

**L'évaluation post-AMM
des médicaments :
Un modèle d'analyse et de prévision
du niveau de service médical rendu
des médicaments remboursables en France**

Claude LE PEN (*),
avec la collaboration de
*Gaël PRIOL et Hervé LILLIU (**)*

(*) **Université Paris-Dauphine**

(**) **CLP-Santé**

**Rapport pour les
Laboratoires Internationaux de Recherche**

Avril 2001

1 - LA SYNTHÈSE

« Panier de soins » et médicaments : Une analyse critique de la réévaluation du Service Médical Rendu par les médicaments remboursables

Le contexte

Le malaise des professions de santé, la crise entre l'Etat et les organismes d'assurance-maladie et la faillite du mécanisme de régulation budgétaire ont mis en évidence la nécessité d'une réforme en profondeur de l'assurance-maladie en France. Dans ce cadre, un des thèmes majeurs de discussion concerne la définition des biens et services remboursables, ce qu'il est convenu d'appeler, de manière impropre et inélégante, « le panier de soins ».

Une certaine confusion règne sur ce sujet. Il faut d'abord insister sur le fait que le « panier de soins » n'est pas une simple liste de biens et services remboursables et encore moins un tarif. Le concept de « panier de soins » est en effet l'union d'une liste de biens et services et d'une enveloppe financière servant à les financer. Le « panier de soins » est ainsi un outil de régulation des dépenses et pas seulement un instrument de définition de droits.

Par ailleurs, le concept de « panier de soins » implique en toute rigueur une définition « sociale » des critères d'entrée, alors que les procédures actuelles privilégient les critères médicaux. En caricaturant un peu, on peut dire que la logique du panier de soins conduit à inclure tous les biens et services que la collectivité souhaite prendre en charge dans le cadre du financement socialisé, indépendamment de leur efficacité médicale. Les procédures actuelles valorisent au contraire principalement – voire exclusivement – la dimension médicale.

Dans le cas du médicament, il est possible d'objecter que le critère de Service Médical Rendu « suffisant », introduit en 1999, correspond à un élargissement de la problématique traditionnelle du remboursement et fait prévaloir des notions de santé publique. Le décret 99-915 du 27 octobre 1999 stipule en effet que l'appréciation du Service Médical Rendu « [...] prend en compte l'efficacité et les effets indésirables du médicament, sa place dans la stratégie thérapeutique, notamment au regard des autres traitements disponibles, la gravité de l'affection à laquelle il est destiné, le caractère curatif symptomatique du traitement médicamenteux et son intérêt pour la santé publique. »

En utilisant les résultats publiés de la réévaluation, nous montrerons que le « SMR en pratique » est beaucoup plus restrictif que le « SMR en théorie » et que l'efficacité et la tolérance du médicament, déjà analysé au moment de la délivrance des AMM, joue un rôle décisif dans sa détermination.

Méthodes

Nous sommes partis des résultats de l'évaluation du SMR portant les 1453 produits de la première tranche, correspondant à quatre domaines thérapeutiques, la cardiologie, la rhumatologie, les maladies du métabolisme et la psychiatrie.

Pour chaque présentation nous disposons des données relatives à leurs caractéristiques, telles qu'elles avaient été appréciées par la Commission de la Transparence :

- La gravité de la pathologie (classées en trois catégories)
- L'efficacité et la sécurité du produit (4 niveaux)
- La nature du traitement (Symptomatique, préventif, curatif)
- La place dans la stratégie thérapeutique (Traitement de 1^{ère} intention, de 2^{ème} intention, de recours, d'appoint)
- L'existence d'alternatives thérapeutiques (Oui, non)
- L'intérêt en termes de santé publique (Oui, Non)

On disposait en outre du classement final en termes de Service Médical Rendu :

- Pas de SMR attribué (23% des présentations)
- SMR Faible (7,5%)
- SMR modéré (6%)
- SMR important ou majeur (63%)

A l'aide d'un modèle de régression logistique multinomial cumulatif à hasard proportionnel ("*proportional odds model*"), nous avons recherché l'existence d'un lien entre les caractéristiques des produits et leur classement final. Il s'agissait d'«expliquer» au sens statistique du terme le classement final sur l'échelle de SMR à partir des caractéristiques reconnues aux produits.

Résultats

Les modèles obtenus sont hautement significatifs et donnent une excellente valeur prédictive. En considérant le meilleur modèle comme un « test de SMR important » (probabilité d'avoir un SMR important ou majeur, compte tenu des caractéristiques), sa sensibilité¹ serait de 97%, sa spécificité² de 89% et sa valeur prédictive³ de 94%.

En ce qui concerne la relation caractéristiques des produits et le niveau de SMR, les principales conclusions de l'analyse sont les suivantes :

- Alors que la méthode prévoit, conformément à la loi, une liste exhaustive de critères d'évaluation, il apparaît qu'en réalité deux d'entre eux, l'efficacité et la gravité de la pathologie, suffisent à expliquer très largement (à plus de 90%) le classement sur l'échelle du SMR. Les autres critères apportent finalement peu de valeur ajoutée.
- Ces deux critères jouent de manière « lexicographique » : un produit d'efficacité importante a une probabilité très élevée d'être classé SMR important, indépendamment des

¹ Pourcentage de médicaments avec SMR important ayant été classés comme tels (900/926).

² Pourcentage de médicament sans SMR importants ayant été classés comme tels (468/527).

³ Pourcentage de médicaments avec une probabilité de SMR important >0.50, ayant effectivement ce classement.

autres critères. La gravité de la pathologie joue en revanche un rôle déterminant pour les produits d'efficacité modérée ou modeste.

- Certains critères souffrent visiblement d'un manque de clarification, ce qui conduit à une interprétation variable, selon les classes, voire selon les produits ou groupes de produit. C'est, en tout premier lieu, le cas du critère « intérêt pour la santé publique » dont le pouvoir explicatif n'est pas négligeable, mais dont on peine à voir comment il a été utilisé : dans la plupart des cas, il est redondant avec l'efficacité et/ou la gravité. Dans d'autres cas, au contraire, il s'en écarte sans qu'il se dégage une ligne d'interprétation très claire
- Un point important est la relative hétérogénéité inter-classes thérapeutiques du processus d'évaluation, au moins pour les niveaux intermédiaires de SMR. Un produit d'efficacité moyenne dans une pathologie de sévérité moyenne possède moins de chance d'obtenir un SMR important en cardiologie que dans les autres classes. Cela tient à un nombre important de produits que la Commission a sans doute souhaité déclasser dans cette catégorie. Toujours est-il qu'une procédure d'harmonisation inter-classes, voire à un niveau plus fin, devrait en toute logique compléter le dispositif.
- Enfin, il existe des cas de discordance entre le classement « prédit » sur la base de l'analyse statistique et le classement effectif. Ces discordances sont peut-être justifiées au regard d'autres critères que ceux officiellement pris en compte. Mais il importe alors que ceux-ci soient explicités et justifiés.

Conclusion

L'analyse suggère finalement, que l'attribution d'un niveau de SMR repose sur une logique technico-médicale très proche de celle utilisée dans le cadre de l'AMM. La plupart des produits « sans niveau de SMR attribué » sont d'ailleurs des produits anciens mis sur le marché avant la généralisation des méthodes modernes d'évaluation.

Par ailleurs, l'analyse n'est pas exempte d'incohérences internes. Elle manque notamment d'homogénéité inter-classes : les probabilités conditionnelles d'obtenir un SMR important ne sont pas identiques d'une classe thérapeutique à l'autre.

Clairement, d'autres facteurs « invisibles » interviennent dans la décision. Si on considère par exemple que le Viagra® est un produit efficace, dans une pathologie de gravité intermédiaire (de grade B) et qu'il s'agit d'un traitement curatif de première intention pour lequel il existe des alternatives thérapeutiques, alors 201 produits de la base ont un profil identique et ils sont tous remboursés !

Il est enfin curieux que les notions économiques soient totalement absentes d'une évaluation ayant pour finalité la prise en charge financière des produits par la collectivité !

Finalement, l'analyse confirme que la pratique actuelle des listes et tarifs n'obéit pas à la finalité et à la logique qui sous-tend la notion de « panier de soins » et d'une réelle réflexion méthodologique sur les critères de prise en charge collective des médicaments – innovants ou non – et sur leur impact sur la santé publique reste à venir.

