

RÉSUMÉ

L'effet d'un médicament peut être différent en pratique quotidienne de ce qu'il était au cours de la recherche clinique ou au moment de la négociation de son prix avant sa mise sur le marché des médicaments remboursables. Les contrats conditionnels, fondés sur les résultats de santé, ont été créés pour se prémunir contre ce risque. Moyennant un accord rapide sur le prix, le laboratoire pharmaceutique prend l'engagement de rembourser en partie ou en totalité la valeur qui a été négociée si les résultats du traitement ne sont pas à la hauteur des espoirs suscités. Ces contrats d'accès au marché des médicaments remboursables sont étudiés ici à la lumière des schémas d'études qui président à leur mise en œuvre. Ils reposent pour l'essentiel sur des études descriptives ou normatives sans groupe témoin, études qui ne permettent pas de mesurer l'impact du médicament sur la santé. La fixation du prix du médicament sur de telles bases risque d'être fragilisée par la modeste qualité des preuves avancées pour en justifier l'octroi. Pour remédier à cette situation, nous proposons la mise en place d'études observationnelles comparatives, dont les populations devront être rendues comparables par la mise en œuvre des techniques micro-économétriques appropriées.

ABSTRACT

Pay-for-performance and conditional price setting of drug prices

When a drug is put on the market, it may not perform exactly the same way as it did at prior to obtaining market authorisation (namely, at the clinical research stage or during the pricing negotiation). This risk is now covered by a "conditional contract". A quick agreement on the price is granted provided pharmaceutical companies undertake to partially or fully refund the sum initially negotiated should the treatment be less efficacious than originally stated. Using a framework setting out application guidelines, we shall analyse these market access contracts for reimbursable products under the Social Security system. For the most part, these conditional contracts are drawn up on the strength of descriptive or normative studies without a control group. This means that the therapeutic impact of the drug cannot be measured. In such a context, price fixing for drugs may be jeopardised by the scanty quality of available evidence. In order to address this issue, we suggest setting up comparative observational studies and comparing groups with the aid of appropriate micro-econometric studies.

Paiement à la performance et fixation conditionnelle du prix du médicament

Robert Launois *, Salah Ghabri **, Lucia Fiestas Navarrete ***,
Jean-Gabriel Le Moine ****, Olivier Ethgen *****

* Réseau d'évaluation en économie de la santé (REES), Paris.

** Société française d'économie de la santé.

*** Réseau d'évaluation en économie de la santé (REES), Paris.

**** Réseau d'évaluation en économie de la santé (REES), Paris.

***** Unité de recherche en évaluation économique de l'innovation médicale (EMIR), université de Liège.

Introduction

Traditionnellement, l'introduction d'un médicament sur le marché européen est subordonnée à une autorisation de mise sur le marché (AMM), délivrée la plupart du temps par les instances administratives européennes après qu'elles se sont assurées de l'efficacité, de la sécurité et de la qualité du nouveau produit. Aujourd'hui, cette seule autorisation ne suffit plus. Tout ou partie du coût d'un traitement médicamenteux est désormais pris en charge par la collectivité, ce qui entraîne une intervention des pouvoirs publics dans la fixation de son prix et le choix des modalités de son remboursement. Les prix des médicaments sont négociés dans le cadre d'une convention signée entre le laboratoire et le Comité économique des produits de santé (CEPS) après que celui-ci a reçu les avis des deux commissions concernées de la Haute Autorité de santé : l'avis consultatif de la Commission de la transparence sur le service médical rendu et l'amélioration du service médical rendu et l'avis d'efficacité de la Commission d'évaluation économique et de santé publique. Le taux de remboursement est fixé par l'Union nationale des caisses d'assurance maladie. L'inscription sur la liste positive des médicaments remboursables est décidée par arrêté interministériel. Dans notre pays, le pouvoir du « payeur » est ainsi tricéphale : la Haute Autorité de santé, l'Assurance maladie et l'État se le partagent. Du fait de la rareté des ressources disponibles, chacune de ces parties prenantes tend à durcir, dans son domaine d'intervention propre, les conditions d'accès au marché du médicament remboursable.

Ces trois acteurs ont de plus en plus conscience d'être exposés à un double risque. Le premier risque porte sur la valeur réelle du médicament en termes d'efficacité. Il se peut que la spécialité pharmaceutique n'entraîne pas, sur la population traitée, les mêmes effets bénéfiques que ceux qui avaient été mis en avant au moment de la demande d'inscription sur la liste des médicaments remboursables. L'efficacité et les risques ont toutes les chances d'être différents en pratique réelle de ce qu'ils étaient dans le cadre d'une recherche clinique et cela pour trois raisons : l'observance est toujours meilleure dans le cadre d'un essai qu'elle ne l'est dans la vie courante, les populations incluses ne sont pas représentatives de la population cible – en cardiologie par exemple, les essais portent davantage sur les hommes que sur les femmes –, et les malades, recrutés sur la base de critères d'inclusion extrêmement restrictifs, reflètent mal la diversité des cas pris en charge en pratique médicale quotidienne. Le payeur risque donc de payer au « prix fort » des produits qui, en pratique quotidienne, se révèlent beaucoup moins performants qu'ils ne prétendaient l'être au moment de la négociation.

Toutes les conditions sont réunies pour qu'un second risque vienne se greffer sur le premier : « l'aléa moral ». Pour que le système de Sécurité sociale puisse fonctionner conformément aux missions qui lui ont été assignées, il faut que la probabilité du risque et le coût du remboursement des frais médicaux ne puissent

être influencés par le comportement des individus. Lorsque ces deux conditions ne sont pas remplies, le risque perd son caractère purement aléatoire, il devient subjectif et impondérable. L'assujettissement à la Sécurité sociale peut créer chez les assurés le sentiment qu'ils « ont droit » à un médicament en échange de leurs cotisations, même si l'effet n'est pas si bénéfique que prévu.

Face à cette situation, les pouvoirs publics peuvent être amenés à proposer aux industriels la signature d'accords dérogatoires aux conditions conventionnelles de fixation du prix du médicament. Cela sera le cas lorsque la négociation pour l'inscription sur la liste des médicaments remboursables entamée dans le cadre classique aura débouché sur une impasse. La mise en place des contrats conditionnels répond ainsi à un triple objectif : protéger les pouvoirs publics contre les incertitudes liées à l'estimation de l'utilité clinique d'un médicament, s'assurer de la soutenabilité du financement à mettre en place par les organismes sociaux et garantir à la population un accès rapide à l'innovation.

Les contrats passés entre les industriels et les autorités chargées de la négociation des prix sont désignés dans la littérature sous de multiples appellations : « *Managed Entry Agreements* » en Europe (European Commission, 2012), « *Risk-Sharing Agreements* » aux États-Unis (Neumann *et al.*, 2011), « *Patient Access Scheme* » au Royaume-Uni (Chapman *et al.*, 2003), « *Deeds of agreement* » en Australie (Australian Government – Department of Health and Ageing, 2009), « Accessibilité par la production de preuves » au Canada (McCabe *et al.*, 2010). La définition de leur objet et la standardisation des différentes classifications qui permettraient de procéder à leur regroupement s'avèrent d'autant plus délicates qu'au-delà des différences de terminologie, il existe des divergences plus profondes entre les auteurs sur la définition même du périmètre qui délimite leurs champs d'application.

Selon une conception restrictive proposée par Carlson (2010), De Pourville et Mongredien (2012) et le groupe de travail de l'International Society for Pharmacoeconomics and Outcome Research (Garrison et Towse, 2012), l'appellation « contrats de partage de risque fondés sur les performances » devrait être réservée aux seuls cas où un lien direct pourrait être établi entre les performances du médicament et la fixation de son prix, ou entre l'octroi d'un financement temporaire dédié et la collecte obligatoire de nouvelles données comparatives en vie réelle. Lorsque l'une de ces deux conditions est remplie, les engagements pris correspondent à un véritable contrat axé sur la recherche d'effets de santé. Devrait en revanche être exclu de la définition, toujours selon les mêmes auteurs, tout contrat à finalité financière puisqu'il ne s'agirait pas dans ce cas de véritables accords de partage du risque, mais plutôt d'accords entre les industriels et l'État sur le partage de la rente liée à l'innovation.

