifongiques utilisés dans les traitements prophylactiques chez les patients à haut risque. Les patients neutropéniques traités par chimiothérapie pour des cancers hématologiques et les bénéficiaires de transplantation de cellules souches hématopoïétiques allogéniques (HSCT) ont un risque élevé de contracter des infections fongiques invasives (IFI). L'utilisation de la prophylaxie antifongique reste un sujet de controverse et il n'y a pas de consensus clair sur le choix du traitement.

Treize essais randomisés contrôlés et leurs résultats ont été pris en compte pour établir différents modèles bayésiens hiérarchiques. Sur la base des résultats, les auteurs estiment qu'il existe des preuves suggérant que la prophylaxie par antifongiques est supérieure au placebo et que la prophylaxie avec le voriconazole est supérieure à la réduction de fluconazole en ce qui concerne l'incidence des IFI.

Une autre étude a comparé les différents traitements antidépresseurs utilisés chez les patients âgés souffrant de dépression majeure. Une revue exhaustive de la littérature a permis d'identifier les différents essais réalisés dans ce domaine indépendamment du comparateur utilisé et deux MTC ont été conduites en utilisant soit une approche bayésienne soit une meta-regression en ajustant les résultats sur la nature du comparateur (placebo ou autre), la sévérité de la maladie, la durée du traitement, l'âge... lorsque les données étaient suffisantes. Les résultats n'ont pas mis en évidence de différence majeure d'efficacité ou d'efficacité en vie réelle entre les différents anti-dépresseurs dans cette population.

Finalement si tous les intervenants du congrès s'accordent sur le fait que les essais randomisés contrôlés restent la meilleure méthode d'estimation de l'efficacité relative des interventions en santé, les comparaisons indirectes commencent à être de plus en plus utilisées et elles constituent désormais une option possible en l'absence d'essais randomisés pragmatiques. L'utilisation des comparaisons en réseau peut reposer sur plusieurs méthodes qui doivent toutefois être encore évaluées pour accroître la crédibilité des résultats.

2.4 Place des études observationnelles

Le congrès européen 2010 a été l'occasion pour la task-force de l'ISPOR sur les études observationnelles prospective de présenter ses réflexions, le rapport final devant être publié en juin 2011. Ce groupe de travail s'articule avec un autre groupe chargé de définir les bonnes pratiques en matière d'exploitation des bases de données rétrospectives pour l'évaluation économique des traitements⁽³⁾. Ces nouvelles recommandations de bonnes pratiques s'appuieront également sur des recommandations provenant d'autres sources telles que le Réseau Européen des Centres de Pharmacopépidémiologie et Pharmacovigilance (ENCePP).

Quand elles sont bien conçues, les études observationnelles ont un rôle important dans la production de données nécessaires aux évaluations d'efficacité relative dans la pratique clinique et dans les évaluations médico-économique. Elles recouvrent les registres, les études observationnelles prospectives et les études cas-témoin et s'opposent aux études naturalistiques portant sur un produit ou un groupe de produits.

Les questions abordées portent sur les nombreux biais et facteurs de confusion liées aux comparaisons entre différents traitements faites à partir de données observationnelles. Parmi ces derniers ont été surtout évoqués les problèmes posés par l'étude comparée de traitements commercialisés depuis un temps différent.

- Ainsi, l'observation d'un traitement récemment commercialisé pose le problème de la taille naturellement limitée de l'échantillon défini par le recours au nouveau traitement.
- Il y a alors un différentiel dans l'expérience des professionnels et des patients quant à l'usage du

³ Berger M, Mamdani M, Atkins D, Johnson M. Good Research Practices for Comparative Effectiveness Research: Defining, Reporting and Interpreting Nonrandomized Studies of Treatment Effects Using Secondary Data Sources: The ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report—Part I. *Value in Health* 2009 ; 12(8) :1044-52

médicament par rapport à d'autres traitements plus anciens. Le taux de mauvais usage ainsi que les pratiques vont évoluer dans le temps, au fur et à mesure de la diffusion de nouvelles informations sur le traitement.