2 - L'ETUDE

I. Une problématique nouvelle : l' « évaluation post-AMM » des médicaments

Dans beaucoup de pays occidentaux, on assiste depuis quelques années à une tendance forte au développement de ce que nous appellerons « l'évaluation post-AMM » des médicaments. On définira ainsi l'ensemble des procédures d'évaluation destinées à statuer sur le prix, le remboursement et l'usage des médicaments qui ont fait préalablement l'objet d'une autorisation de mise sur le marché délivrée par les agences d'enregistrement selon les règles en vigueur. Cette « évaluation post-AMM » doit être distinguée du suivi de pharmacovigilance et des études de « phase IV ». Le premier, destiné notamment à repérer les effets secondaires rares non identifiés durant le développement clinique, existe depuis longtemps et se justifie pleinement dans une perspective de santé publique. Quant aux études de « phase IV », elles sont mises sur pied par les firmes après commercialisation, le plus souvent dans un but médico-promotionnel. A l'inverse, l'évaluation post-AMM est en général imposée aux laboratoires par les instances publiques ou privées qui prennent en charge financièrement le médicament, dans le but d'en fixer les conditions de prise en charge, en fonction de son bénéfice et, éventuellement, de son coût.

Si l'AMM gouverne l'accès au marché, l'évaluation post-AMM a une influence importante sur la taille du marché, en volume et en valeur, à travers notamment le remboursement et le prix. La démarche repose sur l'idée générale que l'évaluation classique, effectuée selon les règles de Bonne Pratique Clinique (GCP, « Good Clinical Practice ») dans le cadre du développement clinique et de la délivrance des autorisations de mises sur le marché, ne suffit à régler ni la question du prix, ni celle remboursement, ni même celle du « Bon Usage » des produits. C'est une idée controversée : selon un point de vue couramment exprimé, notamment dans l'industrie, l'évaluation clinique, qui est censée garantir l'efficacité, la sécurité et la qualité des produits, devrait se suffire à elle-même et aucune évaluation postérieure, aucun « fourth hurdle », ne peut intervenir sans restreindre, d'une manière ou d'une autre, le droit d'accès des patients à des médicaments innovants, efficaces et sûrs et la liberté de prescription des médecins.

Les partisans d'une « évaluation post-AMM » rétorquent ⁴ :

- Que l'évaluation effectuée lors de la délivrance des AMM est conçue pour mettre en évidence l'efficacité du produit (« *efficacy* ») dans les conditions les plus favorables à ce dernier, notamment à travers le choix de la population cible et du comparateur. De même les conditions de surveillance optimale du traitement sont optimales. Une évaluation postérieure prenant en compte la dimension réelle de l'utilisateur en médecine courante donc justifiée.
- Que les libellés des AMM portent essentiellement sur le produit et non sur les stratégies

⁴ Voir par exemple la présentation « NICE » sur le site internet du NICE, www.nice.org.uk

thérapeutiques, au sens large. Il serait donc légitime de proposer – ou d'imposer – en aval de l'évaluation clinique, des recommandations de pratique clinique réglementant l'usage des médicaments, dans une perspective large et comparative.

- Et surtout, que la décision de mise sur le marché n'est pas de même nature que la décision de remboursement et de tarification. Des autorités de régulation – ou des caisses d'assurance maladie – peuvent légitimement élaborer des critères de prise en charge qui soient différents des critères de mise sur le marché, et notamment des critères économiques.

De nombreux exemples de cette « *évaluation post-AMM* » existent déjà. C'est notamment le cas en France où, depuis le début des années 80, la « Commission de la Transparence », désormais rattachée à l'Agence de Sécurité Sanitaire des Produits de Santé (AFSSAPS) évalue le « service médical rendu » par les nouveaux produits et décide de l'opportunité d'un remboursement fondé sur ce critère. Une évaluation du même type, quoique moins formelle, se retrouve dans les pays où des « listes positives » règlent l'accès au remboursement. Nous y reviendrons largement dans cet article.

Mais la mise sur pied d'une telle évaluation post-AMM est certainement l'un des traits saillants des réformes intervenues récemment dans le NHS britannique. La « *restricted list* » des médicaments, établie en 1985, a par exemple, connu un regain d'intérêt notamment à travers la décision de procéder à un remboursement conditionnel du Viagra® [1]. Mais c'est surtout, la création au sein du NHS du désormais fameux National Institute for Clinical Excellence (NICE) qui matérialise le mieux cette tendance. Sa mission est de céder à une évaluation de type « post-AMM » des services médicaux qu'il finance et, en tout premier lieu, des médicaments, nouveaux et anciens [2]. Les avis rendus sur l'antiviral Rlenza®, sur les interférons β dans la sclérose en plaques ou sur le Xenical® dans l'obésité, ont donné lieu à de nombreux commentaires et à quelques polémiques [3].

La pratique de l'évaluation post-AMM n'est pas limitée aux systèmes d'assurance-maladie publics. Aux Etats-Unis, les assureurs privés développent depuis longtemps des « formulaires », c'est-à-dire des listes plus ou moins restrictives de médicaments qu'ils acceptent de prendre en charge de manière préférentielle. Dans beaucoup des pays, les hôpitaux, publics ou privés, se sont dotés de « P&T committees », des comités de prescription et de thérapeutique, qui fixent la politique d'achat et d'utilisation des nouvelles molécules. Mais l'enjeu de l'évaluation « post-AMM » est évidemment plus important dans les systèmes publics où elle détermine très largement la taille du marché d'un nouveau produit.

Les conditions de cette évaluation sont très hétérogènes : hétérogénéité des objectifs, des acteurs, des procédures. Selon les pays, voire même à l'intérieur des pays, les objectifs de cette évaluation varient selon les institutions qui la pratiquent. Pour les unes, l'objectif est avant tout le remboursement et/ou le prix. Pour les autres, c'est surtout l'établissement de « guidelines » d'utilisation en pratique quotidienne. D'autres peuvent poursuivre les deux objectifs simultanément. De même, les acteurs impliqués sont divers et recourent des fonds d'assurance maladie privés ou publics, des établissements hospitaliers, des agences générales spécialisées dans l'évaluation des technologies médicales ou des agences spécialisées dans l'évaluation des technologies médicales. Enfin, les procédures d'évaluation post-AMM, contrairement à celles qui régissent le développement clinique, ne sont pas harmonisées et les exigences concernant le niveau de preuve et les modalités de son administration sont extrêmement variables. Elle concerne les produits nouveaux mais aussi les produits anciens, déjà sur le marché : l'agence française, l'AFSSaPS, a procédé à une telle évaluation du stock

de produits remboursé et l'agence portugaise, Infarmed, vient d'annoncer une évaluation du même type pour 500 produits existants⁵.

Un élément essentiel de l'évaluation post-AMM est l'importance qu'y jouent les facteurs économiques. Si aucune grande agence d'enregistrement ne semble, pour l'instant, disposée à inclure une dimension économique dans la délivrance des autorisations de mise sur le marché, ce n'est évidemment pas le cas de l'évaluation post-AMM. C'est même une de ses raisons d'être que de constituer, en aval de l'enregistrement, un point d'ancrage de cette fameuse « fourth hurdle », qui est redoutée par les industriels, mais qui existe bel et bien sous cette forme dans un nombre croissant de pays.

Cela étant, les modalités d'intégration du facteur économique reste variable : elle peut être directe ou indirecte, explicite ou implicite. L'évaluation du NICE inclut explicitement le coût des traitements comme critère du jugement ; ce n'est pas le cas de la Commission de la Transparence en France, mais la préoccupation économique n'en est pas absente, d'autant que ses décisions ont une implication économique immédiate, à savoir le remboursement des produits. Un point important est le statut des études médico-économiques, ou pharmaco-économiques, qui sont largement produites dans ce cadre, même si elles ne sont reconnues officiellement et formellement que par très peu de pays (le Canada et l'Australie font plutôt figure d'exception).

Paradoxalement, cette évaluation post-AMM, qui tend à se développer rapidement, a été relativement peu étudiée en tant que telle. L'objectif du présent article est de contribuer à une réflexion méthodologique sur ce thème, à travers l'analyse d'un exemple concret qui tend à faire école dans certains pays d'Europe, à savoir l'évaluation des conditions d'accès au remboursement en France, après leur modification en 1999. Notre objectif est d'analyser les concepts et méthodes de cette évaluation. Si la modification des règles d'évaluation est récente, le gouvernement a décidé de les appliquer rétrospectivement à l'ensemble des médicaments remboursables, pour savoir s'ils répondent tous aux nouveaux critères de prise en charge. Une sortie du remboursement est envisagée, pour les produits qui ne répondraient pas aux critères.

L'avantage de cette « réévaluation » est de nous permettre de disposer d'un large échantillon de produits évalués à une même époque selon une même méthode, permettant donc une analyse statistique du processus d'évaluation post-AMM.

L'article est organisé comme suit : nous décrivons d'abord le système français d'évaluation post-AMM pour le remboursement des médicaments, nouveaux et anciens. Puis nous rappellerons la manière dont les critères ont été appliqués rétrospectivement à tout le stock de médicaments existant et nous donnerons les résultats de cette « réévaluation ». Nous présenterons ensuite une analyse statistique de ces données, en utilisant notamment des modèles de régressions logistiques multinomiaux multivariés. Nous en tirerons, pour terminer, quelques conclusions et quelques réflexions pour les méthodes d'évaluation post-AMM.