D'autres auteurs, dont Adamski (2010), intègrent les deux dimensions, scientifique et financière, car en tant que payeurs, ils se sentent investis de cette double responsabilité. À leurs yeux, les contrats régulés d'accès au marché du médicament remboursable recouvrent « tout type d'accord passé entre les entreprises du médicament et les pouvoirs publics en vue de limiter les conséquences de la prise en charge des traitements nouveaux ou anciens sur le budget des organismes assureurs, que ces conséquences soient dues à l'incertitude relative sur l'efficacité

réelle du traitement en condition normale d'usage ou aux contraintes imposées par l'existence de budgets limités».

Cette définition extensive des contrats organisant l'accès au marché du médicament remboursable reflète le point de vue des organismes payeurs européens, auxquels appartenait la majorité des coauteurs de la publication d'Adamski. C'est donc celle que nous retiendrons, en incluant dans notre examen les contrats financiers. L'élimination de ceux-ci du champ des engagements contractuels entre les industriels et l'État nous apparaîtrait en effet d'autant plus arbitraire que les mêmes méthodes d'audit sont utilisées dans le cadre de ces contrats pour vérifier le degré de réalisation de leurs objectifs, par rapport à des valeurs cibles prédéfinies.

Ces contrats pris au sens large, qui réintroduisent une flexibilité dans le système administré de fixation du prix des médicaments, constituent aujourd'hui un enjeu fondamental pour les industriels. Les règles qui structurent les procédures d'accès négocié au marché du médicament remboursable deviennent, de ce fait, un domaine de recherche à part entière. Ce domaine se différencie à la fois de la recherche clinique et des études postérieures à l'inscription sur la liste des médicaments remboursables.

Les essais randomisés contrôlés¹ de la recherche clinique permettent de répondre à la question sur l'efficacité théorique du traitement *in vitro* : « Est-ce que cela peut marcher ? ». Ils sont menés dans des conditions idéales définies de manière extrêmement rigoureuse par des protocoles qui prévoient le recrutement d'une population homogène de patients, une liste limitative des comorbidités et des traitements concomitants autorisés, une durée de suivi limitée à quelques mois ou à quelques années. De telles conditions constituent autant de limites à la généralisation de leurs conclusions.

Après que l'autorisation de mise sur le marché a été accordée, la question n'est plus de savoir si le traitement peut être efficace, mais s'il fonctionne sur des patients tout-venant, le cas échéant porteurs de poly-pathologies. Les études de terrain se proposent de vérifier que les résultats annoncés sont bien au rendez-vous, puisque rien ne garantit que les conclusions d'un essai soient transposables à toutes les situations cliniques. Ces enquêtes sont qualifiées d'observationnelles lorsqu'il est possible d'affirmer que le chercheur qui observe une population ou un phénomène n'influence pas les comportements ni n'interfère avec le cours naturel des choses (dosage et durée des traitements, modalités de prise en charge et de surveillance, risques encourus, etc.). L'intérêt de ces études est qu'elles prennent en compte la réalité des contextes pour les décrire de la manière la plus fidèle possible. Leur inconvénient majeur est que, le plus souvent, elles ne comportent pas de comparateur. Elles suggèrent l'existence de tendances dans les consommations ou les résultats sans en apporter la preuve formelle.

1 La méthode de l'essai contrôlé randomisé consiste à partager en deux groupes, par tirage au sort, les individus recrutés dans la population cible : l'un recevant le traitement (le groupe traité), l'autre ne le recevant pas (le groupe de contrôle). L'essai est dit « contrôlé » du fait de l'existence d'une population de référence, il est dit « randomisé » puisque la répartition des sujets entre les groupes est le fruit du hasard (aléatoire). De par la loi des grands nombres, les personnes des deux groupes ont en moyenne les mêmes caractéristiques observables et non observables, lorsque la taille des deux populations est grande. Dans cette hypothèse, les différences entre les résultats observés dans les deux groupes sont, toutes choses égales par ailleurs, exclusivement attribuables au traitement.

Comment peut-on combler l'écart entre l'expérimentation et la « vraie vie » ? Quels schémas d'étude permettent d'établir un *continuum* entre le contexte toujours un peu artificiel des études expérimentales et la réalité d'une prise en charge au jour le jour des études d'observation descriptives, tout en offrant la garantie qu'il existe un lien direct et exclusif entre le traitement administré et les résultats obtenus ? L'arbitrage entre la transposabilité des conclusions de l'étude (validité externe) et la démonstration d'un lien de cause à effet (validité interne) est au cœur de la réflexion sur la nature des études exigibles pour que les accords d'accès négocié au marché du médicament remboursable apportent ce que l'on en attend en termes de résultats sanitaires et réduisent l'incertitude liée à la mise sur le marché du traitement.

L'objectif de cet article sera de classer, dans un premier temps, les accords conclus en fonction de la nature des schémas d'études qui conditionnent leur mise en place. Dans un deuxième temps, nous soulignerons les limites d'une conception administrative du contrôle de l'efficacité et de la sécurité de l'usage du médicament en termes de normes à respecter. Nous proposerons enfin que les contrats dérogatoires de prix conditionnels soient établis non pas en fonction des « performances » obtenues par le médicament au regard des engagements virtuels négociés, mais sur la base de « l'efficacité directement et exclusivement attribuable au traitement » dans le cadre d'une étude observationnelle où un groupe de contrôle sera systématiquement introduit et les biais neutralisés grâce à l'utilisation des techniques économétriques appropriées.

Classification des contrats selon leurs obligations en termes de schémas d'étude : « performance » contre « résultat de santé »

Le préalable à toute procédure d'évaluation est d'être au clair sur ce que l'évaluateur cherche à prouver. Veut-il simplement vérifier que les promesses des essais cliniques sont corroborées par les résultats observés en pratique médicale quotidienne ? Ou souhaite-t-il s'assurer de ce que les résultats thérapeutiques constatés dans la « vraie vie » sont directement et exclusivement attribuables à un traitement ?

La première question est une question administrative, dont la réponse relève des techniques de l'audit : on compare ce qui est à ce qui devrait être. Dans la perspective de l'audit, c'est une véritable chaîne de contrôle qui se met en place, allant de l'édiction d'une norme à la vérification de son respect et à la sanction de sa violation.

La seconde est une question scientifique sur l'effet propre du traitement, qui permet de savoir si l'objectif thérapeutique, qui a été atteint en vie réelle, l'a été exclusivement de son fait. Les informations à recueillir dans les deux cas ne sont pas les mêmes. Dans le premier, il suffit de mettre en place une étude de suivi et de surveillance. Dans le second, il faut rechercher et neutraliser les interférences qui pourraient être partiellement à l'origine du résultat (« facteurs de confusion ») pour que seule la plus-value thérapeutique attribuable au médicament soit

spécifiée dans le cadre d'une véritable étude d'évaluation, afin que sa juste valeur soit reconnue et gratifiée d'un prix plus élevé. Chacune de ces questions appelle des schémas d'étude différents pour y apporter une réponse (Gertler *et al.*, 2010; Khandker *et al.*, 2009; Savedoff *et al.*, 2006).

Les deux démarches, ont été abusivement assimilées par Garrison (Garrison et Towse, 2012). Il recourt en effet à la même appellation pour les désigner sous le nom de «contrats de partage de risque fondés sur les performances», ce qui est, selon nous, à l'origine de bien des ambiguïtés. L'écart absolu entre les objectifs négociés et les résultats constatés dans la vraie vie devient, pour cet auteur, le déterminant essentiel du juste prix. Dans ce cadre, le prix du médicament serait ajusté à la hausse comme à la baisse en fonction du degré de réalisation des objectifs de santé contractuellement fixés, sans que l'on sache exactement à quoi est due cette convergence. Un suivi individualisé et exhaustif de l'évolution de la santé des patients dans le cadre de registres² permettrait, toujours dans ce contexte, de vérifier le respect des engagements pris sans s'inquiéter de leurs causes exactes.