- Le nouveau traitement sera d'abord utilisé par des médecins acceptant mieux l'innovation, donc différents des autres professionnels. Ils peuvent aussi être ceux qui traitent les patients les plus malades. Le même type de biais peut se présenter quand la communication du laboratoire s'adresse directement aux malades, entraînant les malades les plus à l'écoute à recourir ou exiger plus souvent le traitement.
- Lorsque l'on considère comme critère de sélection initial dans l'étude observationnelle, le fait de débiter une thérapie donnée, cette situation correspond souvent à l'échec d'un traitement antérieur. Il est alors difficile de comparer les pratiques et résultats du médicament étudiés avec d'autres comparateurs depuis beaucoup plus longtemps sur le marché.
- Finalement, les résultats observés en début de phase de commercialisation d'un traitement ne sont pas aussi bons que ceux rapportées dans des essais cliniques et ce, qu'il s'agisse d'efficacité ou de tolérance.

En résumé, les principaux problèmes rencontrés dans l'utilisation des données observationnelles sont liés aux évolutions qui surviennent dans le temps et qui ne peuvent être maîtrisées comme dans un essai randomisé. Toutefois, si les biais potentiels sont identifiés, il est généralement possible de prendre en compte comme facteur d'ajustement chacun des paramètres caractérisant ces derniers par exemple par l'intermédiaire d'un Score de Propension (PS) des distributions.

En ce sens l'utilisation des données observationnelles par les payeurs reste encore limitée et est source d'interrogations bien que ces données soient de plus en plus souvent exigées (études post-commercialisation).

Une illustration des problèmes rencontrés a été apportée dans le cadre d'une étude en vie réelle sur un traitement des douleurs réfractaires. Les auteurs montrent que les résultats du traitement étudié par rapport à ses comparateurs sont

largement influencés par la trajectoire thérapeutique antérieure des patients.

Une autre présentation a fait état des résultats d'une enquête auprès de 17 experts issus d'universités, d'agences d'HTA, de caisses d'assurance maladie, d'organismes de recherche et de l'industrie pharmaceutique - Royaume-Uni, France, Allemagne, Pays-Bas et États-Unis – et portent sur l'utilisation des données de vie réelle dans les évaluations médico-économiques de traitements. Tous les experts interrogés pensent que les RCT continueront d'être le support privilégié des évaluations en raison des limites des analyses en vie réelle et principalement des risques de biais.

Les nouvelles approches méthodologiques privilégiant les études de base de données aléatoires ou l'utilisation de techniques statistiques élaborées (par exemple les scores de propension intégrant de multiples paramètres) réduisent seulement partiellement les difficultés.

Les experts estiment toutefois que les évolutions actuelles plaident pour la mise à disposition et le partage de bases de données de dossiers médicaux électroniques sur des plateformes permettant l'analyse automatique et en temps réel des données en vie réelle. Mais ils pensent qu'il faudra du temps pour résoudre les difficultés techniques posées par le fait de relier les bases de données et, surtout, les questions de propriété des données et de protection de la vie privée.

Plusieurs experts prédisent qu'à court terme, les évolutions porteront surtout sur le développement de dispositifs électroniques de collecte de données : téléphones intelligents, tablettes et moniteurs à domicile, piluliers électroniques, etc. capable d'alimenter automatiquement des bases de données.

En conclusion, les mesures d'efficacité en vie réelle peuvent prendre une place importante dans le développement de médicaments, mais cela ne sera possible que grâce à des progrès graduels.

Plusieurs études montrent que les résultats d'efficacité en vie réelle sont souvent nettement moins favorables que ceux des essais expérimentaux ce qui a des conséquences à la fois sur les résultats des évaluations médico-économiques prises en compte dans certains pays pour décider du remboursement mais aussi sur les niveaux de prix qui sont acceptés au niveau européen et qui apparaissent ainsi surestimés.

3. NOUVEAUX MODES DE REGULATION

Le congrès européen de l'ISPOR a été aussi l'occasion d'échanges sur les stratégies adoptées par les différents

états pour réguler les dépenses de médicaments avec un focus important sur les démarches mises en œuvre au Royaume-Uni.