⁵ Scrip n°2617, Février 2001

II.1 Le système d'évaluation post-AMM des médicaments remboursables

Le système français de remboursement et de tarification des médicaments a été décrit à plusieurs reprises [4, 5]. Il s'agit d'un système de « liste positive » sur laquelle les médicaments admis au remboursement sont inscrits après avis d'une commission spécialisée, la Commission de la Transparence, rattachée à l'agence française d'enregistrement (l'AFSSAPS). Cette commission est composée d'un président, d'un vice-président et de quatorze membres, qui sont des représentants de l'administration, des caisses d'assurance maladie et de l'Ordre des Médecins. Cinq personnalités « choisies en raison de leur compétence médicale, scientifique ou économique dans le domaine du médicament » la complètent. Les travaux d'évaluation sont confiés à des experts extérieurs non membres, qui élaborent des rapports sur la base desquels la Commission statue.

Selon les textes, cette commission est chargée de donner un avis sur « l'intérêt » des médicaments sollicitant leur remboursement. Elle est également chargée d'évaluer « les coûts comparés des médicaments à même visée thérapeutique ». Les avis de la Commission sont publics et leur diffusion auprès des médecins prescripteurs est prévue. Une procédure d'audition est possible en cas de contestation des termes de l'avis par les firmes⁶.

L'avis de la Commission est décisif pour un médicament car, outre le remboursement, il en détermine également dans une large mesure, le prix de marché. C'est, en effet, en grande partie sur la base de cet avis que l'Etat fixe les prix de vente au public, à travers une instance spécialisée inter-ministérielle, le Comité Economique des Produits de Santé (CEPS).

II.2 Un nouveau critère d'évaluation : le Service Médical Rendu

La procédure d'inscription sur la liste positive des médicaments remboursables a récemment été modifiée par le décret 99-915 du 27 octobre 1999. Selon ce texte, « les médicaments dont le service médical rendu est insuffisant au regard des autres médicaments ou thérapies disponibles ne sont pas inscrites sur la liste ». Ce décret a donc introduit un nouveau critère pour l'évaluation post-AMM : le Service Médical Rendu (SMR)⁷. Ce critère n'est pas défini directement par le décret. Celui-ci donne cependant des indications pour son attribution : « Cette appréciation prend en compte l'efficacité et les effets indésirables du médicament, sa

⁶ L'inscription doit être sollicitée par les firmes. Il leur revient de décider de l'opportunité de demander le remboursement de leurs produits, sachant que dans ce cas, leur prix sera réglementé. C'est pour éviter cette contrainte que certaines firmes préfèrent ne pas demander de remboursement pour des produits de prescription. C'est notamment le cas des produits dits « Life-style » comme le Viagra® ou le Xenical®.

⁷ Nous ne nous attardons pas sur l'historique du système et sur le fait que le Critère du Service Médical Rendu remplace un autre mécanisme plus complexe. Dans le cadre de la réglementation antérieure, la Commission de la transparence évaluait non pas le niveau du Service Médical Rendu mais son amélioration par rapport aux produits de référence. Le médicament pouvait être remboursé : i) s'il apportait une amélioration de service médical rendu (ASMR) ; ii) dans le cas contraire, s'il entraînait une économie dans coût de traitement médicamenteux. Dans le nouveau système, c'est le niveau absolu du service médical rendu qui est évalué et non son amélioration. Par ailleurs, un produit de SMR insuffisant ne peut être remboursé, même s'il entraîne une économie de coût de traitement médicamenteux.

place dans la stratégie thérapeutique, notamment au regard des autres traitements disponibles, la gravité de l'affection à laquelle il est destiné, le caractère curatif symptomatique du traitement médicamenteux et son intérêt pour la santé publique. » Concrètement, les firmes demandant le remboursement de leurs spécialités doivent donc soumettre à la Commission un dossier comportant les éléments de preuve permettant d'étayer et de quantifier ces différentes dimensions du SMR. Le dossier comporte une large partie comparative où le nouveau médicament est situé par rapport aux autres stratégies thérapeutiques, médicamenteuses ou non. L'Amélioration du Service Médical Rendu (ASMR), c'est-à-dire la différence de SMR entre le nouveau médicament et le traitement de référence, si elle existe, est également documentée.

II.3 La réévaluation des produits existants

La procédure d'évaluation du SMR décrite précédemment concerne prioritairement les produits nouveaux. Le gouvernement français a pris cependant la décision, fin 1998, de faire procéder, dans les trois années suivantes, à une évaluation rétrospective de l'ensemble des médicaments remboursables sur la base de ce même critère de Service Médical Rendu⁸. Cette opération a conduit la Commission de Transparence à mettre au point une méthodologie d'évaluation qui a été explicitée dans de courtes notes internes [4, 5]. C'est pour nous ainsi l'occasion de mieux comprendre la logique de cette évaluation.

La procédure suivie par la Commission comporte quatre phases. Les produits sont d'abord regroupés par indication. Ces regroupements sont transmis aux firmes concernées pour recueillir leurs avis et éviter les erreurs ou omissions. Les groupes d'experts attribuent ensuite à chaque produit un niveau de SMR selon l'échelle décrite ci-dessous. La Commission de Transparence valide ce premier classement qui est transmis pour informations aux entreprises concernées. Celles-ci peuvent demander un réexamen de leur produit en cas de désaccord. La dernière étape est l'adoption définitive du classement, après réexamen du produit éventuel et audition de la firme.

La réévaluation a été menée en plusieurs tranches. La première, qui concernait l'évaluation des produits cardio-vasculaires, de rhumatologie, des troubles de métabolisme et de psychiatrie, est terminée. Une deuxième tranche, avec les médicaments de l'appareil respiratoire, de l'appareil hépato-gastro-entérologie, les antibiotiques et les antiviraux ainsi que les vaccins est en cours d'achèvement. Une troisième et dernière tranche couvrant l'anesthésie – réanimation, la cancérologie, l'hématologie, la stomatologie, l'ophtalmologie, l'urologie, la néphrologie et les corticoïdes a été lancée courant 2000. Ce sont les résultats complets de la première tranche qui porte sur plus de 1400 produits et présentations que nous analyserons ici.

II.4 La méthodologie de la réévaluation du SMR

La commission a d'abord distingué 3 niveaux de SMR : « *Majeur* ou *important* », « *Modéré* », et « *Faible* ». Une quatrième mention « *Pas de SMR attribué* » est portée dans les cas où la Commission juge ne pas disposer des éléments scientifiques nécessaires à l'évaluation. Il s'agit en général de produits anciens pour lesquels il n'existe pas de preuves cliniques d'efficacité et dont le SMR est nul ou faible.

⁸ Cette intention figure dans les attendus de la loi de financement pour la sécurité sociale pour 1999, du 23 décembre 1998

Certains produits possèdent plusieurs indications. Dans ce cas, la commission a attribué un niveau de SMR pour chacune d'entre elles. Le niveau de SMR d'un produit correspond au niveau le plus élevé, à condition toutefois que l'indication correspondante représente une part « substantielle » des volumes prescrits. Si ce n'est pas le cas, le niveau le plus élevé correspondant à l'indication la plus importante sera retenu.

Pour attribuer un niveau de SMR à chaque produit, la Commission a examiné les différentes « dimensions » de celui-ci, dans l'esprit des critères cités dans le décret de 1999. Le tableau 1 explicite ces différentes dimensions et les modalités de cotation. La gravité de la pathologie est cotée sur une échelle à trois niveaux selon la sévérité du risque vital ou le niveau de handicap. L'efficacité-tolérance du produit est codée sur une échelle classique à 4 degrés. Le critère « *place dans la stratégie thérapeutique* » distingue les utilisations en 1^{ère} ou en 2^{ème} intention, comme traitement d'appoint ou comme traitement de recours. Une mention « *pas de place dans la stratégie thérapeutique* » est affectée aux médicaments pour lesquels la Commission s'interroge quant à leur utilité pratique. La caractéristique « *intérêt pour la santé publique* » et l'existence d'alternatives thérapeutiques sont codées de manière dichotomique (oui/non), sans autre explication. La même classification s'applique aux « *conditions d'utilisation* », le « *oui* » correspondant aux cas d'indications multiples.