Notre approche clarifie au contraire les enjeux en établissant un classement explicite des schémas d'étude par ordre croissant de fiabilité, selon la qualité des preuves qu'ils sont susceptibles de produire. Confrontés à l'arbitrage entre validité interne des études et transposabilité de leurs conclusions, nous proposons une hiérarchisation des plans d'expérience à mettre en œuvre en vie réelle. Cette hiérarchisation privilégie l'identification des effets de la cause au contrôle du respect des normes. Elle distingue clairement :

- les études de dénombrement des succès ou des échecs (études descriptives);
- les études de surveillance continue de la performance (études normatives);
- les études d'impact³ (études comparatives).

Les études descriptives et les études de surveillance continue de la performance ne permettent pas d'estimer l'impact du traitement. L'incertitude concernant l'efficacité en vie réelle et le coût du médicament persiste donc si l'on mène ces deux types d'étude. L'inscription d'un médicament sur la liste des produits remboursables et la fixation de son prix définitif ne peuvent être décidées qu'après

2 La création de registres dans le cadre d'engagements contractuels permettrait d'enregistrer de manière prospective et exhaustive les médicaments prescrits, sur une zone géographique donnée. Il conviendra : 1. de fixer clairement et limitativement les objectifs qui président à leur mise en place – respect des indications AMM, dénombrement des effets indésirables, recherche d'efficacité comparative. Chacun de ces objectifs implique des définitions différentes de la population cible et de la fenêtre temporelle de suivi; 2. de vérifier l'absence de double emploi avec les travaux actuels des observatoires des médicaments, des dispositifs médicaux, et des innovations thérapeutiques; 3. éventuellement de choisir une autre terminologie pour les désigner, afin de les distinguer clairement des registres «qualifiés» dont la création relève d'une décision de la commission d'évaluation des registres, mis en place auprès de l'Institut national de la santé et de la recherche médicale et de l'Institut national de veille sanitaire.

3 Une dénomination qui n'est pas sans rappeler les expressions utilisées par les épidémiologistes pour désigner la recherche de facteur causaux : «études étiologiques» lorsque la méthode consiste à mettre en évidence une relation entre un facteur de risque et une maladie ou «études exposés/non exposés» lorsque l'exposition est dichotomique, à cette différence près qu'il ne s'agit plus de rechercher la cause d'une maladie, mais d'identifier les causes des retombées positives ou négatives d'une action de santé ou d'un traitement médicamenteux ou non médicamenteux. D'autres appellations et acronymes recouvrent, à peu de choses près, les mêmes schémas d'études : *Comparative Effectiveness Research* aux États-Unis (Patient-Centered Outcomes Research Institute, 2012); *Relative Effectiveness Assessment* en Europe (European Network of Centres for Pharmacoepidemiology and Pharmacovigilance, 2010); analyse d'effectivité comparative ou AEC en France (plus couramment désignée sous le nom d'amélioration du service médical rendu), étude d'évaluation par opposition aux études de suivi et de surveillance précédemment citées. Le comparateur utilisé dans le cadre des quatre derniers schémas est toujours un traitement actif et non un placebo.

avoir identifié sa part de responsabilité dans les résultats obtenus en condition normale d'usage. Là où les études de performance ou les analyses descriptives se contentent de comparer les effets observés aux objectifs fixés ou de constater les faits en s'appuyant sur l'observation directe, les études observationnelles comparatives s'efforcent de neutraliser l'influence des facteurs de confusion pour qu'un lien de causalité direct et exclusif puisse être dégagé en vie réelle, entre le médicament prescrit et le résultat de santé obtenu.

Trois types de contrats axés sur des résultats de santé peuvent être distingués en fonction des schémas d'étude exigés pour suivre leur mise en place : suivi des performances par rapport à des référentiels normatifs prédéfinis, réalisation d'études de terrain à vocation purement descriptive ou réalisation d'études visant à mesurer l'effet propre du traitement.

Les contrats de performance par rapport à un référentiel normatif

Seront regroupés sous le nom de «contrats de performances par rapport à un référentiel normatif prédéfini» les contrats axés sur le respect des engagements pris sur des critères de morbi-mortalité ou d'efficience, ainsi que les contrats dont les clauses recouvrent essentiellement des engagements budgétaires. Les modalités de leur mise en place reposent, dans les deux cas, sur l'adoption de standards de référence qui définissent «sur le papier» le niveau de performance à atteindre pour que les engagements pris soient honorés. Un ensemble clair d'indicateurs est mis en place dans le cadre des activités de suivi, pour comparer les avancées réelles du programme à l'aune de ce qui a été prévu et convenu. Cela suppose que des valeurs de référence aient été contractuellement arrêtées à titre dérogatoire par le Comité économique des produits de santé et les industriels avant l'instauration du traitement et que des valeurs cibles réalistes aient été choisies afin de suivre les avancées futures par rapport à la situation initiale. Le calibrage de ces objectifs s'établira selon toute probabilité à mi-chemin entre les résultats publiés des essais randomisés et les résultats attendus dans la vie réelle. La comparaison des résultats obtenus et des objectifs contractuellement arrêtés mesurera la performance. Le caractère contractuel des standards ne doit cependant pas faire illusion, leurs clauses s'appliqueront à toute la population édictant *de facto* à son égard autant de normes qui s'imposeront à elle.

Engagements par rapport à des objectifs de morbi-mortalité ou d'efficience

Aux termes de l'article 10 ter de l'accord cadre signé entre le CEPS et le syndicat des entreprises du médicament (LEEM) le 10 octobre 2012 (Comité économique des produits de santé et LEEM, 2012), le payeur peut proposer au laboratoire une fixation conditionnelle du prix de son médicament. Le maintien, la baisse ou la hausse du prix sera fonction du «degré d'atteinte des objectifs» attesté par des «résultats précis observés sur les indicateurs de résultats contractuellement

arrêtés». L'intérêt de cette démarche contractuelle réside dans la définition claire des attentes des deux parties signataires (Eichler, 2001). Seront donc précisés ce que le payeur pourra demander en retour du prix qu'il accorde et les performances que les industriels s'engageront à réaliser sur des objectifs sanitaires ou en termes d'efficience par rapport à d'autres indicateurs fixés par les deux parties au contrat. Le respect de ces engagements permettra de pérenniser le prix accordé, voire de l'augmenter si les performances relevées sur le terrain sont supérieures en valeur absolue au niveau des résultats de santé qui avait été négocié *a priori*. Leur inexécution entraînera un remboursement total ou partiel des dépenses médicamenteuses ou éventuellement de certaines dépenses de soins liées à la prescription de la spécialité. Ce remboursement prendra la forme d'une remise tenue secrète ou d'un reversement des sommes mises sous séquestre, sans modification du prix nominal.

Ces engagements souscrits dans le cadre d'un référentiel défini *a priori* peuvent porter aussi bien :

- sur le contenu du service rendu au malade, défini soit en termes de résultats intermédiaires, soit en termes de résultats finaux sur ce que l'on est en droit d'attendre du traitement (garantie de résultat);
- sur l'activité de soins elle-même. Aux normes portant sur le produit s'ajoutent des normes qui précisent les modalités de prise en charge des malades. L'emploi d'un test biomoléculaire peut être ainsi spécifié dans le cadre d'une médecine personnalisée⁴ pour s'assurer, avant la mise en œuvre du traitement, de ce que les patients ont toutes les chances d'être répondeurs (garantie de bonne pratique);
- sur le comportement attendu des patients; ainsi en va-t-il de l'observance dont l'industriel pourra se porter garant (garantie de bon usage).