3.1 Dispositifs de régulation au Royaume-Uni

Une revue des données publiées sur le site web du NICE entre 2007 et 2010 concernant des médicaments indiqués dans différents types de cancer a été présentée. Parmi les 10 évaluations ayant conduit à des avis positifs, 4 ont bénéficié d'un remboursement inconditionnel. Par exemple, le rituximab a obtenu ce statut en raison de son efficacité relative comparativement à la chimiothérapie dans le traitement des LLC (gain en survie sans progression > 10 mois) ET parce qu'il répondait au seuil normatif fixé sur le plan de son ratio coût-efficacité dans cette indication (<30 000 £ par QALY gagné).

L'évaluation des 6 autres traitements a conduit à une décision de remboursement conditionnel. Ces produits étaient des thérapies ciblées où la survie totale ou sans progression était comprise entre 3 mois et 9 mois et/ou avait un coût par QALY gagné > 45 000 £. Cinq des six fabricants ont accepté soit un schéma d'accès au marché (PAS : Patient Access Scheme), se traduisant par des baisses de prix (cetuximab ; gefitinib) soit un dispositif de partage de risque (PFS : performance based schemes) imposant un remboursement du NHS par le laboratoire en cas d'échec du traitement (sunitinib, bortezomib, lenalidomid).

Globalement, toutes les thérapies ciblées offrant un gain de survie sans progression ou de survie totale supérieur à 3 mois sont donc prises en charge au Royaume-Uni mais lorsque les gains sont inférieurs à 9 mois ou que le ratio coût-efficacité est jugé élevé le prix du produit est, de fait renégocié à la baisse (éventuellement par le biais d'un PFS).

Une analyse approfondie du concept de PAS au Royaume-Uni a été également présentée. Conçue dans une optique de maîtrise par le ministère de la santé du Royaume-Uni il est intéressant de remarquer que ces PAS s'articulent avec le dispositif d'HTA et s'appuient sur les résultats de cette dernière. Le PAS est ainsi justifié s'il y a une (1) incertitude autour de l'effet au long terme de ce produit ou (2) la valeur de l'ICER est remise en question.

Dans le cas (1), le PAS peut consister en une baisse de prix négociée ou en l'approvisionnement gratuit du produit par l'industriel (ce qui revient pas exactement au même). Dans le cas (2), le PAS tente de diminuer l'ICER par une baisse de prix ou en liant paiement et résultats ou encore en mélangeant les deux solutions.

L'argument coût-efficacité seul explique la plupart des décisions du NICE. Le seuil estimé pour les traitements en

cancérologie se situerait en pratique autour de 40 000 £/QALY gagné donc au-delà du seuil officiel (20 000-30 000 £/QALY gagné). Cette position serait justifiée par une plus forte propension des patients à payer pour des traitements dans ce domaine.

Il est intéressant de remarquer l'effort du NICE pour assurer une certaine transparence dans ses décisions et sa volonté de devenir de plus en plus explicite sur la manière dont il utilise les résultats coût-efficacité dans la prise de décision, mais aussi les autres facteurs qu'il prend en compte.

Le développement des PAS et des PFS est présenté comme une solution permettant de faciliter l'accès au marché des traitements coûteux. La complexité de la gestion de ces dispositifs a imposé la création en Octobre 2009 du PASLU (Patient Access Scheme Liaison Unit) dans un système de santé qui a besoin de plus en plus de main d'œuvre administrative

3.2 Les accords de partage de risque

Les schémas de type PFS entrent dans le cadre général des accords de partage de risque (RSS : risk sharing schemes). Ces derniers ont été définis en 2006 par G. de Pourville de la manière suivante : "a contract between two parties who agree to engage in a transaction in which there are uncertainties concerning its final value. Nevertheless, one party has sufficient confidence in its claims of either effectiveness or efficiency that it is ready to accept a reward or a penalty depending on the observed performance."