Tableau 1 : la liste et la cotation des dimensions du SMR

Critères du décret	Modalités de codage par la CT
Gravité de la pathologie	A : « <i>l'affection engage le pronostic vital du patient immédiatement ou par suite de complications</i> » B : « <i>l'affection se caractérise par une évolution vers un handicap et/ou une dégradation marquée de la qualité de vie</i> » C : « <i>l'affection n'engage ni le pronostic vital du patient, n'entraîne pas de complications graves, ni de handicap, ni de dégradations marquées de qualité de vie</i> »
Efficacité-sécurité	efficacité faible efficacité modeste efficacité moyenne efficacité importante
Nature du traitement	Symptomatique Préventif Curatif Préventif et curatif
Place dans la stratégie thérapeutique	Traitement de 1 ^{ère} intention Traitement de 2 ^{ème} intention Traitement de recours Traitement d'appoint Place limitée dans la stratégie thérapeutique Pas de place dans la stratégie thérapeutique
Existence d'alternatives thérapeutiques	Oui Non
Intérêt en termes de santé publique	Oui Non
Conditions d'utilisation	Plusieurs indications (dans ce cas prise en compte de l'indication prévalente) Une seule indication

On notera que certaines dimensions (« *gravité de pathologie* », « *efficacité-sécurité* ») sont des variables « ordinales » en ce sens qu'elles expriment un gradient de qualité ou de gravité. Les autres sont plutôt des variables catégorielles ou polytomiques qui n'expriment pas un ordre préférentiel. La variable « *santé publique* » est mal définie. Sa place en tant que critère indépendant est d'ailleurs discutable puisqu'en toute logique, l'intérêt pour la santé publique est une notion synthétique qui devrait résulter des autres caractéristiques des produits. Si l'intérêt pour la santé publique constituait réellement un critère indépendant, on devrait pouvoir imaginer un médicament efficace dans une pathologie grave n'offrant cependant pas d'intérêt pour la santé publique ! C'est évidemment absurde, même si nous en trouverons des exemples !

II.5 Les résultats de la réévaluation

a) Le SMR

Le tableau 2 présente le résultat final de l'évaluation pour les 1453 produits et présentations de la première tranche.

Tableau 2 : La répartition des produits des différentes classes en niveau de SMR

	Cardiologie	Métabolisme	Psychiatrie	Rhumatologie	Ensemble
n	812	163	223	255	1453
Pas de SMR	28,0%	18,0%	20,5%	13,0%	23,0%
faible	1,5%	0,5%	8,5%	30,0%	7,5%
modéré	3,5%	3,5%	10,5%	9,0%	6,0%
important	66,0%	78,0%	60,5%	48,0%	63,0%
majeur	1,0%				

On note que la plupart d'entre eux appartiennent aux deux grands groupes extrêmes. 23% d'entre eux ont été classés « *sans SMR attribué* », ce qui signifie en réalité un niveau de SMR « *insuffisant* » pour justifier le remboursement. Ce pourcentage varie de 28% dans le domaine cardio-vasculaire à 13% en rhumatologie. Dans le domaine cardio-vasculaire il s'agit essentiellement de médicaments utilisés dans :

- L'insuffisance veineuse et veino-lymphatique (diosmine, genko-biloba, etc.).
- Le traitement d'appoint des déficits cognitif et neuro-sensoriel du sujet âgé (pirocétam, ribasine, etc.)
- Le traitement d'appoint ou symptomatique de la claudication intermittente des artériopathies des membres inférieurs (genko-biloba, nicergoline, etc.).

Dans les troubles du métabolisme et de la nutrition, on trouve essentiellement les traitements des carences magnésiennes, et, en rhumatologie, les traitements topiques des douleurs musculaires ou ligamentaires. En psychiatrie, ce sont principalement des traitements de phytothérapie dans les manifestations mineures de l'anxiété et dans les troubles légers du sommeil, ainsi que quelques médicaments utilisés dans des syndromes mal définis comme la « *spasmophilie* ».

Pour le reste, les deux tiers des produits et présentations ont un niveau de SMR « *important* ». Le niveau « *majeur* » a été attribué à des hypocholestérolémiants comme la pravastatine ou la

simvastatine (qui existent en France sous plusieurs noms de marque en raison des co-marketings) pour les distinguer des autres hypolipidémisants dont les bénéfices n'ont pas été prouvés par de grands essais de prévention primaire et secondaires des événements cardiovasculaires majeurs (décès cardio-vasculaire, infarctus aigu du myocarde, procédures de revascularisation).

b) Les dimensions du SMR

Ce classement résulte de l'analyse des dimensions du SMR. Le tableau 3 ci-dessous donne les classements des produits en ce qui concerne l'efficacité et la gravité des pathologies traitées :

Tableau 3 : Le classement sur les différentes dimensions du SMR

Efficacité	N	%
Faible	110	7,4%
Modeste	341	23,0%
Moyenne	275	18,5%
Importante	725	48,9%

Pathologie	N	%
A – risque vital	739	49,8%
B – risque moyen	422	28,5%
C – bénin	291	19,6%

Nature du traitement	N	%
symptomatique	711	47,9%
curatif	202	13,6%
préventif	509	34,3%
curatif et/ou préventif	29	2,0%

Place stratégique thérapeutique	N	%
Trt 1 ^{ère} intention	765	51,6%
Trt 2 ^{ème} intention	207	14,0%
Trt appoint	328	22,1%
Trt recours	37	2,5%
Pas de place	109	7,3%
Place limitée	6	0,4%

Intérêt santé publique	N	%
non	516	35,5
oui	937	65,5

Alternatives	N	%
possible	1329	89,6%
impossible	123	8,3%

Ainsi, selon la grille de la Commission, la moitié des produits et présentations étudiés ont une efficacité « importante » ; dans une moitié de cas, également, leur indication correspond à une pathologie grave, mettant en jeu le pronostic vital. Mais, dans 48% des cas, il s'agit de traitement symptomatique, moins de 16% ayant été jugés comme possédant une capacité curative. Pour 90% des médicaments, il existe, selon la Commission, des alternatives thérapeutiques et dans 64,5% des cas seulement, le produit a un « intérêt pour la santé publique ».

III.1 Objectifs

La question que suscitent immédiatement ces données est celle de la relation existant entre le niveau du SMR et le jugement porté par la Commission sur chacune des « dimensions ». Les dimensions ont-elles toutes le même poids ? Certaines sont-elles plus déterminantes que d'autres et lesquelles ? Le système de pondération est-il stable d'une classe thérapeutique à l'autre ou varie-t-il selon les familles de produits ?

C'est pour répondre à cette question que nous avons bâti un modèle d'analyse statistique des données de la réévaluation qui poursuit deux objectifs complémentaires. Le premier consiste à comprendre la logique de la Commission et la manière dont elle apprécie les critères du SMR. La connaissance des poids des critères peut permettre ensuite un modèle prédictif de la probabilité d'obtenir un SMR d'un niveau donné en fonction des caractéristiques du produit. Par exemple, estimer la probabilité d'obtenir un SMR de niveau "important" pour un produit d'efficacité-sécurité modeste, indiqué dans une pathologie dont les complications sont susceptibles d'engager le pronostic vital, qui est prescrit en première intention, qui n'a qu'une seule indication et qui présente un intérêt manifeste pour la santé publique.

III.2 La méthode

L'analyse a été menée en deux temps.

1. On a d'abord mené des analyses simples (« univariées ») confrontant le niveau de SMR à chacune des « dimensions » distinguées précédemment. Ces analyses ont été effectuées classiquement sous la forme de tableaux à double entrée, la dépendance entre l'appréciation sur la dimension considérée et le niveau de SMR étant appréciée à l'aide d'un test classique du Chi-2. Ces analyses ont l'intérêt de livrer une analyse fine du processus et des redondances possibles entre dimensions ou entre modalités de ces dernières.
2. A partir des enseignements des analyses simples, on a mené, dans un second temps, des analyses de régression multiples (« multivariées ») de manière à quantifier la relation existante entre la variable "expliquée" (le niveau de SMR) et des variables "explicatives", sélectionnées au terme de la phase précédente. La variable expliquée étant une variable catégorielle multinomiale ordinaire (le niveau de SMR en 3 ou 4 classes), nous avons utilisé un modèle de régression logistique multinomial cumulatif à hasard proportionnel ("*proportional odds model*"), adapté à ce type de données. Ce type de modèle permet d'estimer la probabilité d'un niveau donné de SMR, conditionnellement à la valeur obtenue sur chacune des variables explicatives, elles-mêmes exprimées sous forme de variables catégorielles. Un bon exposé de ces méthodes figure dans Agresti [6] ou Andersen [7]. L'hypothèse de hasard proportionnel a été testée systématiquement avant l'application des modèles. Les calculs ont été effectués à l'aide de la procédure LOGIT du progiciel statistique SASTM release 8.1 en environnement Windows 2000 [8].

III.3 Les données

Les données utilisées ont été présentées plus haut. Pour l'analyse, un certain nombre de reclassements ont du être opérés.

On a d'abord regroupé les niveaux « faible » et « modéré » du SMR, ainsi que les niveaux « important » et « majeur ». Ces regroupements se justifient sur la base des effectifs donnés au tableau 2. Les niveaux regroupés correspondent à des effectifs relativement restreints. Par ailleurs, le niveau « faible » a été attribué à des produits que la Commission souhaitait maintenir dans le périmètre du remboursable, quitte à en abaisser le taux de prise en charge, et on ne pouvait donc pas les assimiler au niveau « pas de SMR » : l'implication logique de l'attribution de ce niveau est en effet la sortie du périmètre du remboursable.