L'accord passé en 1999 entre la North Staffordshire Health Authority et le laboratoire Parke-Davis (aujourd'hui Pfizer) sur la prescription de statines dans les hypercholestérolémies fournit un bon exemple du premier type de dispositif susceptible d'être mis en place (Garrison et Towse, 2012). L'industriel entendait démontrer, sur la base du résultat de ses essais cliniques, qu'une baisse significative des taux de cholestérol LDL pouvait être obtenue dans la vie réelle, grâce son produit, l'atorvastatine (Tahor®). Le critère d'efficacité choisi était la proportion de patients dont le taux de concentration du cholestérol LDL était inférieur au seuil de 3,0 mmol/l après traitement. Le taux de réponse observé dans les essais chez les patients présentant une hypercholestérolémie initiale modérée (taux de LDL

4 Il existe à ce jour plusieurs définitions de la médecine personnalisée. Pour les uns, l'appellation devrait être réservée aux seuls cas où des biomarqueurs sont mis en œuvre et recouvrerait essentiellement l'usage des tests-compagnons qui permettent d'administrer le bon traitement, au bon patient, au bon moment et à la bonne dose : « Un biomarqueur se réfère à une caractéristique objectivement mesurée et évaluée qui permet de distinguer les processus biologiques normaux des cas pathologiques ou de servir d'indicateur de réponses pharmacologiques à une intervention thérapeutique. Les tests-compagnons permettent de déterminer le statut du patient vis-à-vis de biomarqueurs d'intérêt clinique ». Cité in Marty R. et Roze S. (2013), « Valorisation et modélisation médico-économique du couple test diagnostique compagnon et thérapies ciblées », Groupe de travail « Biomarqueurs » du LEEM, Les Entreprises du médicament – HEVA HEOR, Paris.

Pour les autres, le terme désignerait moins la personnalisation effective d'un traitement que la constitution de sous-groupes de populations, ciblés en fonction du profil médical des individus qui les composent. Un grand nombre de descripteurs peuvent alors contribuer à la définition de ceux-ci. Les tenants de cette approche préfèrent parler « de médecine stratifiée » en fonction des effets hétérogènes du traitement (EHT) plutôt que de médecine personnalisée. Nous avons fait nôtre la seconde définition.

baseline compris entre 3,0 et 4,8 mmol/l) atteignait 89% lorsqu'ils recevaient 10 mg d'atorvastatine.

Partant du principe que les taux d'efficacité en pratique quotidienne ne pouvaient pas être identiques à ceux relevés dans les essais du fait d'une moins bonne observance des patients tout-venant, le laboratoire fit preuve d'une grande prudence dans la négociation de ses engagements. Il accepta le principe de devoir rembourser au National Health Service (NHS) des dépenses associées à la consommation de son produit, uniquement lorsque la proportion de patients qui avaient répondu au traitement en vie réelle se révélait être inférieure à 71%.

Un programme d'audit fut mis en place sur un financement du laboratoire. Un taux de réponse de 88% fut relevé. Puisque le seuil d'efficacité acceptable avait été fixé par convention à 71%, aucun reversement ne fut effectué.

L'analyse des audits réalisés en France sur la consommation des molécules onéreuses et des dispositifs médicaux implantables dans les hôpitaux permet de mieux appréhender la nature profonde des contrats de performance. Les agences régionales de santé (ministère de la Santé et des Sports, 2010) procèdent en deux temps :

- sur la base de critères qualifiés de «médico-administratifs», elles s'assurent de la conformité des prescriptions par rapport à l'autorisation de mise sur le marché ou aux référentiels établis par l'Institut national du cancer (INCa), l'Agence nationale de sécurité du médicament et la Haute Autorité de santé. Par exemple, dans le cas d'un traitement contre le cancer : décompte du nombre de cures de chimiothérapie administrées par rapport au nombre recommandé par le résumé des caractéristiques du produit, calcul des doses cumulées, etc.;
- sur la base de critères dits «économiques», elles vérifient que les taux de croissance fixés chaque année pour encadrer l'évolution des dépenses des molécules onéreuses qui sont facturées par l'hôpital à la Sécurité sociale, en sus des tarifs forfaitaires qu'elle paye dans le cadre du programme de médicalisation des systèmes d'information (PMSI)⁵, ont bien été respectés. Aucune allusion n'est faite à une éventuelle démarche en termes d'efficience, et pour cause, puisque les coûts supportés cessent d'être rapportés aux résultats obtenus.

Dans les deux illustrations choisies, qu'il s'agisse du cas anglais ou du cas français, les évaluations conduites sont des «évaluations normatives», au sens où l'entendait Contandriopoulos *et al.* (1993) ou la Cour des comptes américaine (US General Accounting Office, 1991). Autrement dit : «Un jugement qui est porté sur une intervention en comparant les ressources mises en œuvre, les services ou les biens produits et les résultats obtenus à des critères et à des normes de référence». L'évaluation des résultats sanitaires relève donc exclusivement, dans ce cadre, d'une approche en termes d'écarts par rapport à des valeurs cibles. Le fait que ces normes puissent être négociées dans un cadre contractuel ne changera rien quant à la nature purement juridique et virtuelle du rôle qui est le leur en tant que comparateur (Gridchyna, 2012).

5 Au titre des groupes homogènes de séjours (GHS).

Engagements sur des objectifs budgétaires à respecter

Les « ententes » (Bourassa Forcier et Noël, 2012) à vocation financière qui sont passées entre le payeur et les compagnies pharmaceutiques sont fondamentalement différentes dans leurs finalités, sinon dans leurs modalités, de celles des contrats de santé. Elles ont en effet pour objet de contenir les conséquences « économiques » du prix négocié. Ces contrats financiers peuvent être regroupés en deux catégories, selon la formulation des engagements qu'ils comportent, au niveau de l'individu ou au niveau de la population cible.

Les contrôles individuels des dossiers des patients par les médecins de la Sécurité sociale nécessitent une informatisation des dossiers médicaux et requièrent le suivi de nombreux paramètres. La collecte et le traitement de l'information peuvent donc s'avérer fastidieux pour les professionnels, d'autant qu'ils risquent d'alourdir une activité clinique déjà chargée.

Les contrats financiers passés par rapport à une population cible ont essentiellement pour objet de veiller au respect des enveloppes budgétaires. Ils reposent sur le suivi de données agrégées (nombre de patients traités, nombre de doses administrées...). Ils sont donc assez faciles à mettre en œuvre et à examiner.

Pendant, le classement des contrats de performances sur référentiel défini *a priori* en termes d'engagements sur des résultats thérapeutiques ou budgétaires ne doit pas masquer la porosité des frontières qui les séparent. Certains auteurs ne voient ainsi dans le contrat bortezomib (Velcade®) signé entre Janssen-Cilag et le NHS dans l'indication « myélome multiple » qu'un accord financier, alors que pour d'autres, il s'agit clairement d'un contrat de santé. En réalité, le contrat emprunte ses caractéristiques à chacun des deux systèmes. Les versements en cas d'inexécution des engagements sont déclenchés par un critère de jugement clinique, une réduction du taux de protéine M inférieure à 40% au bout de quatre semaines de traitement. Néanmoins, l'accord comporte également une clause de plafonnement des dépenses par patient, laquelle relève d'une logique purement financière.

Les études de terrain descriptives sur l'efficacité du médicament

Des études observationnelles avec ou sans comparateur peuvent être mises en place afin de s'assurer de la réalité des gains de santé obtenus (Carlson *et al.*, 2010; Carbonneil *et al.*, 2009).

Dans le cas des études observationnelles sans comparateur, les financements sont accordés sous condition : la prise en charge du produit est subordonnée à la démonstration *ex post* de son efficacité. Cette garantie repose sur le simple constat de ce qui s'est passé dans la vie réelle des patients, sans qu'il soit nécessaire de faire appel à une définition conceptuelle ou contractuelle des valeurs cibles. Aucune référence normative ne préside à l'interprétation des résultats. On se contente de décrire et de dénombrer les événements observés conformément au principe de « voir pour payer ». De simples observations permettent d'appréhender directement l'évolution de la maladie sous traitement et de distinguer les patients qui répondent au traitement de ceux qui y sont réfractaires dans le cadre

d'études descriptives : le traitement a réussi ou il a échoué. En cas d'échec, une procédure de récupération des règlements indus doit être mise en œuvre au cas par cas. Aucune référence n'est faite aux antécédents médicaux du sujet, à son état de santé général, en bref à tous les facteurs de confusion qui peuvent être à l'origine de l'échec du traitement.