Les industriels comme les agences d'HTA ont besoin de se doter de capacités analytiques pour évaluer a priori leur intérêt à travers des modélisations financières. Des outils analytiques doivent être développés pour permettre des révisions au fur et à mesure que les données cliniques ou de pratique deviennent accessibles

Ces stratégies sont de plus en plus utilisées dans différents pays comme le Royaume-Uni et l'Italie où, dans certains domaines thérapeutiques comme l'oncologie, et ces accords remplacent alors quasiment le mécanisme de remboursement normal. Les autres systèmes de santé semblent suivre. Les payeurs allemands, particulièrement les caisses d'assurance maladie, semblent très ouverts à ces approches ainsi que le payeurs néerlandais. Les modèles de RSS vont devenir de plus en plus populaires dans le futur proche, notamment en Allemagne. Les industriels doivent donc s'y préparer. Bien que d'autres modèles de risk sharing aient été introduits en Europe, les approches britanniques semblent constituer un repère pour les payeurs de nombre de pays européens.

3.3 Que faut-il financer ?

Les réflexions sur l'allocation des ressources en santé imposent de s'interroger sur notre capacité à financer tous les progrès observés. Ce questionnement a été illustré par deux présentations sur le thème de la cancérologie.

Selon l'OMS, plus de 11 millions de patients sont diagnostiqués comme atteints d'un cancer chaque année. Ce nombre atteindra 16 millions en 2020

Les coûts liés à la morbidité et aux décès prématurés dus au cancer sont considérables mais les coûts financiers engendrés par les traitements dans ce domaine sont également conséquents et représentent un poids important pour les patients, leur famille et la société dans son ensemble (le coût des cancers a été estimé représenter 192,4 milliards de dollars aux Etats-Unis en 2004).

D'un point de vue économique, ce ne sont pas les effets d'une pathologie (mortalité, morbidité) qui sont d'intérêt lors de la prise de décision relative à l'allocation des ressources. L'élément majeur à prendre en compte est la capacité à intervenir de façon efficiente afin de réduire les effets négatifs et il apparaît que les gains de santé permis par les nouveaux traitements ne sont pas toujours importants dans le secteur de l'oncologie (22% ne permettraient aucun gain de santé exprimés en termes d'amélioration de la survie totale et seulement 12% permettraient un gain supérieur à 1 QALY par cas).

Pour autant les traitements du cancer sont parfois traités différemment des autres médicaments pour des raisons tenant à la fréquence et à la sévérité de cette pathologie, à son intérêt médiatique, aux intérêts politiques en jeu, etc.

Une idée courante voudrait ainsi que l'on finance tous les traitements contre le cancer indépendamment de leurs coûts.

Ainsi, au Royaume-Uni, des dispositifs viennent d'être mis en place pour faciliter l'accès aux thérapeutiques innovantes sans qu'aucune évaluation médico-économique n'ait été réalisée. L'OLS Innovation Pass permet à certaines innovations d'être disponibles au sein du NHS (au moins pour un temps) sans qu'aucune évaluation n'ait été conduite par le NICE et on peut penser que ce seront essentiellement des traitements pour le cancer qui en bénéficieront.

Un tel dispositif suscite de vives réactions de la part des économistes de la santé pour différentes raisons :

- Accorder aux traitements contre le cancer un financement indépendant des gains de santé et les traiter différemment des autres médicaments revient à valoriser davantage la vie de patients atteints d'un cancer relativement aux autres patients ;
- Commercialiser des traitements sans tenir compte des coûts supprime les incitations encourageant les industriels à fixer les prix des traitements en fonction de leur gain thérapeutique.

Cette décision (qui s'inscrit en partie dans le cadre d'un programme de soutien aux industries des sciences de la vie au Royaume-Uni) est donc considérée comme non justifiée d'un point de vue éthique comme du point de vue de l'allocation des ressources en santé.

4. LA PLACE DES PATIENTS

Dans les systèmes de santé européens, la place accordée aux patients est croissante et l'usage des *Patient-Reported Outcomes* (PRO), mesures d'état de santé auto-évaluée par le patient, est grandissant. Les PRO recouvrent différents types de mesures. Les patients auto-évaluent ainsi leur état de santé, leur qualité de vie, leur adhérence au traitement, leur satisfaction, etc. De nouvelles technologies, de nouvelles méthodes de recueil des données (autres que les questionnaires papiers) sont maintenant disponibles et leur utilisation de plus en plus fréquente n'est pas sans poser question.

Sur un plan très différent la place des mesures de QALY continue à faire l'objet de débat et différentes interventions lors du congrès ont mis en évidence un intérêt pour les méthodes de révélation des préférences collectives comme critère d'aide à la décision en santé.