On dispose ainsi de trois niveaux de SMR correspondant à trois niveaux (théoriques) de prise en charge :

- SMR « insuffisant » (23%) : pas de remboursement,
- SMR « moyen » (13,5%) : remboursement avec large co-paiement (65%),
- SMR « important » (63,5%) : remboursement avec co-paiement faible ou nul (35% ou 0%).

En ce qui concerne les « dimensions », on a regroupé les niveaux d'efficacité « faible » (7,4%) et « modeste » (23%), également pour des raisons d'effectifs (cf. tableau 3).

De même, en ce qui concerne l'objectif du traitement, on a regroupé les items « *traitement curatif et préventif* » (2% des produits) et « *traitement curatif* » (13,6%). On a fait prévaloir le caractère curatif sur le caractère préventif, dans la mesure où il semble davantage valorisé et où il concerne moins de produits.

Le critère « *existence d'alternatives thérapeutiques* » n'a pas été retenu dans l'analyse : 90% des produits ayant été reconnus comme possédant une alternative, ce critère n'est pas discriminant.

Enfin, s'agissant de la « place dans la stratégie thérapeutique », on a considéré :

- Que les « *traitements de recours* » (2,5%) étaient des « *traitements de seconde intention* » (14%) ;
- Que les traitements classés comme « *ayant une place limité* » (0,4%) étaient assimilables à des « *traitements d'appoint* » (22,1%).

IV.1 Les analyses simples

La relation efficacité-SMR est décrite au tableau 4.

Tableau 4 : Niveau de SMR et Efficacité(*)

SMR	Insuffisant	Moyen	Important	total
Efficacité				
Faible	336	112	4	452
Moyenne	0	67	208	275
Importante	0	12	713	725
Total	336	191	925	1452

P (Chi-1)<0,0001

Deux-tiers des produits avec une efficacité faible ont un SMR faible. Mais quatre d'entre eux ont un SMR important. Il s'agit :

- Du gluconate de calcium dans la calcithérapie d'urgence (tétanie hypocalcémique),
- Du phosphoneuros solution buvable, dans le rachitisme hypophosphorémique et dans le diabète phosphaté héréditaire ou acquis,
- Du Didronel® dans le traitement de l'ostéoporose post-ménopausique avec au moins un tassement vertébral.

Dans les trois premiers cas, les traitements ont été jugés par ailleurs comme des traitements « symptomatiques », « d'appoint » et « sans intérêt pour la santé publique ». C'est donc la gravité de la pathologie (« importante ») qui est sans doute à l'origine du classement. Dans le cas de Didronel®, en revanche, la pathologie est classée de gravité moyenne (« B ») mais le traitement est curatif et son intérêt pour la santé publique est reconnu.

L'impact de la gravité est donné par le tableau 5.

Tableau 5 : Niveau de SMR et gravité de la pathologie (*)

SMR	Insuffisant	Moyen	Important	total
Sévérité				
Sévère	51	48	660	739
Moyenne	54	105	263	422
Bénigne	231	38	22	291
Total	336	191	925	1452

P (Chi-1)<0,0001

87% des traitements indiqués dans une pathologie sévère sont classés comme ayant un SMR important. A l'inverse, 95% des traitements indiqués dans des pathologies légères se voient attribuer un SMR « insuffisant » ou « moyen ».

En ce qui concerne l'effet des traitements, la Commission a attribué une « prime » aux produits et présentations considérés comme « curatifs » (tableau 6). Inversement le caractère « symptomatique » ne constitue pas un facteur favorable pour l'obtention des niveaux élevés de SMR. On observe une « décote » pour les traitements préventifs : seuls 81% d'entre eux se voient attribuer un SMR majeur contre 92,5% des traitements curatifs.

Tableau 6 : Niveau de SMR et effet des traitements (*)

SMR	Insuffisant	Moyen	Important	total
Effet				
Curatif	4	11	216	231
Préventif	24	73	412	509
Symptomatique	307	107	297	711
Total	336	191	925	1451

P (Chi-1)<0,0001

Le tableau 7 montre une nette dichotomie entre les traitements auxquels une place dans les stratégies thérapeutiques est reconnue, qu'il s'agisse d'un traitement de 1^{ère} ou de 2^{ème} intention, et ceux qui n'ont que des traitements d'appoint ou des traitements pour lesquels la Commission s'interroge sur leur place.

Tableau 7 : Niveau de SMR et place dans la stratégie thérapeutique (*)

SMR	Insuffisant	Moyen	Important	total
Place				
1^{ère} intention	0	46	719	765
2^{ème} intention	0	41	203	244
Trt. appoint	239	92	3	334
Pas de place	97	12	0	109
Total	336	191	925	1452

P (Chi-1)<0,0001

On note enfin curieusement que 61 produits et présentations auxquels la Commission n'a pas reconnu d'intérêt pour la santé publique reçoivent néanmoins un SMR important (tableau 8). Deux grands groupes de produits sont dans ce cas. D'abord les corticothérapies locales dans certaines indications comme :

- L'arthrose en poussée ;
- Les talalgies, le syndrome du canal carpien ou la maladie de Dupuytren ;
- Les tendinites et bursites.

Dans l'arthrite inflammatoire, en revanche, l'intérêt pour la santé publique est reconnu.

Tableau 8 : Niveau de SMR et place dans la stratégie thérapeutique (*)

SMR	Insuffisant	Moyen	Important	total
Intérêt Santé pub.				
Oui	336	119	61	516
Non	0	72	865	937
Total	336	191	925	1453

P (Chi-1)<0,0001

Pour cette classe de médicaments, on est en droit de se demander si le jugement sur la santé publique est bien indépendant de celui sur les autres critères. En effet, dans tous les cas où l'intérêt pour la santé publique n'est pas reconnu, l'efficacité est notée comme « moyenne » comme d'ailleurs la gravité de la pathologie. En revanche, l'efficacité est « importante » et la pathologie est « grave » dans l'arthrite inflammatoire

L'autre exemple d'incohérence apparente est celui de bêtabloquants dans la récurrence des tachycardies ventriculaires menaçant le pronostic vital : la pathologie est grave ; l'efficacité des traitements est reconnue comme « importante » et pourtant, l'intérêt en termes de santé publique n'est pas reconnu !

IV.2 les analyses multiples

L'objet de l'analyse présentée ici est de tenter de dégager un système implicite de pondération des critères, en les faisant entrer tous dans l'équation de régression logistique multinomiale et multivariée qui a été décrite plus haut.

a) Les variables

Les variables retenues pour l'analyse sont les suivantes :

- L'efficacité du produit en trois modalités (faible, moyenne, importante)
- La gravité de la pathologie en trois modalités (bénigne, moyenne, sévère)
- L'effet du traitement en deux modalités (curatif ou préventif vs. symptomatique)
- La place dans la stratégie thérapeutique en deux modalités (1^{ère} ou 2^{ème} intention contre « traitement d'appoint » ou « pas de place »)
- L'intérêt pour la santé publique en deux modalités (oui, non).

Comme indiqué précédemment, « l'existence d'alternatives thérapeutiques », qui figurent parmi les critères du décret sur le SMR, n'est pas discriminant et il n'est donc pas pris en compte ici.

Le choix des modalités se justifie sur la base de l'analyse précédente. Il s'agit des modalités qui semblent être déterminantes pour l'octroi des niveaux de SMR.

b) le modèle général

Le modèle général porte sur l'ensemble de la base de données, toutes classes confondues. Il sera complété par des modèles par classes. Le modèle donne la probabilité qu'un médicament reçoive un niveau déterminé de SMR, compte tenu de ses caractéristiques.

Les variables étant des variables catégorielles, il est nécessaire de définir un « cas de référence » par rapport auquel sera évalué l'effet des différents facteurs. Le cas de référence est celui d'un médicament d'efficacité faible, indiqué dans une pathologie bénigne, dont

l'effet est symptomatique, qui occupe une place d'appoint ou qui n'a pas de place dans la stratégie thérapeutique et qui n'a pas d'intérêt pour la santé publique. On calculera les probabilités qu'un tel médicament obtienne un SMR important, moyen ou faible et on étudiera l'effet de la modification de ses caractéristiques sur ces probabilités.

Les résultats des estimations figurent au tableau 9 ci-dessous. Cinq spécifications de l'équation de régression ont été testées. La première correspond au modèle complet avec toutes ces variables. La spécification 2 est la même que la précédente à l'exception de la variable d'effet qui n'est pas significative. Les deux spécifications suivantes retiennent les variables de gravité de la pathologie et d'efficacité du traitement et ajoutent respectivement la variable « place dans la stratégie thérapeutique » (spécification 3) et « intérêt pour la santé publique » (spécification 4). La dernière n'inclut que les variables d'efficacité et de gravité de la pathologie qui sont les seules dont on peut garantir qu'elles sont indépendantes l'une de l'autre.