Les études de terrain comparatives sur l'efficacité du médicament

À l'inverse des études observationnelles sans comparateurs, les financeurs peuvent souhaiter « payer pour voir », en accordant des financements temporaires dédiés à la réalisation d'études comparatives en pratique médicale quotidienne. Dans ce cas, il s'agit de réunir de nouveaux éléments de preuves en vie réelle sur l'utilité clinique ou médico-économique du produit, grâce à une collecte de données complémentaires dans le cadre d'études observationnelles avec groupe de contrôle⁶ (Haute Autorité de santé, 2011 ; Ranque *et al.*, 2011). Ces schémas d'études visent à comparer un groupe de patients ayant bénéficié du médicament avec un groupe de non-bénéficiaires, qui présentent des caractéristiques similaires et ont eu des trajectoires les plus proches possibles. Ces études estiment l'effet « propre » du traitement en établissant des différences entre les deux groupes. Autrement dit, on cherche à mesurer ce qui se serait passé en l'absence de prescription du médicament pour mieux en mesurer les effets.

Les programmes de recherche lancés en France en 2002 par la direction générale de l'offre de soins (Carbonneil *et al.*, 2009 ; ministère de la Santé, 2012) dans le cadre des programmes « Stratégies innovantes et coûteuses » (PSTIC) ou en 2005 par Medicare aux États-Unis sous l'appellation « *Coverage with evidence development* » (Jaroslowski *et al.*, 2011), constituent les véritables prototypes de ce type d'accords. Ils ont été conçus l'un et l'autre avec le même objectif : faciliter l'accès des patients à des médicaments prometteurs tout en permettant de collecter des données probantes complémentaires. Le financement temporaire dédié peut être réservé aux seuls patients enrôlés dans l'étude ou bénéficier à l'ensemble des patients éligibles. Dans ce second cas, le financement des traitements est subordonné à la réalisation d'une étude clinique sans que tous les patients soient obligés d'y participer pour bénéficier de la couverture financière.

⁶ Cohortes prospectives ou rétrospectives, registres, études cas témoin, banques de données médico-administratives, etc.

Les limites des études de performances et des études de terrain descriptives

Les faits ne sont pas nécessairement des preuves

L'existence de biais est consubstantielle aux études observationnelles en condition normale d'usage. L'épidémiologiste David Sackett (1979) avait identifié 35 types de biais. La dernière publication faite sur le sujet en recense 70 (Delgado-Rodriguez et Llorca, 2004). Recueillir des données factuelles en situation réelle d'usage n'implique donc pas qu'un lien de cause à effet ait été scientifiquement démontré. Les risques d'interférences sont multiples : il peut s'agir tout aussi bien de biais de sélection que de facteurs de confusion.

Dans le premier cas, celui du biais de sélection, la présence de cette erreur systématique fausse l'estimation des données recueillies et invalide le calcul de leurs résumés statistiques. L'estimation des paramètres d'intérêt au niveau de l'échantillon étudié est une mesure inexacte de leurs vraies valeurs dans la population. L'appellation recouvre à la fois le biais d'auto-sélection (un patient participe à une étude sur la base du volontariat), le biais d'indication (un médecin conseille à son patient de pratiquer un sport à titre préventif, ce qui biaise l'étude de l'association entre sédentarité et maladies cardiovasculaires) ou le biais par survie sélective (si certains patients dans l'indication décèdent précocement, ce qui entraîne pour eux le fait de ne pas être inclus dans l'étude). Toute analyse ne prenant pas en compte ces trois types de biais aboutit à des indicateurs erronés.

Dans le second cas, celui des facteurs de confusion, c'est la présence de tiers facteurs qui entraîne des erreurs dans l'interprétation des résultats. Prenons un exemple simple (Sonis, 1998) : il existe une forte relation entre le fait d'avoir des cheveux gris et le fait d'être frappé d'un infarctus du myocarde. Le rapport de risque liant ces deux grandeurs est significativement supérieur à 1 (il s'élève à 1,73). En réalité, derrière cette liaison immédiate, se cache un troisième facteur, l'âge, qui est étroitement associé à la couleur du cheveu et qui est également lié à la survenue des infarctus. Lorsque l'on introduit cette donnée dans l'analyse en calculant un risque relatif ajusté sur l'âge, l'association entre les cheveux gris et l'infarctus est égal à 0,80. Autrement dit, il n'existe aucun rapport réel entre ces deux grandeurs. De même, la guérison d'un rhume banal après l'administration d'un traitement antiviral peut être liée à une résolution spontanée de la maladie, beaucoup plus qu'au traitement lui-même. Dans ce cas, l'évolution naturelle de la maladie est un facteur de confusion, c'est-à-dire un facteur qui produit un effet similaire à celui que l'on attend du traitement et qui peut être mis sur le compte du traitement, alors même que celui-ci est dépourvu d'efficacité.

La confusion de la terminologie dans la littérature, le caractère laudatif et moraliste du discours sur les performances cachent l'essentiel. Le simple constat des faits, dressés dans le cadre d'études observationnelles sans comparateur, ni le degré de réalisation des objectifs fixés par les contrats à la performance, n'apportent de nouvelles preuves permettant de consolider les résultats obtenus dans le cadre des essais randomisés. L'absence de réponse au traitement ou le constat

d'un décès du patient dans le cadre d'un financement conditionnel (Carlson *et al.*, 2010) ne sauraient donc justifier le déclenchement de la procédure de récupération des versements indus, dont les industriels seraient censés avoir bénéficié. En effet, de multiples causes peuvent avoir été à l'origine de l'échec du traitement. De la même façon, aussi longtemps que l'influence des interférences n'a pas été neutralisée, il est impossible d'évaluer si les objectifs ont été atteints du fait de l'intervention médicamenteuse ou du fait d'autres facteurs.

L'atteinte des valeurs cibles ne permet pas d'évaluer la plus-value thérapeutique

L'atteinte des valeurs cibles fixées dans le contrat ne permet pas non plus d'apprécier la valeur ajoutée du nouveau traitement. D'une part, les études observationnelles avec référentiel comparent la réalité et la règle, ce qui est et ce qui devrait être. L'incertitude qui entoure l'utilité clinique du médicament dans la vie réelle demeure. Une situation qu'illustrerait parfaitement l'apologue de la mise en service d'un passage clouté dans une impasse, passage qui n'a aucun impact sur les accidents de la voie publique, puisque par définition aucun véhicule ne l'empruntera jamais. C'est moins son utilité qui est évaluée, que le respect du code de la route qui fait l'objet de contrôle.

D'autre part, les taux de réussite négociés dans le contrat et observés dans les études peuvent être égaux sans que cette égalité soit directement et exclusivement attribuable à l'usage du médicament. On peut légitimement imaginer, dans une étude observationnelle, que les patients recevant le médicament présentent des caractéristiques particulières et que celles-ci auraient joué en leur faveur, même s'ils n'avaient pas participé à l'étude. Si le traitement s'avère bénéfique, on doit s'attendre à ce qu'ils en retirent un avantage encore plus grand. Le vrai bénéfice attribuable au médicament a dans ce cas toutes les chances d'être surestimé. Ce raisonnement fonctionne en sens inverse lorsqu'il s'agit d'effets indésirables graves. Le raisonnement en termes de performances présente ainsi des failles importantes sur le plan scientifique.

La validation délicate des critères de jugement intermédiaires

Pour disposer rapidement des éléments d'information nécessaires, le décideur est souvent amené à privilégier l'utilisation de critères intermédiaires validés. Ceux-ci doivent avoir présenté, dans le cadre d'essais randomisés, une corrélation avec les paramètres cliniques qu'ils remplacent. Par exemple : mesures unidimensionnelles des lésions cibles et addition des plus grands diamètres pour calculer les taux de réponse en cancérologie ; taux d'hémoglobine glyquée (HbA1c) en diabétologie avec une norme de référence entre 4 % et 6 % chez les individus non diabétiques ; valeur seuil du T score de densité osseuse pour prévenir la survenue de fractures ostéoporotiques, etc. Mais pour que le test puisse être considéré comme

un substitut au critère clinique qu'il est censé représenter, il est nécessaire qu'il prédise complètement la réponse au traitement administré (EUNETHTA, 2012).