4.1 Le développement des recueils électroniques de données

L'utilisation de méthodes électroniques pour réaliser le recueil de données relatif au PRO se développe rapidement. Différents nouveaux dispositifs sont disponibles (version électronique des questionnaires disponible sur Internet, ordinateurs à écran tactile, PDAs, Smartphone, etc.). Ils facilitent la gestion administrative, permettent d'éviter des erreurs de saisie, permettent de transférer des données de façon instantanée, etc. Ces différents avantages incitent à leur utilisation. Cependant, plusieurs aspects doivent être pris en compte lorsque l'on met en place une étude faisant appel à ces outils. Il est ainsi nécessaire de

considérer les caractéristiques des répondants, la complexité du design de l'étude, la taille de l'échantillon, le nombre de sites, les zones géographiques, l'aspect pratique et logistique et les coûts entraînés qui demeurent un élément important de la prise de décision.

Toutes ces technologies ne s'adaptent pas à l'ensemble des situations (certains patients sont dans l'incapacité d'utiliser un ordinateur, d'autres n'ont pas accès à internet). Il est alors tentant de « mixer » dispositifs électroniques et questionnaires papier. Pourtant lorsque l'on utilise les deux types de méthodes pour la même étude, l'intégrité et la fiabilité des données peuvent potentiellement en être affectés.

L'usage de différentes méthodes pour recueillir des données dans une étude peut induire des biais de mesures car les méthodes ne fournissent pas nécessairement des données comparables. Ce type d'erreur réduit alors notablement la puissance statistique de l'étude.

Pourtant offrir la possibilité d'un recueil de données papier en sus d'un recueil électronique permet encore d'accéder à un échantillon plus large et plus représentatif (certaines populations n'ont pas accès à internet ou au téléphone, certains patients peuvent ne pas pouvoir s'adapter à toutes les méthodes) mais aussi de limiter les données manquantes. Des patients trop malades pour répondre à un questionnaire par le biais d'un ordinateur seront à l'origine de données manquantes simplement parce que la méthode de recueil de données n'est pas adaptée à leur situation. Ne pas prendre en compte les patients les plus malades affectera alors les résultats de l'étude et les effets observés seront probablement inférieurs aux effets réels.

Il est donc nécessaire avant de choisir la méthode utilisée de considérer attentivement toutes les sources potentielles de biais de mesure et d'en tenir compte lors de la phase de conception de l'étude.

Il apparaît préférable de minimiser les cas où l'on utilise différentes méthodes de recueil de données. Toutefois, le choix reste ouvert lorsque la situation s'y prête (dans le cas d'essais cliniques internationaux, tous les pays n'ont pas le même niveau d'équipement électronique et il peut être nécessaire de s'adapter). Enfin, lorsque différentes méthodes sont mobilisées, il est nécessaire d'adopter une seule approche pour l'ensemble du questionnaire et de ne pas combiner les modes de recueil pour différentes sections du questionnaire car cela peut poser des problèmes au niveau de l'interprétation des données.

4.2 Evaluation de la qualité de vie en cancérologie

Le cancer étant une pathologie de plus en plus présentes dans nos sociétés (deuxième cause de décès dans le monde

; 13,8% de l'ensemble des décès selon l'OMS en 2008) la quantité d'études réalisées en lien avec l'oncologie est croissante et, de ce fait, trois sessions de l'ISPOR 2010 ont été consacrées à ce sujet. Les études abordées étaient très variées : des études économiques (le coût de l'absentéisme pour cause de cancer, l'évaluation médico-économique de produit anticancéreux, la place des méta-analyses, etc.) mais également nombre d'études portant sur la prise en compte de la qualité de vie des patients atteints de cancer

Les évaluations de la qualité de vie liée à la santé (Health Related Quality of Life, HRQoL) demeurent un sujet de recherche très actif. Le nombre d'études présentant des résultats est de plus en plus important. Ces études sont subjectives (le patient répond au questionnaire), multidimensionnelles (elles prennent en compte la capacité physique, l'aspect émotionnel et l'impact de la pathologie sur la vie sociale) et dynamiques (réalisées à différents moments dans le temps).