Tableau 9 : Les estimations des équations de régression logistiques multinomiales

Variables	Spécification 1		Spécification 2		Spécification 3		Spécification 4		Spécification 5	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
Constante 1	-7,507	<,0001	-7,689	<,0001	-7,752	<,0001	-6.841	<,0001	-6.761	<,0001
Constante 2	-2,058	<,0001	-2,243	<,0001	-2,383	<,0001	-2.220	<,0001	-2.366	<,0001
Grav=B	1,865	<,0001	1,917	<,0001	2,135	<,0001	2.083	<,0001	2.318	<,0001
Grav=A	1,016	0,0008	1,100	0,0001	1,900	<,0001	1.252	<,0001	2.097	<,0001
Ef=moyen	2,326	0,0006	2,346	0,0005	2,840	<,0001	5.252	<,0001	6.052	<,0001
Ef=import	5,442	<,0001	5,511	<,0001	6,105	<,0001	7.877	<,0001	8.773	<,0001
Effet	-0,187	0,405	-	-	-	-	-	-	-	-
Stratégie	3,991	<,0001	3,987	<,0001	4,357	<,0001	-	-	-	-
Santé Pub,	1,861	<,0001	1,854	<,0001	-	-	2.016	<,0001	-	-
Test hasard proport,	<,0001		<,0001		<,0001		<,0001		<,0001	
Test Wald	<,0001		<,0001		<,0001		<,0001		<,0001	
AIC	774		772		811		822		870	
Pourcentage concordance	96,4%		96,4%		94,7%		96,0%		94,7%	
SD	0,950		0,952		0,932		0,945		0,927	

Nota : le cas de référence correspond à un médicament d'efficacité faible, indiqué dans une pathologie bénigne, dont l'effet est symptomatique, qui occupe une place d'appoint ou qui n'a pas de place dans la stratégie thérapeutique et qui n'a pas d'intérêt pour la santé publique. Le test de hasard proportionnel valide l'utilisation de la méthode d'estimation. Le test de Wald est un test de significativité globale des coefficients (vs l'hypothèse nulle, =0 pour toutes les variables). La statistique d'Akaike (Akaike Information Criteria) est une mesure de la qualité des modèles, les valeurs les plus faibles étant préférables. Le pourcentage de concordance et le D de Sommer (SD) sont des tests de la capacité prédictive du modèle. Pour tout couple de produits classés différemment sur l'échelle du SMR, on compare le classement effectif à celui suggéré par les probabilités déterminées par le modèle. Il y a concordance quand les deux classements coïncident et discordance dans le cas contraire. Le test de Sommer est un test classique de corrélation des rangs. La capacité prédictive du modèle est parfaite quand il vaut 1.

Ces résultats suggèrent les conclusions suivantes :

- Toutes les estimations sont hautement significatives avec des capacités prédictives très élevées. Cela constitue une validation du choix de la méthode.
- La variable « effet du traitement » n'est pas significative : ce critère n'a donc aucune influence sur le niveau de SMR obtenu par les produits.
- Le modèle le plus « court » ne comportant que les variables « efficacité du produit » et « gravité de la pathologie » obtient une des valeurs prédictives les plus élevées. Ces deux variables suffisent donc à expliquer à elles seules plus de 90% du classement.
- L'adjonction de variables supplémentaires, notamment la « place dans la stratégie thérapeutique » et « l'intérêt pour la santé publique », n'améliore pas ou améliore très faiblement la valeur prédictive du modèle.
- Le « meilleur » modèle est néanmoins le modèle 2 qui, outre l'efficacité et la gravité, comporte également les variables de « place dans la stratégie thérapeutique » et d'« intérêt pour la santé publique ». C'est donc lui qui sera utilisé dans la suite de l'exposé.

Pour ce modèle, il existe 72 combinaisons différentes de modalités des variables explicatives. Pour chacune d'entre elles, il permet de calculer les probabilités d'un niveau donné de SMR. En fait seules 20 combinaisons existent réellement. Les autres sont virtuelles. Par exemple, la combinaison « produit efficace », « indication grave », « pas de place dans la stratégie thérapeutique » et « pas d'intérêt pour la santé publique » n'existe pas ! Plus de 97% des produits et présentations se retrouvent dans 12 combinaisons seulement. Le tableau 10 indique les probabilités (et les effectifs) qui correspondent.

Tableau 10 : Les probabilités de niveau de SMR pour les combinaisons de critères les plus fréquentes

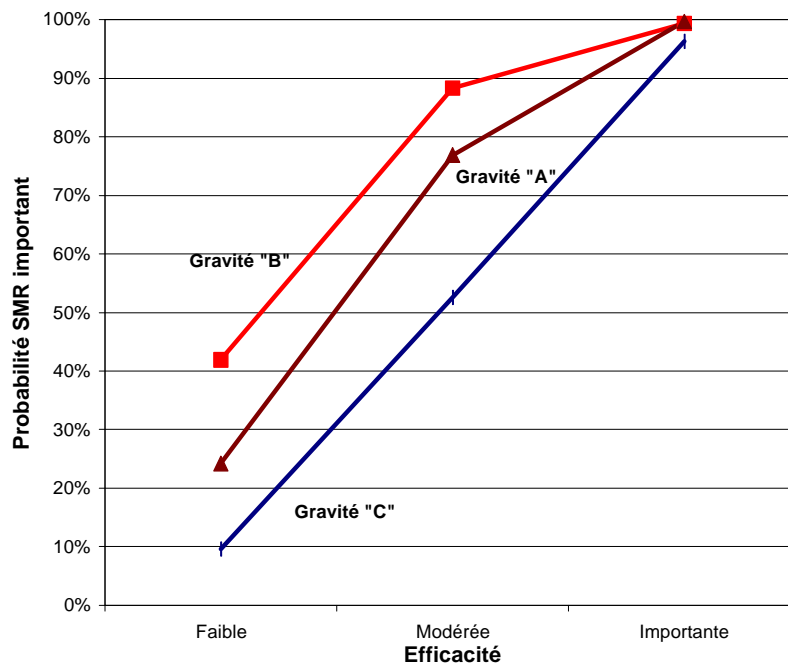
Critères du SMR				Effectifs	Niveau de SMR		
Gravité	Efficacité	Place dans la stratégie	Intérêt Santé Pub.		Faible	Moyen	Important
A	Forte	Oui	Oui	497	0,0%	0,8%	99,2%
C	Faible	Non	Non	237	90,4%	9,5%	0,0%
B	Forte	Oui	Oui	197	0,0%	0,4%	99,6%
A	Moyenne	Oui	Oui	147	0,1%	16,7%	83,2%
B	Faible	Non	Non	121	58,1%	41,6%	0,3%
B	Moyenne	Oui	Oui	68	0,1%	23,6%	76,3%
A	Faible	Non	Non	54	75,8%	24,0%	0,1%
C	Moyenne	Oui	Non	31	1,6%	77,9%	20,5%
B	Moyenne	Oui	Non	23	0,2%	36,1%	63,7%
A	Forte	Oui	Non	18	0,0%	5,1%	94,8%
A	Faible	Non	Oui	15	5,5%	87,6%	6,9%
C	Faible	Non	Oui	10	59,6%	40,1%	0,3%
				1418			

Un produit d'efficacité forte, indiqué dans une pathologie grave, ayant une place dans la stratégie thérapeutique et ayant un intérêt de santé publique, a pratiquement 100% de chances de se voir reconnaître un SMR important. Ce type de produit représente 34% des effectifs. A l'inverse, le produit « minimal » (efficacité faible, pathologie bénigne, place limitée ou nulle dans la stratégie thérapeutique et sans intérêt pour la santé publique), qui représente 16% des effectifs de la base, a une probabilité pratiquement nulle d'avoir un SMR important et une probabilité de plus de 90% d'avoir un SMR faible.

L'analyse des chiffres montre :

- Ni l'efficacité semble plus « rentable » que la gravité de la pathologie en termes de SMR élevé ;
- Que pour une efficacité reconnue et un niveau de gravité de pathologie au moins égal à B, les facteurs « place dans la stratégie thérapeutique » et « intérêt dans la santé publique » augmentent très sensiblement la probabilité de SMR important.