Si les mots ont encore un sens, nous devons garder en mémoire les étapes constitutives de la chaîne de causalité : des moyens techniques, financiers et humains sont mobilisés pour déployer des activités de recherche qui fournissent des produits et des services (le nombre de patients pris en charge) dont les effets (le nombre de réponders au traitement par exemple) sont mesurés par la mise en place d'une batterie d'indicateurs de performance : autant de variables intermédiaires ou d'objectifs temporaires, dont il convient de démontrer que les modifications ou le degré de réalisation contribuent effectivement à l'amélioration de l'état de santé de la personne malade et au prolongement de sa vie. Ce prolongement est l'impact final de toute la chaîne, le seul vrai résultat, celui qui compte véritablement aux yeux du patient puisque le soin *in fine* n'a qu'une seule raison d'être : son utilité clinique. Les termes utilisés pour désigner le critère de jugement ultime (impact, résultat de santé, résultat différentiel, efficacité comparative, ou utilité clinique relative) sont autant de synonymes dont la mesure, après rapprochement avec les données similaires qui ont été relevées sous d'autres traitements, est la seule qui permette de dégager la vraie valeur ajoutée de l'innovation médicamenteuse.

Cette règle parfaitement générale s'applique notamment vis-à-vis des biomarqueurs. Le nombre de ceux pour lesquels une bonne corrélation avec le critère de jugement clinique a été démontrée est encore très réduit. Il est exceptionnel d'obtenir une réponse du type « tout ou rien » (Jaroslowski *et al.*, 2011). Il existe, le plus souvent, des « faux positifs » (qui alimentent l'inquiétude des patients en suscitant de « fausses alarmes ») et des « faux négatifs » (qui sont autant de « fausses sécurités »). Comme le soulignent les députés Alain Claeys et Jean-Sébastien Vialatte (Office parlementaire d'évaluation des choix scientifiques et technologiques, 2013), « la caractérisation strictement biologique des tumeurs demeure très partielle ». En 2012, selon l'INCa, il n'existait que huit biomarqueurs, dont l'usage conditionnait l'emploi de onze thérapies ciblées (Institut national du cancer, 2013). Il serait vain de croire que la médecine dite personnalisée permettra de donner le bon traitement au bon patient, au bon moment de façon quasiment automatique, dès lors qu'un biomarqueur aura été utilisé. Subordonner le financement d'une molécule innovante à la mise en œuvre préalable du test compagnon correspondant demeure encore l'exception.

La valeur pronostique des critères de résultats temporaires demeurera controversée tant que leur corrélation avec le critère de jugement ultime, c'est-à-dire avec le prolongement de la vie ou l'amélioration de la qualité de vie, n'aura pas été clairement établie.

Comment administrer la preuve de la causalité en vie réelle ?

La vocation de ces études observationnelles avec comparateur, que certains auteurs qualifient d'études d'impact, n'est plus de répondre à la question « Est-ce

que ça peut marcher ?» ni à la question «Est-ce que ça marche dans la vraie vie ?», mais de comprendre «Pourquoi ça marche ?». Leur objectif est d'isoler l'effet propre du traitement dans les résultats obtenus en distinguant la part des résultats qui est attribuable au médicament et la part qui est due à d'autres causes (caractéristiques particulières du patient, préférences du prescripteur, modalités de fonctionnement du système de soins, etc.)

L'imputabilité individuelle de l'effet d'un traitement : un idéal inaccessible

Pour évaluer le lien de cause à effet qui existe entre un traitement et un résultat de santé pour un individu donné, il faudrait comparer le bénéfice que le sujet a retiré de la mise en place du traitement au résultat qui aurait été le sien s'il n'avait pas été traité ou s'il avait reçu un traitement différent. On compare une situation réelle et observée à une situation virtuelle et inobservée, communément qualifiée de «situation contrefactuelle». Cette comparaison est nécessairement une fiction, puisque le même sujet ne peut pas au même instant être bénéficiaire et non-bénéficiaire d'un traitement. Il y a forcément l'une des deux situations qui n'existe pas dans la réalité. Une situation contrefactuelle ne peut être estimée qu'en reconstruisant la situation inobservée à partir de la situation observée chez ceux qui n'ont pas reçu le traitement et en calculant la différence d'effet moyen (Rubin, 1974).

Une simple comparaison «avant-après» ne saurait constituer une méthode d'évaluation crédible, puisqu'elle ne tient pas compte des facteurs externes qui peuvent changer au cours du temps. Ce raisonnement conduit à utiliser des observations concernant des populations d'individus en identifiant les «clones» des bénéficiaires pour définir le groupe de contrôle. La méthode statistique choisit donc de ne pas se concentrer sur l'effet causal individuel mais sur le calcul de l'effet causal moyen du traitement dans la population des patients traités.

Dans les études avec comparateur les biais d'indication doivent être maîtrisés

Les traitements ne sont pas prescrits au hasard. L'indication du médecin est posée en raison des caractéristiques observables, cliniques, fonctionnelles ou comportementales des patients. Or ces paramètres sont autant de facteurs pronostiques de l'évolutivité de la maladie. Les choix thérapeutiques dépendent également de facteurs de risque subtils (par exemple le désir du médecin d'apporter une réponse à une souffrance psychologique ou sociale du patient liée au chômage), qui sont autant de facteurs parasites impossibles à prendre en compte dans ce type d'étude.

Dans les essais contrôlés randomisés, l'affectation aléatoire garantit la comparabilité des populations entre les deux groupes car elle neutralise les caractéristiques personnelles observables et non observables des patients inclus dans

chacun d'entre eux. L'introduction d'un placebo permet d'isoler l'effet propre du traitement en retranchant ce qui est observé sous traitement de ce qui est observé dans le groupe contrôle, celui-ci n'ayant d'autre objet que de quantifier les effets des différents facteurs confondants. Les études observationnelles, même avec comparateur, n'offrent aucune de ces sécurités. En l'absence de randomisation, les populations comparées peuvent être dissemblables.

Les différences observées dans la survenue des événements peuvent ne pas être liées au choix thérapeutique mais aux raisons pour lesquelles ce choix a été fait. L'évaluateur peut se méprendre sur l'origine des résultats constatés et attribuer, à tort, leur caractère bénéfique ou néfaste au seul médicament, alors que certains peuvent être liés à la présence des facteurs de confusion. Ceux qui oublient la présence de ces biais risquent d'attribuer à tort l'intégralité de l'effet total du traitement mis en œuvre à l'effet propre du médicament, alors qu'une partie ne lui est aucunement attribuable.

La comparaison de la moyenne des deux populations risque, en effet, de mélanger deux effets : l'effet propre du traitement et l'effet de population, lequel résulte des différences préexistantes entre les niveaux de santé initiaux de ceux qui ont bénéficié du traitement et de ceux auxquels le traitement n'a pas été administré.

Pour assurer la validité des études permettant de fixer le prix d'un médicament, il est impératif de distinguer et d'isoler l'effet des caractéristiques observables ou inobservables des sujets de l'effet spécifique du médicament sur les variables d'intérêt, en mettant en œuvre les techniques économétriques appropriées pour neutraliser les différents types de biais (Behaghel, 2012 ; Givord, 2010) : appariements *a priori* par score de propension (Denis *et al.*, 2006 ; Dhainaut *et al.*, 2007 ; Payet *et al.*, 2007 ; Riou Franca *et al.*, 2008 ; Rosenbaum et Rubin, 1984), mise en œuvre de la technique de la double différence, utilisation de variable instrumentale (Grootendorst, 2007) ou ajustements *a posteriori* par analyse multivariée.