En oncologie, les données de HRQoL permettent d'avoir une vision complète du traitement (effets cliniques et impact sur le patient et sa vie), spécialement dans le cas où les bénéfices cliniques additionnels sont difficiles à mettre en évidence. Un nouveau traitement peut avoir peu de bénéfices au niveau de la survie du patient mais se traduire par une amélioration notable de la qualité de cette survie relativement à d'autres traitements.

Les travaux de l'European Organisation for Research and Treatment of Cancer (EORTC) ont montré un décalage à ce niveau entre les déclarations des médecins et les déclarations des patients (les médecins ne percevant pas ce que les patients vivent réellement) ce qui justifie la nécessité d'interroger le patient directement.

4.3 Evolution des méthodes

Les méthodes d'élicitation des préférences telles que l'évaluation contingente et la méthode des choix discrets ont fait l'objet de plusieurs communications.

Ces méthodes visent à répondre au problème classique en économie posé par l'évaluation des biens non marchands. Par définition, pour estimer la valeur de ces biens on ne dispose pas de résultats monétaires directs et aucun marché ne permet de révéler les préférences des individus. Les méthodes d'élicitation des préférences que sont l'évaluation contingente et la méthode des choix discrets permettent alors de mettre à jour les préférences d'une population déterminée. La révélation de ces préférences est réalisée au moyen d'un marché contingent (fictif) construit à partir des réponses des individus à un questionnaire. Ce sont donc les intentions des individus qui sont observées et non pas leur comportement réel. Bien que ces méthodes soient encore d'application limitée en économie de la santé (spécialement dans certains pays), on note un

intérêt grandissant des différents acteurs (institutions, chercheurs et industrie) pour leur utilisation et elles pour-

raient être amenées à prendre une place plus importante dans l'aide à la décision en santé.

CONCLUSION

L'évaluation économique des stratégies de santé est à nouveau apparue être un champ en plein essor au cours de ce congrès de l'ISPOR 2010. On peut penser que l'intérêt de plus en plus marqué des agences de régulation pour intégrer une dimension économique dans leurs décisions va encore renforcer ce mouvement dans le futur. Les points de vue tendent à se rapprocher. Les tenants d'une régulation exclusivement financière et budgétaire ont évolué vers la prise en compte des questions d'efficacité. Inversement, la dimension parfois dogmatique des approches économiques en santé disparaît au profit de positions de compromis.

De manière significative, la question principale posée tout au long du congrès portait sur l'harmonisation des approches à l'échelle européenne et non sur leur acceptabilité.

On a pu ainsi voir combien il reste difficile de concilier les points de vue des différents acteurs et de faire évoluer les processus de décision dans un sens qui réponde à la fois aux contraintes économiques majeures auxquelles sont confrontées toutes les économies européennes mais aussi qui convienne à tous : patients, professionnels, autorités de régulation et payeurs.

Si les intervenants s'accordaient généralement sur des sujets importants : nécessité de continuer à traiter séparément l'évaluation de l'efficacité et l'évaluation de la valeur ajoutée, nécessité de dépasser les résultats produits par les approches expérimentales classiques pour intégrer au plus tôt des résultats d'efficacité relative si possible en vie réelle, intérêt pour les techniques nouvelles de comparaison indirecte ou les études observationnelles, ils ont dans le même temps fait part de leurs craintes par rapport aux changements en cours. Les questions de perte potentielle de souveraineté, de répartition des rôles entre agences du médicament et agences d'HTA étaient ainsi à l'ordre du jour.

Les apports méthodologiques et travaux de recherche ont donc été cette année en partie masqués par la prééminence des questions politiques. Une part significative du congrès non rapportée ici a concerné par ailleurs les problèmes spécifiques rencontrés par les pays européens de l'Est face aux enjeux du financement des soins et aux problèmes de régulation dans des économies aux ressources encore plus contraintes que celles des pays d'Europe de l'ouest.

Bruno **DETOURNAY** (CEMKA-EVAL)

Cette synthèse du congrès ISPOR est le fruit d'une collaboration entre le Collège des Economistes de la Santé et les Laboratoires Novartis France.



Tirage : 650 exemplaires.

Réalisation : Louyot.