Figure 1 : Effet de l'efficacité du traitement et de la gravité de la pathologie sur la probabilité d'un SMR important



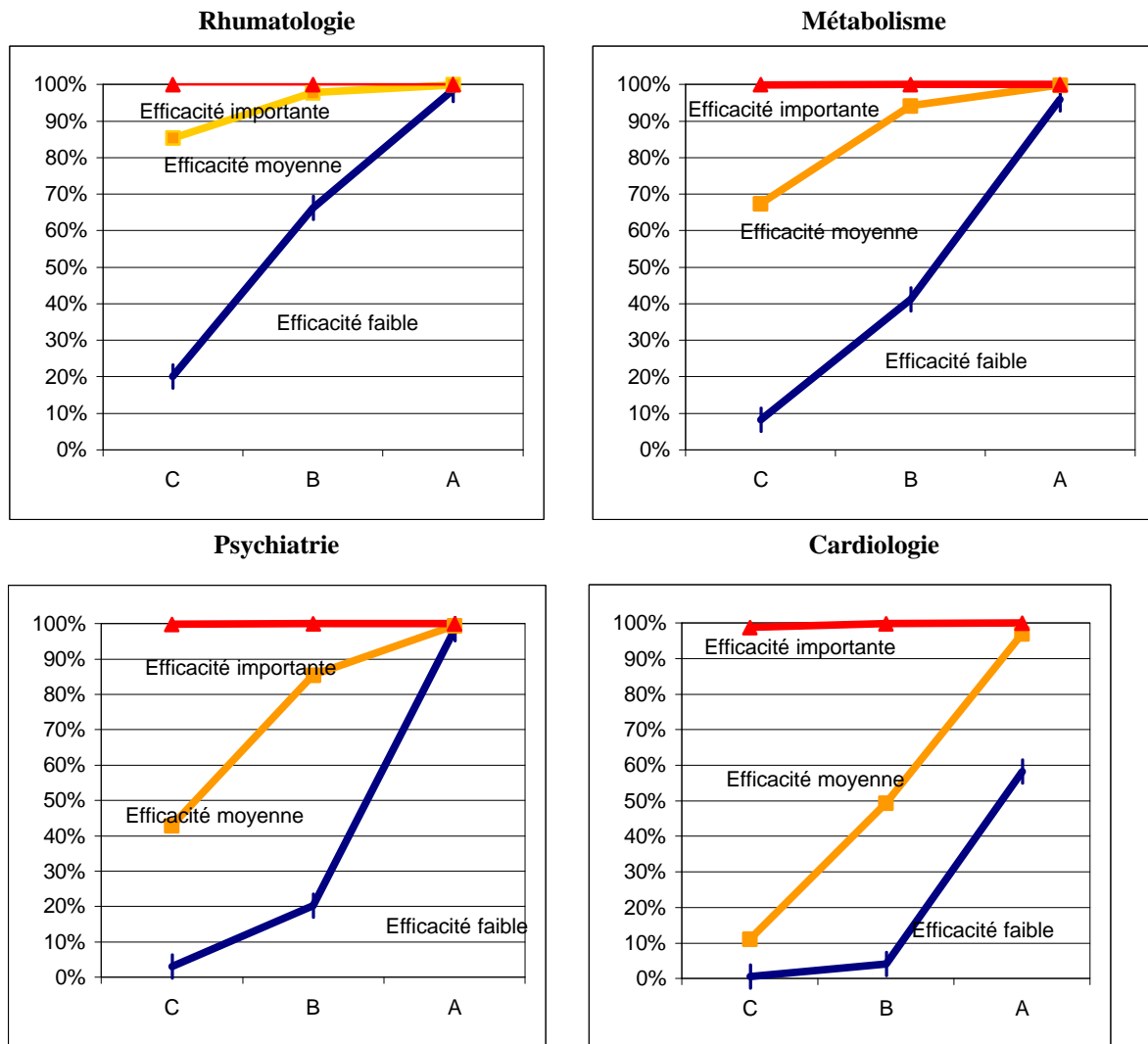
Nota : calculs effectués sur la base de la spécification 2 pour un produit ayant une place reconnue dans la stratégie thérapeutique et un intérêt pour la santé publique

Le graphique 1 illustre ce premier point. Un médicament d'efficacité forte dans une pathologie bénigne a sensiblement plus de chance d'obtenir un SMR important qu'un médicament d'efficacité faible dans une pathologie sévère.

c) Les modèles par classes

La figure 2 présente la probabilité d'obtenir un SMR important en fonction de l'efficacité et de la gravité pour chacune des 4 classes thérapeutiques analysées.

Figure 2 : les profils par classes thérapeutiques



Nota : Calculs effectués sur la base de la spécification 2 pour un produit ayant une place reconnue dans la stratégie thérapeutique et un intérêt pour la santé publique. En ordonnée, probabilité d'obtenir un SMR « important ». Gravité de la pathologie côté de C (pathologie bénigne) à A (pathologie sévère).

On observe :

- Que probabilité d'obtenir un SMR élevé pour un produit jugé d'efficacité importante est très forte dans toutes les classes ;
- En revanche, la probabilité pour les produits d'efficacité moyenne d'obtenir un SMR important diffère selon les classes : elle est toujours très élevée en rhumatologie, alors qu'elle en dépend fortement du niveau de gravité de la pathologie dans les autres classes. Ainsi un produit d'efficacité moyenne dans une pathologie bénigne a t'il plus de 80% de

« chances » d'obtenir un SMR élevé en rhumatologie contre 45% environ en psychiatrie et à peine plus de 10% en cardiologie.

- Enfin, un produit d'efficacité faible dans une indication sévère a près de 100% de chances d'obtenir un SMR important dans toutes les classes sauf en cardiologie où cette probabilité ne dépasse pas 60% !

Une leçon de cette analyse est donc que le processus d'évaluation ne semble pas complètement homogène d'une classe à l'autre.

c) La validité prédictive du modèle

Enfin, nous avons confronté les probabilités d'obtenir un niveau de SMR avec le niveau de SMR effectivement obtenu. Nous rapporterons ici les résultats pour le niveau « SMR important ou majeur » qui sont représentés dans le tableau 11 et le graphique 3 ci-dessous.

Tableau 11 : Valeur prédictive du modèle pour le niveau de SMR important

	SMR moyenne ou insuffisante	SMR importante	Total
P(SMR=imp.)<0.50	468	26	494
P(SMR=imp.)>0.50	59	900	959
Total	527	926	1453

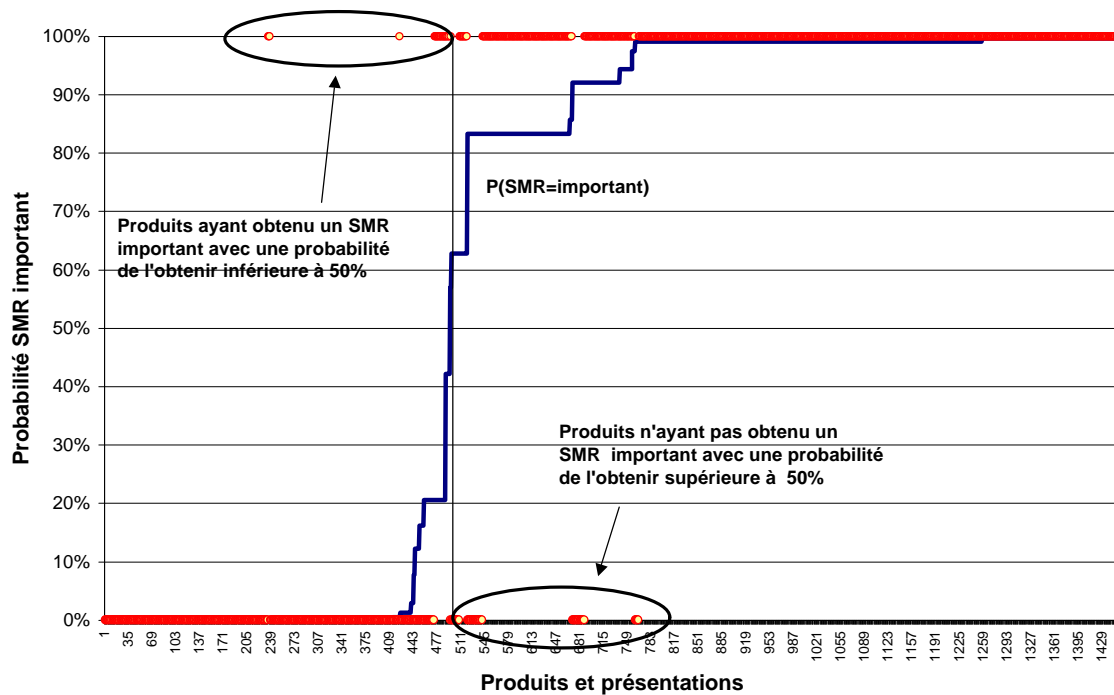
De manière générale, on note la très bonne concordance entre le niveau « prédit » et le niveau effectif. En considérant le modèle comme un « test de SMR important », sa sensibilité⁹ serait de 97%, sa spécificité¹⁰ de 89% et sa valeur prédictive¹¹ de 94%, ce qui constituent des valeurs excellentes.

⁹ Pourcentage de médicaments avec SMR important ayant été classés comme tels (900/926).

¹⁰ Pourcentage de médicament sans SMR importants ayant été classés comme tels (468/527).

¹¹ Pourcentage de médicaments avec une probabilité de SMR important >0.50, ayant effectivement ce classement.