Neutralisation des biais dus aux facteurs observables

Au moment du choix du schéma d'étude, il convient d'adopter une démarche en trois temps :

- concevoir en premier lieu, un modèle d'exposition indépendamment des effets attendus du traitement, en estimant *a priori* la probabilité que les sujets en bénéficient sur la base du résumé de leurs caractéristiques observables, quel que soit le groupe auquel ils ont été affectés ;
- procéder ensuite à l'appariement des individus du groupe traité et de ceux du groupe de contrôle qui ont des caractéristiques proches, en les mettant en correspondance sur la base de leur score de propension respectif⁷. Le fait d'être bénéficiaire ou non du traitement devient alors aléatoire (ou tout au moins indépendant de la variable de résultat), comme dans un essai randomisé, et le risque de biais est neutralisé aussi bien sur le plan clinique qu'en termes de coût ;
- construire, dans un troisième temps, un modèle d'effet causal en étudiant la différence entre la moyenne observée chez les bénéficiaires du traitement et la

7 Ce qui suppose que les densités de distribution de ces scores aient un domaine de définition en commun.

moyenne relevée chez les non-bénéficiaires qui présentent les mêmes scores de propension par quintiles.

Neutralisation des biais dus aux facteurs non observables

Les éventuels biais dus aux facteurs non observables peuvent être neutralisés en mettant en œuvre la technique de la double différence. Le principe de la méthode est le suivant : un suivi initial est mis en place avant l'introduction du nouveau traitement, puis répété après son introduction. Les mêmes observations sont relevées de façon homothétique sur le groupe témoin. On calcule la différence pour chacun des groupes avant et après traitement et l'on soustrait de la différence qui a été observée sur le groupe expérimental, celle qui a été constatée sur le groupe témoin. Cette méthode permet de discriminer la part du changement de l'état de santé qui est imputable au médicament de la part liée à l'évolution générale du contexte, en neutralisant ainsi l'évolution des facteurs d'environnement qui est supposée avoir été la même dans les deux bras de l'étude.

Une seconde méthode est actuellement très largement utilisée pour exploiter les bases de données médico-administratives, elle consiste à tirer parti des informations disponibles sur les types de consommations de soins pour en déduire l'état de santé des assurés sociaux qui en ont bénéficié (Schneeweiss *et al.*, 2009). La littérature internationale considère à l'heure actuelle qu'il convient de séparer les « nouveaux utilisateurs » d'un produit médicamenteux et les « cas prévalents », afin d'analyser la consommation et la fréquentation du système de soins des seuls « cas incidents », dans les six mois qui ont précédé leur mise sous traitement. L'identification de ces consommations dans le cadre d'un score de propension haute densité (hd-PS) et l'appariement des groupes traités et non traités à l'aide de cet instrument permettent de repérer, en fonction des variables « empiriques » retenues (consultations, hospitalisations, etc.), les différences d'état de santé qui préexistaient, avant traitement, entre les populations.

L'état de santé d'un patient est évalué à partir des médicaments qui lui ont été délivrés, sur prescription du praticien qui a posé le diagnostic, à la suite de la consultation que le patient a sollicitée et à partir des symptômes ressentis. Ces variables de substitution dépendent elles-mêmes des facilités d'accès aux soins, de la gravité de la maladie, des habitudes de prescription du médecin et de l'importance des tickets modérateurs laissés à la charge du patient. L'ajustement des consommations sur des variables qui sont de bons substituts à des variables non observables revient à ajuster les résultats sur les facteurs inobservables eux-mêmes. Ces variables empiriques doivent donc être intégrées en plus des variables cliniques dans les scores de propension enrichis. Plus nombreuses sont les variables de substitution intégrées dans le score de propension, plus grandes sont les chances de prendre en compte les éléments de contexte inobservables qui sont autant de facteurs de confusion à neutraliser.

Conclusion : de l'argent pour la valeur ou de la valeur pour son argent ?

Dans les pays où il existe des valeurs seuils définissant l'effort collectif socialement acceptable par unité monétaire dépensée pour la santé (Pays-Bas, Autriche, Royaume-Uni, Nouvelle-Zélande, Canada), l'objectif premier des contrats de performances est d'obtenir un meilleur rapport coût-efficacité. Cet objectif risque d'être dénaturé dans les pays qui ne disposent pas, explicitement ou implicitement, de ce type de repère. À défaut de seuils de référence, nous n'aurons à notre disposition que des données de cadrage médico-économique intuitives pour conduire la négociation. Celle-ci reposera sur la conviction que la science peut démontrer ce qu'elle veut et que l'échange humain paraît beaucoup plus sûr. En l'absence de référentiel économique, les clauses d'efficience risquent ainsi de se transformer en simples accords sur la pertinence des soins.

Dans cette hypothèse d'accord *a minima*, les composantes du raisonnement seront amenées à être dissociées les unes des autres et étudiées séparément. Les coûts, désormais appréhendés pour eux-mêmes, seront entendus dans leur acception comptable la plus étroite, sans être rapportés à l'apport thérapeutique des produits. L'évaluation de la valeur sera analysée par rapport à des critères de résultats prédéfinis en termes exclusivement médicaux, sans être rapprochée des coûts.

Cette méthode signe l'abandon des analyses coût-efficacité et le retour au système d'assurance-qualité des années 1990, lequel est associé en Europe à des schémas de réduction budgétaire.

L'accord cadre signé le 5 décembre 2012 entre les industries du médicament et le comité économique des produits de santé ne privilégie *a priori* aucune des deux formes d'étude possibles (études d'impact avec comparateur ou études descriptives ou normatives à un bras). Son article 10 ter permet de conditionner le prix de certains médicaments aux résultats « démontrés en vie réelle en termes de santé publique » ou en fonction de l'écart constaté entre les résultats escomptés et les engagements contractuels pris avec le laboratoire concerné, sur la base des indicateurs retenus par les parties signataires. Le choix des méthodes à mettre en œuvre pour évaluer les effets de l'utilisation des médicaments nouveaux en vie réelle n'est pas spécifié plus avant, mais la HAS sera appelée à jouer un rôle clé dans la définition des études. L'article 11, alinéa 7 prévoit que « préalablement à la rédaction du protocole, la HAS et le CEPS s'accordent sur les objectifs de l'étude et les questions à traiter en conséquence. L'alinéa 8 du même article spécifie que « le protocole de l'étude est soumis à la HAS afin qu'elle donne son avis sur la capacité de l'étude à répondre aux questions posées ». Enfin, l'alinéa 9 précise que la « HAS évalue les résultats et fixe avec le CEPS le degré d'atteinte des objectifs ». Une étroite collaboration entre les deux institutions sera de mise.

Derrière le débat apparemment technique entre les auteurs qui se font les chantres de la simplicité et ceux qui militent pour la mise en place de véritables études scientifiques en l'absence de randomisation, se cachent deux conceptions radicalement différentes de la régulation du prix du médicament. Pour les uns, le prix doit être modulé en fonction de la valeur thérapeutique du produit constatée

par rapport à un niveau de performance contractuellement fixé en valeur absolue (Garrison et Towse, 2012). Pour les autres, la détermination du prix au moment de la réévaluation du produit doit prendre appui à la fois sur les évaluations comparatives d'efficacité et d'efficience en vie réelle (Schneeweiss *et al.*, 2011). C'est cette seconde voie que nous recommandons d'emprunter.

Références bibliographiques

Adamski J., Godman B., Ofierska-Sujkowska G., Osinska B., Herholz H., Wendykowska K., Laius O., et al. (2010), « Risk Sharing Arrangements for Pharmaceuticals: Potential Considerations and Recommendations for European Payers », *BMC Health Serv Res*, vol. 10, n° 153, p. 1-29.

Australian Government – Departement of Health and Ageing (2009), Guidelines for deeds of agreement for the pharmaceutical benefits scheme, Canberra.

Banerjee A. et Duflo E. (2009), « L'approche expérimentale en économie du développement », *Revue d'économie politique*, n° 119, p. 691-726.

Bequemont, L., Bordet, R. & Cellier, D. (2012). « La médecine personnalisée : comment passer du concept à l'intégration dans un plan de développement clinique en vue d'une AMM ? » *Thérapie*, 67, pp.339-348.

Behaghel L. (2012), *Lire l'économétrie*, Paris, Éditions La Découverte, collection « Repères ».

Bourassa Forcier M. et Noël N. (2012), « Ententes entre gouvernements et compagnies pharmaceutiques », Montréal, Centre interuniversitaire de recherche en analyse des organisations.

Carboneil C., Quentin F., Lee-Robin S. H. (2009), « A Common Policy Framework for Evidence Generation on Promising Health Technologies », *Int J Technol Assess Health Care*, n° 25 suppl 2, p. 56-67.

Carlson J.J., Sullivan S.D., Garrison L.P., Neumann P.J., Veenstra D.L. (2010), « Linking Payment to Health Outcomes: a Taxonomy and Examination of Performance-based Reimbursement Schemes Between Healthcare Payers and Manufacturers », *Health Policy*, n° 96, p. 179-90.

Chapman S., Reeve E., Rajaratnam G., Neary R. (2003), « Setting up an Outcomes Guarantee for Pharmaceuticals: new Approach to Risk Sharing in Primary Care », *BMJ*, n° 326, p. 707-709.

Comité économique des produits de santé et LEEM (2012), Accord cadre du 5 décembre 2012 entre le Comité économique des produits de santé et les entreprises du médicaments, Paris.

Contandriopoulos A.P., Champagne F., Denis J.-L., Pineault R. (1993), « L'évaluation dans le domaine de la santé : Concepts et Méthodes », *Bulletin*, n° 33, p. 12-17.

De Pourville G. et Mongrédien L. (2012), « L'accès au marché remboursé pour les médicaments : Les contrats de partage de risque fondés sur les résultats », Cergy-Pontoise, Le Kremlin-Bicêtre, ESSEC Business School/Collège des économistes de la santé.

Delgado-Rodriguez M. et Llorca J. (2004), « Bias », *J Epidemiol Community Health*, n° 58, p. 635-641.

Denis P., Launois R., Devaux M., Berdeaux G. (2006), « Comparison of diurnal intraocular pressure control by latanoprost versus travoprost: results of an observational survey », *Clin Drug Investig.* n° 26, p. 703-714.

Dhainaut J.F., Payet S., Vallet B., Franca L.R., Annane D., Bollaert P.E., Le Tulzo et al. (2007), « Cost-effectiveness of Activated Protein C in Real-life Clinical Practice », *Crit Care*, n° 11, R99.

Eichler R. (2001), «Utilisation des paiements basés sur la performance pour améliorer les programmes de santé», *Le Management*, n° 10, p. 1-22.

EUNETHTA (2012), Levels of Evidence – Applicability of Evidence in the Context of a Relative Effectiveness Assessment of Pharmaceuticals, Copenhagen, European Network for Health Technology Assessment.

European Commission (2012), Proposal for a Directive of the European Parliament and of the Council relating to the transparency of measures regulating the prices of medicinal products for human use and their inclusion in the scope of public health insurance systems, Brussels.

European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (2010), ENCePP Guide on Methodological Standards in Pharmacoepidemiology, London.

Garrison L. et Towse A. (2012), Performance-Based Risk-Sharing Arrangements – Good Practices for Design, Implementation and Evaluation: An ISPOR task force Report, International Society for Pharmacoeconomics and Outcomes Research (ISPOR).

Gertler P.M., Martinez S., Premand P., Rawlings L.B., Vermeersch C.M.J. (2010), « Impact Evaluation in Practice », Washington DC, The International Bank for Reconstruction and Development/The World Bank.

Givord P. (2010), Méthodes économétriques pour l'évaluation de politiques publiques, Paris, INSEE.

Gridchyna I. (2012), «Utilisation de la norme juridique comme instrument de régulation du marché des médicaments innovants en Europe et en France», Doctorat, université Bordeaux 2.

Grootendorst P. (2007,) « A Review of Instrumental Variables Estimation in the Applied Health Sciences », *Health Serv Outcomes Res Methods*, n° 7, p. 159-179.

Haute Autorité de santé (2011), Choix méthodologiques pour l'évaluation économique à la HAS, Paris.

Institut national du cancer (2013), «Situation de la chimiothérapie des cancers», Rapport 2012, Paris.

Jaroslowski, S. & Toumi, M. (2011). «Design of patient access schemes in the UK: influence of health technology assessment by the National Institute for Health and Clinical Excellence». *Appl Health Econ Health Policy*, 9, 209-15.

Khandker S.K., Koolwal G.B., Samad H.A. (2009), *Handbook on Impact Evaluation: Quantitative Methods and Practices*, Washington DC, The International Bank for Reconstruction and Development/The World Bank.

McCabe C.J., Stafinski T., Edlin R., Menon D. (2010), « Access with evidence development schemes : a framework for description and evaluation », *Pharmacoeconomics*, n° 28, p. 143-52.

Ministère de la Santé et des Sports (2010), Circulaire DSS/1C/DGOS/PF2 n° 2010-389 du 12 novembre 2010 relative à la mise en œuvre des dispositions relatives à la maîtrise des produits de santé des listes en sus/actions locales à conduire en 2010 et 2011 (application du dispositif de régulation).

Ministère de la santé (2012), «Prise en charge de l'innovation par la DGOS : de la recherche clinique à la diffusion aux patients» [en ligne]. <http://www.sante.gouv.fr/prise-en-charge-de-l-innovation-par-la-dgos-de-la-recherche-clinique-a-la-diffusion-aux-patients.html>, consulté le 19 décembre 2013.

Neumann P.J., Chambers J.D., Simon F., Meckley L.M. (2011), « Risk-sharing arrangements that link payment for drugs to health outcomes are proving hard to implement », *Health Aff (Millwood)*, n° 30, p. 2329-2337.

Office parlementaire d'évaluation des choix scientifiques et technologiques (2013), Étude de faisabilité de la saisine sur «les enjeux scientifiques, technologiques et éthiques de la

médecine personnalisée», Paris, Assemblée nationale [en ligne]. <http://www.assemblee-nationale.fr/opepst/ba14.asp>, consulté le 19 décembre 2013.

Patient-Centered Outcomes Research Institute (2012), *Our Questions, Our Decisions: Standards for Patient-Centered Outcomes Research*, Washington DC.

Payet S., Riou Franca L., Le Lay K., Vallet B., Dhainaut J. F., Launois R. (2007), «Évaluation coût-efficacité de la drotrécogine alfa comparée à la prise en charge conventionnelle dans le traitement du sepsis sévère en pratique réelle», *Journal d'économie médicale*, n° 25, p. 207-223.

Ranque B., Mechtouff L., Grabar S. (2011), «Épidémiologie étiologique : du facteur de risque à la cause», *Sang Thrombose Vaisseaux*, n° 23, p. 242-252.

Riou Franca L., Payet S., Le Lay K., Launois R. (2008), «Drotrecogin alfa's impact on intensive care workload in real life practice: a propensity score approach », *Value in Health*, n° 11, p. 1051-1060.

Rosenbaum P.R., Rubin D.B. (1984), « Reducing Bias in Observational Studies Using Subclassification on the Propensity Score », *Journal of the American Statistical Association*, n° 79, p. 516-524.

Rubin D. (1974), « Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies », *Journal of educational psychology*, n° 66, p. 688-701.

Sackett D.L. (1979), « Bias in analytic research », *J Chronic Dis*, n° 32, p. 51-63.

Savedoff W.D.L., Levine R., Birdsall N. (2006), « When will we ever learn? Improving Lives through Impact Evaluation », Washington DC, Center for Global Development.

Schneeweiss S., Gagne J.J., Glynn R.J., Ruhl M., Rassen J.A. (2011), « Assessing the Comparative Effectiveness of Newly Marketed Medications: Methodological Challenges and Implications for Drug Development », *Clin Pharmacol Ther.* n° 90, p. 777-790.

Schneeweiss, S., Rassen, J. A., Glynn, R. J., Avorn, J., Mogun, H. & Brookhart, M. A. (2009). «High-dimensional propensity score adjustment in studies of treatment effects using health care claims data». *Epidemiology*, 20, 512-22.

Sonis J. (1998), « A Closer Look at Confounding », *Fam Med*, n° 30, p. 584-588.

US General Accounting Office (1991), *Designing evaluations*, Washington.