Figure 3 : Concordance entre le niveau de SMR « prédit » et le niveau obtenu



Nota : pour chaque produit, la probabilité d'obtenir un niveau de SMR important est donnée par la courbe continue. Les points indiquent le niveau réellement obtenu avec la valeur 1 si le SMR est important et 0 sinon. Les cas de discordance sont ceux où le SMR est important alors que la probabilité de l'obtenir est faible (<50%) et, à l'inverse, ceux où le SMR n'est pas important, alors que la probabilité est élevée (>50%).

Sur les 926 produits ayant obtenu un niveau de SMR « important », 26 (2,8 %) avaient une probabilité de l'obtenir inférieure à 50%. Le tableau 12 ci-dessous résume quelques uns de ces cas de « surclassement » qui concernent essentiellement des corticoïdes dans certaines indications rhumatologiques et des phosphates ou calcium dans certains état de carences.

Tableau 12 : Cas de discordance où le SMR « prédit » est inférieur au SMR obtenu

Produits	Indications	Gravité Pathol.	Efficacité	Place Stratégique (*)	Santé Publique	Prob. SMR important	SMR obtenu
Gluconate de calcium	Calcithérapie d'urgence	C	Faible	Non	Non	<1%	Important
Phosphoneuros®	Rachilisme hypophosphorémique	C	Faible	Non	Non	<1%	Important
Hydrocortancyl®	Talalgie, maladie de Dupuytren, etc.	C	Moyenne	Oui	Non	21%	Important

(*) Oui : traitement de 1^{ère}, de 2^{ème} intention ou de recours ; non : traitement d'appoint ou sans place dans la stratégie thérapeutique.

Inversement, sur les 527 produits et présentations ayant obtenu un SMR autre qu'important, 59 (11,2%) avaient une probabilité supérieure à 50% d'obtenir un SMR important.

Tableau 13 : Cas de discordance où le SMR « prédit » est inférieur au SMR obtenu

Produits	Indications	Gravité Pathol.	Efficacité	Place Stratégie (*)	Santé Publique	Prob. SMR important	SMR obtenu
Aldomet ®	HTA sévère ou modérée	A	Moyenne	Oui	Oui	97%	Modéré
Catapressan ®	HTA	A	Moyenne	Oui	Oui	97%	Modéré
Cardioquine	Prév. récurrences tachycardies supraventricul.	A	Moyenne	Oui	Oui	97%	Modéré
Isorhythm ®	Idem	A	Importante	Oui	Oui	100%	Modéré
Moclamine ®	Ep. dépressifs majeurs, y compris sévères	A	Moyenne	Oui	Oui	97%	Modéré
Densical ® vit.D3	Traitement ostéoporose (patients carencés)	B	Moyenne	Oui	Non	63%	Modéré
Mogadon ®	Insomnie transitoire	B	Moyenne	Oui	Oui	92%	Modéré

(*) Oui : traitement de 1^{ère}, de 2^{ème} intention ou de recours ; non : traitement d'appoint ou sans place dans la stratégie thérapeutique.

Ces produits appartiennent pour l'essentiel à quelques groupes, dont le tableau 13 donne quelques exemples. On notera le cas des 5 présentations du disopyramide dans l'indication de la prévention des récurrences des tachycardies supraventriculaires, pour lesquelles la gravité de la pathologie est sévère, l'efficacité est importante, la place dans la stratégie thérapeutique et l'intérêt pour la santé publique sont reconnus, et dont le niveau de SMR est « modéré ». Le modèle « prédit » pour ces produits un SMR important avec une probabilité pratiquement égale à 1.

V. Discussion – Conclusion

L'importance prise récemment par l'évaluation post-AMM pour le remboursement et l'utilisation des médicaments justifie largement qu'on s'interroge sur la pertinence de ses méthodes. L'expérience de la réévaluation du service médical rendu effectué en France par la Commission de la Transparence constitue un cas d'école dont l'étude est facilitée par une méthodologie explicite.

L'analyse des résultats de cette opération nous conduit aux conclusions suivantes :

- Alors que la méthode prévoit, conformément à la loi, une liste exhaustive de critères d'évaluation, il apparaît qu'en réalité deux d'entre eux, l'efficacité et la gravité de la pathologie, suffisent à expliquer très largement le classement sur l'échelle du SMR. Les autres critères apportent finalement peu de valeur ajoutée.
- Parmi les critères, certains souffrent visiblement d'un manque de clarification, ce qui conduit à une interprétation variable, selon les classes, voire selon les produits ou groupes de produit. C'est, en tout premier lieu, le cas du critère « intérêt pour la santé publique » dont le pouvoir explicatif n'est pas négligeable, mais dont on peine à voir comment il a été

utilisé : dans la plupart des cas, il est redondant avec l'efficacité et/ou la gravité. Dans d'autres cas, au contraire, il s'en écarte sans qu'il se dégage une ligne d'interprétation très claire. Sans doute une réflexion complémentaire sur ce critère est-elle nécessaire. Mais c'est alors l'ensemble de l'opération qu'il faut revoir, dans la mesure où ce critère devrait être un critère synthétique, résumant les différentes dimensions de l'évaluation post-AMM.

- Un point important est la relative hétérogénéité inter-classes thérapeutiques du processus d'évaluation, au moins pour les niveaux intermédiaires de SMR. Comme nous l'avons montré, un produit d'efficacité moyenne dans une pathologie de sévérité moyenne a moins de chance d'obtenir un SMR important en cardiologie que dans les autres classes. Cela tient à un nombre important de produits que la Commission a sans doute souhaité déclasser dans cette classe. Toujours est-il qu'une procédure d'harmonisation inter-classes, voire à un niveau plus fin, devrait compléter le dispositif.
- De même, il existe des cas de discordance entre le classement « prédit » sur la base de l'analyse statistique et le classement effectif. Ces discordances sont peut-être justifiées au regard d'autres critères que ceux officiellement pris en compte. Mais il importe alors que ceux-ci soient explicités et justifiés.
- On peut également s'interroger sur l'absence de critère économique explicite. Alors que tout l'enjeu de la procédure est de nature économique, à aucun moment les questions de prix n'ont été évoquées. Le jugement sur le rapport coût-efficacité des médicaments est cependant un critère important pour juger de l'opportunité de la prise en charge. Les autorités de tutelle françaises semblent avoir du mal à considérer le coût des thérapeutiques modernes comme un élément objectif susceptible d'être apprécié scientifiquement, comme le sont l'efficacité ou la sécurité des médicaments.
- En ce qui concerne l'outil d'analyse qui a été présenté ici, les tests montrent qu'il a une très forte valeur prédictive et qu'il permet de révéler les cas d'incohérence et d'hétérogénéité. Il contribue donc à la transparence de l'opération et il devrait être systématiquement mis en œuvre dans les procédures d'évaluation post-AMM.

En ce qui concerne ce que nous avons appelé « l'évaluation post AMM » des médicaments, il apparaît clairement, comme nous l'avons montré en introduction, qu'elle est appelée à se développer, indépendamment de l'opinion que nourrit l'industrie pharmaceutique qui a tendance à la trouver inutile et redondante par rapport à l'autorisation de mise sur le marché. L'expérience française d'évaluation du SMR montre cependant qu'il est possible de jeter sur le médicament un regard différent de celui des instances d'enregistrement. Mais elle montre également les difficultés d'une telle évaluation. Si cette dernière a un sens, il doit en effet s'écarter d'une procédure technico-médicale analogue à celle utilisée dans le cadre de l'AMM pour mettre en œuvre une appréciation orientée vers la santé publique et la prise en charge collective des médicaments. Cette démarche implique par exemple que le jugement soit déconnecté de l'efficacité technique des médicaments. Des médicaments de faible efficacité répondant à un réel besoin doivent pouvoir être pris en charge au nom d'impératifs de santé publique, alors que d'autres, techniquement plus efficaces, peuvent ne pas l'être, pour les mêmes raisons.

L'importance du critère « efficacité » dans le classement montre que la Commission n'a pas complètement rompu avec l'appréciation technico-médicale.

L'opération de réévaluation a été conduite dans un contexte dominé par des considérations d'opportunité conjoncturelle : il s'agissait de justifier le déremboursement de certains médicaments sur des arguments « présentables » à l'opinion publique, notamment leur inefficacité. C'est de cette manière qu'elle a été officiellement présentée par les responsables politiques. Paradoxalement, une fois les résultats acquis, ces mêmes pouvoirs publics ont

renoncé à en tirer les conclusions qui s'imposaient et ont maintenu dans le périmètre du remboursable des médicaments à SMR insuffisant. Tout au plus ont-ils réduit les prix de ces produits, ce qui est une demi-mesure peu satisfaisante et peu logique.

Mais, indépendamment de ces aléas politiques, la question posée par l'expérience française est celle de l'élaboration d'une vraie doctrine de l'évaluation post-AMM. Cela passe par la fixation de critères d'évaluation qui procède d'une réelle réflexion méthodologique sur les critères de prise en charge des médicaments – innovants ou non – et sur leur impact sur la santé publique. La principale leçon de cette expérience, c'est qu'on en est encore loin.